

Investigating the Effect of User Variables on Perceiving the Emotional Nonverbal Behaviors of an Empathic Virtual Agent

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 Dissertation

von

Adineh Hosseinpanah
aus
Rasht (Iran)

1. Gutachterin: Prof. Dr. Nicole Krämer
 2. Gutachter: Prof. Dr. Albert Newen
- Tag der mündlichen Prüfung: 19.02.2020

DuEPublico

Duisburg-Essen Publications online

UNIVERSITÄT
DUISBURG
ESSEN

Offen im Denken

ub | universitäts
bibliothek

Diese Dissertation wird über DuEPublico, dem Dokumenten- und Publikationsserver der Universität Duisburg-Essen, zur Verfügung gestellt und liegt auch als Print-Version vor.

DOI: 10.17185/duepublico/71809

URN: urn:nbn:de:hbz:464-20200604-143356-4

Alle Rechte vorbehalten.

Acknowledgment

Completing this dissertation was a unique experience for me and it would not have been possible without the support of people whom I would like to thank deeply.

First and foremost, I would like to express my sincere gratitude to my supervisor Prof. Dr. Nicole Krämer. Thank you, Nicole, for trusting me and giving me the wonderful opportunity to undertake my Ph.D. Thank you for your continuous support, professionally and emotionally, whenever I needed. I have truly learned a lot from you, and I could not have imagined a better mentor along the way.

I would also like to thank Prof. Dr. Albert Newen, first, for presenting the international master's program *Cognitive Science* in Bochum that raised my interest in research, and second, for accepting to be the second supervisor of this dissertation. Doing my master's studies in Bochum was a life-changing experience and a big motivation to proceed with my academic career.

Many thanks to the former and current members of SP, for being such amazing colleagues. I am deeply grateful for your support, insight, encouragement, and constructive feedback. And also, to the people of LE206: thank you for being the best office mates ever.

I would also like to express my special thanks to Carolin Straßmann and Ramin Yaghoubzadeh, for all the shared knowledge and valuable insights that contributed to this dissertation. Furthermore, I would like to acknowledge Ms. Aylin Demirel, who contributed to the third study of this dissertation by conducting the study in the lab.

Besides, my deepest thanks go to my dear friends, near or 4,800 kilometers away. Thank you for the honor of being my friends. Your friendship means a lot to me and I appreciate having each one of you endlessly. Also, a very special thanks to my friend of 10 years, Ed. K. Parast, for his unwavering support and for being my private Sigmund Freud.

And to the sister I never had, Matin. You defined true friendship to me by making my life easier away from home. Thank you for always being a text message away, for all the moments that you were the ear to me, for being my telepathic partner, and for always being there for me. I appreciate your love and "I'll be there for you", too.

And for being by my side, my very special thanks go to Koen Johan. Dank je wel, Koen, for believing in me and always reminding me of what I have achieved in life. Thanks for broadening my view, for your patience, and above all, for being such a kind and wonderful human being. And, of course, thank you, for all the rides, walks, talks, and delicious meals. They all made the process of writing this dissertation way easier.

And last but not least, I would like to thank my family: Amir, you are the best big brother a girl could ask for. Thank you for your love and for always backing me up. A million thanks to my beloved father, for showing me how to be hopeful and persistent by example. Thank you for your unconditional love, encouragement, and support, baba. I could not have reached this far without you. And, finally, to my rock, my mom, the strongest woman of my life: Merci maman, for that afternoon talk, back home, for telling me not to give up on my dreams and not to lose my *hope*.

Parts of this dissertation have been presented in or accepted for talks and posters at conferences and workshops. Please refer to the following publications for more information:

Study 1

Hosseinpanah, A. Krämer, C. (July 2017). *Age-related Differences in Evaluation of Empathic Nonverbal behaviors of a Virtual Agent*. Poster presented at 3rd Workshop on Virtual Social Interaction, Bielefeld, Germany.

Hosseinpanah, A. Krämer, C. (2018, April). *Age-related Differences in Evaluation of Empathic Nonverbal behaviors of a Virtual Agent*. Poster presented at American Psychological Association's Conference on Technology, Mind & Society, Washington, DC, The United States.

Hosseinpanah, A. (2018, May). *Empathy for Everyone?: The Effect of Age When Evaluating a Virtual Agent*. Poster presented at the annual meeting of the ICA 2018 (International Communication Association), Prague, Czech Republic.

Hosseinpanah, A., Krämer, N. C., & Straßmann, C. (2018, December). *Empathy for Everyone?: The Effect of Age When Evaluating a Virtual Agent*. In *Proceedings of the 6th International Conference on Human-Agent Interaction* (pp. 184-190). ACM.

“Out beyond ideas of wrongdoing and rightdoing, there is a field.

I’ll meet you there.”

— Rumi

Abstract

The present dissertation investigated the effect of individual variables on the perception of emotional nonverbal behaviors of an empathic virtual agent. Prior studies have reported the importance of implementing empathy in virtual agents regarding maximizing the quality of human-agent interactions. As studies in human-human interactions demonstrate, nonverbal behaviors play a significant role in the transfer of meanings and intentions between the interlocutors. Therefore, the aim of this dissertation was to examine how individual variables influence the perception of the nonverbal behaviors of an agent. The nonverbal behaviors were specifically created to express the empathizing intention of the virtual agent. In this regard, three empirical studies have been conducted to investigate the effect of users' age (Study 1), users' cultural background (Study 2), and users' cultural stereotypes toward the agent (Study 3) on the perception of an empathic agent.

The results of Study 1 provide important insights into the understanding of how users of varying ages perceive an agent's emotional nonverbal responses. Older users observed to be more sensible to nonverbal behaviors of the agent since they perceived the agent as more empathic compared with younger adults. The higher perception of the older users was based on the agent's displayed nonverbal behaviors, indicating the importance of implementing emotional nonverbal behaviors specifically for this target group. Moreover, there were differences between older and younger adults with respect to perceiving specific nonverbal behaviors as empathic.

The culture of the users in Study 2 was also found to have an influence on users' perception of an agent. Comparing Iranian and German users (as the examples of relatively collectivistic and individualistic cultures) showed differences regarding users' perception of the agent's nonverbal behaviors. While Iranian participants rated a wide array of nonverbal behaviors as empathic, their German peers perceived one specific nonverbal behavior only as the most empathic one. This is where the factor of culture becomes a matter of concern. In order for a virtual agent to look more emphatic, users' culture should be considered while implementing nonverbal behaviors.

And finally, in Study 3, it was observed that users do not show stereotypical attitudes toward the agent when they are aware that the agent is designed and produced in a different country. Users who were primed about the country of the agent did not show different ratings of

empathic behaviors of the agent in the presence of emotional nonverbal behaviors. The findings of this study make valuable contributions to the research in the domain of bias toward the agent. It can be stated that as long as there are no cues in the appearance of the agent representing a different culture, users do not get biased toward its emotional responses.

Taken together, this dissertation provides insight into understanding that user variables indeed matter and they can influence accurately perceiving the emotional nonverbal behaviors of virtual agents. Since the purpose of designing an empathic agent is to enhance users' experiences and satisfaction, the factors that have the potential to impact their perception of an empathic agent should be considered when designing a virtual agent.

Zusammenfassung

Die vorliegende Dissertation untersuchte den Effekt von persönlichen Variablen auf die Wahrnehmung des emotionalen nonverbalen Verhaltens eines empathischen virtuellen Agenten. Studien zeigten bereits die Relevanz der Implementierung von Empathie in virtuellen Agenten um die Qualität von Interaktionen zwischen Menschen und Agenten zu maximieren. Wie bereits in Studien bei Mensch-zu-Mensch-Interaktionen demonstriert, spielt nonverbales Verhalten eine wichtige Rolle beim Transfer von Bedeutungen und Interaktionen zwischen den interagierenden Akteuren. Folglich, war es das Ziel dieser Dissertation zu untersuchen, wie sich persönliche Variablen auf die Wahrnehmung auswirken, wenn es um das nonverbale Verhalten eines Agenten geht. Das nonverbale Verhalten wurde explizit generiert um die emphatische Intention des Agenten auszudrücken. Es wurden drei empirische Studien durchgeführt, um zu untersuchen welchen Einfluss das Alter der Nutzenden (Studie 1), deren kulturelle Hintergrund (Studie 2), und die kulturell bedingten Stereotype dem Agenten gegenüber (Studie 3) auf die Wahrnehmung des empathischen Agenten nehmen.

Die Ergebnisse der ersten Studie liefern Erkenntnisse um zu verstehen, wie Nutzende unterschiedlichen Alters die emotionalen nonverbalen Antworten eines Agenten wahrnehmen. Nutzende höheren Alters waren sensibler in Bezug auf das nonverbale Verhalten des Agenten, im Vergleich zu jüngeren Erwachsenen nahmen sie den Agenten empathischer wahr. Diese empathischere Wahrnehmung älterer Nutzenden basierte auf dem vom Agenten gezeigten nonverbalen Verhalten. Dies kann als Indikator gesehen werden für die Wichtigkeit der Implementierung von emotionalem nonverbalem Verhalten explizit für diese Zielgruppe. Zudem, ergaben sich Unterschiede zwischen älteren und jüngeren Erwachsenen hinsichtlich der Wahrnehmung von spezifischem nonverbalem Verhalten als empathisch.

Im Rahmen der zweiten Studie erwies sich der kulturelle Hintergrund der Nutzenden ebenso als Einflussfaktor für die Wahrnehmung eines Agenten. Der Vergleich zwischen iranischen und deutschen Nutzenden (als Beispiele für eher kollektivistische und individualistische Kulturen) ergab Unterschiede hinsichtlich der Nutzerwahrnehmung des nonverbalen Verhaltens von Agenten. Während iranische Probanden ein weites Spektrum von nonverbalen Verhaltensweisen als empathisch einstufen, nahmen deren deutsche Pendanten nur ein spezifisches nonverbales Verhalten als höchst empathisch wahr. Dadurch gewinnt der

Einbezug kultureller Faktoren an Bedeutung. Um virtuelle Agenten empathischer wirken zu lassen, sollte die Kultur der Nutzenden bei der Implementierung nonverbaler Verhalten berücksichtigt werden.

Abschließend wurde in der dritten Studie gezeigt, dass Nutzende den Agenten keine stereotypischen Eigenschaften zuschreiben, wenn Ihnen bekannt ist, dass der Agent in einem anderen Land konzipiert und produziert wurde. Nutzende, die über das Ursprungsland der Agenten im Voraus informiert wurden, lieferten keine abweichenden Bewertungen des empathischen Verhaltens des Agenten bei der Anwesenheit emotionaler, nonverbaler Verhaltensweisen. Die Ergebnisse dieser Studie tragen zur Forschung im Bereich der Voreingenommenheit gegenüber Agenten bei. Es lässt sich sagen, dass solange keine Hinweise an der Agentenerscheinung vorhanden sind, die eine andere Kultur repräsentieren, auch keine Voreingenommenheit der Nutzenden gegenüber seiner emotionalen Antworten herrscht.

Zusammenfassend liefert die vorliegende Dissertation Erkenntnisse darüber, dass Variablen der Nutzenden von Bedeutung sind und dass sie die akkurate Wahrnehmung emotionaler, nonverbaler Verhaltensweisen virtueller Agenten beeinflussen können. Da das Ziel der Gestaltung empathischer Agenten die Optimierung von Nutzererleben und Zufriedenheit ist, müssen diese Faktoren mit einem potenziellen Einfluss auf die Wahrnehmung eines empathischen Agenten berücksichtigt werden, wenn virtuelle Agenten gestaltet werden.

Table of Contents

Abstract	7
Zusammenfassung	9
List of Tables.....	15
List of Figures	17
I INTRODUCTION	18
II THEORETICAL BACKGROUND.....	24
1 Media Equation Theory	24
2 The Concept of Empathy.....	26
2.1 The Role of Empathy in Human-Human Interactions.....	27
3 Theory of Mind	29
3.1 Differences Between Theory of Mind and Empathy.....	31
3.2 Theory of Mind in Late Adulthood	33
3.3 Theory of Mind in Different Cultures	36
3.4 Reflection of Individualism/Collectivism in Theory of Mind and Empathy.....	38
4 Stereotypes Toward Virtual Agents	42
5 Nonverbal Behaviors in Human–Human Interactions.....	46
6 Nonverbal Behaviors in Human–Agent Interactions.....	48
7 Empathy in Human–Agent Interactions	51
III RESEARCH APPROACH	54
IV EMPIRICAL STUDIES	56
8 Study 1: Empathy for Everyone? The Effect of Age When Evaluating a Virtual Agent.....	56
8.1 Hypotheses and Research Questions	57
8.2 Method.....	58
8.2.1 Material	59
8.2.2 Sample	64
8.2.3 Procedure.....	64
8.2.4 Measures.....	65
8.3 Results	65
8.4 Additional Analyses	70
8.5 Discussion	73

8.6	Conclusion.....	76
8.7	Limitations and Future Work	76
9	Study 2: Lost in Interpretation? The Role of Culture on Rating the Emotional Nonverbal Behaviors of a Virtual Agent	78
9.1	Hypotheses and Research Questions	78
9.2	Method.....	80
9.2.1	Material	80
9.2.2	Sample	81
9.2.3	Procedure.....	81
9.2.4	Measure	81
9.3	Results	82
9.4	Additional Analyses	87
9.5	Discussion	92
9.6	Conclusion.....	95
9.7	Limitations and Future Work	95
10	Study 3: Investigating Stereotypical Attitudes of Users Toward a Virtual Agent Designed in a Different Country	97
10.1	Hypotheses and Research Questions	98
10.2	Method.....	103
10.2.1	Material	103
10.2.2	Sample	103
10.2.3	Procedure.....	103
10.2.4	Measures.....	104
10.3	Results	106
10.4	Additional Analyses	115
10.5	Follow-up study (with Pictures).....	120
10.6	Discussion	128
10.7	Limitations and Future Research.....	134
10.8	Conclusion.....	134
	V GENERAL DISCUSSION.....	135
11	Synopsis of Empirical Results.....	135
12	Theoretical Implications.....	146
13	Practical Implications.....	149
14	Limitations and Future Directions.....	151

15 Conclusion..... 153
References155
Apendix171

List of Tables

Table 1 Emotional sentences uttered by the agent	60
Table 2 Emotional nonverbal behaviors and their duration	62
Table 3 Statistical values of the contrasts for the perceived nonverbal behaviors – independent of age (empathy)	66
Table 4 Descriptive values of the perceived nonverbal behaviors – independent of age (empathy)	67
Table 5 Statistical values of the contrasts for the older users (empathy)	67
Table 6 Statistical values of the contrast for the younger users (empathy).....	68
Table 7 Descriptive values of the perceived nonverbal behaviors for the older and younger users (empathy)	68
Table 8 Descriptive values of the perceived nonverbal behaviors for the older and younger users (friendly, intelligent, and helpful)	69
Table 9 Statistical values of the contrasts for the perceived nonverbal behaviors – independent of age (trustworthy)	71
Table 10 Descriptive values of the perceived nonverbal behaviors – independent of age (trustworthy).....	71
Table 11 Statistical values of the contrast for the older users (trustworthy).....	72
Table 12 Statistical values of the contrasts for the younger users (trustworthy)	72
Table 13 Descriptive values of the perceived nonverbal behaviors for the older and younger users (trustworthy).....	73
Table 14 Statistical values of the contrasts for the perceived nonverbal behaviors – independent of culture (empathy)	83
Table 15 Descriptive values of the perceived nonverbal behaviors - independent of culture (empathy)	84
Table 16 Statistical values of the contrasts for the Iranian users (empathy).....	84
Table 17 Statistical values of the contrasts for the German users (empathy)	85
Table 18 Descriptive values of the perceived nonverbal behaviors for the Iranian and German users (empathy)	85
Table 19 Statistical values of the contrasts for the perceived nonverbal behaviors – independent of culture (trustworthy)	87
Table 20 Descriptive values of the perceived nonverbal behaviors – independent of culture (trustworthy).....	88

Table 21 Statistical values of the contrasts for the perceived nonverbal behaviors – independent of culture (helpful)	88
Table 22 Descriptive values of the perceived nonverbal behaviors – independent of culture (helpful).....	89
Table 23 Statistical values of the contrasts for the Iranian users (helpful)	90
Table 24 Statistical values of the contrasts for the German users (helpful).....	90
Table 25 Descriptive values of the perceived nonverbal behaviors for the Iranian and German users (helpful).....	91
Table 26 Statistical values of the contrasts for the perceive nonverbal behaviors – both primed groups (empathy).....	108
Table 27 Moderated regression analysis for item empathy (positive affect)	110
Table 28 Moderated regression analysis for item empathy (negative affect)	111
Table 29 Moderated regression analysis for item empathy (experience with technology).....	111
Table 30 Moderated regression analysis for item empathy (level of empathy)	112
Table 31 Perception of positive emotional responses regarding the happy nonverbal behaviors	114
Table 32 Perception of negative emotional responses regarding the sad nonverbal behaviors ...	114
Table 33 Perception of the neutral responses of the agent	115
Table 34 Statistical values of the contrasts for the perceived nonverbal behaviors – both primed groups (unfriendly and reactive)	116
Table 35 Statistical values of the contrasts for the perceived nonverbal behaviors – both primed groups (warmhearted/cold, incompetent/competent, and inactive/active).....	118
Table 36 Statistical values of the contrasts for the perceived nonverbal behaviors – both primed groups (appropriateness)	120
Table 37 Moderated regression analysis for item empathy in the follow-up study (positive affect)	125
Table 38 Moderated regression analysis for item empathy in the follow-up study (negative affect)	125
Table 39 Moderated regression analysis for item empathy in the follow-up study (experience with technology).....	126
Table 40 Moderated regression analysis for item empathy in the follow-up study (level of empathy).....	126

List of Figures

Figure 1: Key-frame editor and the list of body parts which can be manipulated	61
Figure 2: Examples of Billie showing emotional nonverbal behaviors from left to right: Initial position (neutral), Smile, Sad Face, Head Down, Dropping the Arms plus Sad Face	62
Figure 3: Assignment of Nonverbal Behaviors to the Happy Situations	63
Figure 4: Assignment of Nonverbal Behaviors to the Sad Situation	63
Figure 5: Descriptive values of rating the item empathy for older and younger users	66
Figure 6: Descriptive values of rating the item trustworthy for older and younger users.....	70
Figure 7: Descriptive values of rating the item empathy for Iranian and German users	83
Figure 8: Descriptive values of rating the items trustworthy and helpful for Iranian and German users.....	86
Figure 9: Descriptive values of the four primed groups' ratings (empathy, unfriendly, and reactive).....	122
Figure 10: Descriptive values of the four primed groups' ratings (warmhearted/cold, incompetent/competent, and inactive/active).....	123

I INTRODUCTION

Humans are social beings, surrounded by the company of others, and naturally oriented toward them. The ability to understand the intentions and motivations of others and to predict their behavior is crucial and necessary for social interactions. Having such ability can reduce conflicts and misunderstandings, thereby enabling human beings to survive in complex social environments (Paiva, Leite, Boukricha, & Wachsmuth, 2017). Living in the 21st century makes encountering smart technologies inevitable. The range of smart technologies is wide, from smart home appliances, smartphones, and smart cars to robots and embodied conversational agents. Although in the past some of these technologies, such as robots and embodied agents, were encountered by humans only when watching movies or industrial robots in factories, they have become more involved in their everyday private lives in recent years (Eimler, Krämer, & Von Der Pütten, 2011). This makes it important to evaluate how users respond to these technologies. The purpose of all these technologies is one thing: to enhance the quality of humans' life in a maintainable way. One way to achieve this goal is to enable these technologies to interact with humans in a natural and social way, inspired by the way that humans interact with each other in human–human interactions (Paiva et al., 2017). However, one of the important issues to consider is to ensure that human users understand and perceive the intention of these technologies accurately, in order to maximize their satisfaction and quality of their experience.

With the rapid development of technology, there is an increasing possibility that companion virtual agents and robots will be part of humans' daily life, in the form of, for instance, artificial assistants. Therefore, as Beer and colleagues (2015) suggest, having insights into human–human social interactions is no longer sufficient, as these interactions can occur between humans and agents, robots, and other artificial entities as well. These companion assistants are stated to be capable of entering home environments and health-care settings, for both older people and younger adults (Beer, Smarr, Fisk, & Rogers, 2015).

Embodied conversational agents are considered as an important form of intelligent user interfaces. They are defined as “animated anthropomorphic interface agents that are able to engage a user in real-time, multimodal dialogue, using speech, gesture, gaze, posture, intonation, and other verbal and nonverbal behaviors to emulate the experience of human face-to-face interaction” (Cassell, Sullivan, Churchill, & Prevost, 2000, see Bickmore, & Cassell, 2005, p. 3).

Moreover, agents can have abilities such as being autonomous, proactive, reactive, and social, and the social ability of the agents can be defined either with human users or other agents (Wooldridge & Jennings, 1995). In this dissertation, the focus is on embodied conversational agents as potential social companions for users. Companion virtual agents are defined as personalized, helpful, collaborative, conversational, and emotional computer agents that have the potential to know their owners and form a long-term relationship with them (Danilava, Busemann, & Schommer, 2012).

Companion agents have been used in many research projects (Kim, Baylor, & Shen, 2007; Prendinger & Ishizuka, 2005; Rieser & Lemon, 2011; Yaghoubzadeh, Kramer, Pitsch, & Kopp, 2013). One of the main interests and goals of research in the field of human-agent interaction is to ensure that first of all the agents are accepted by the users, and then to elevate the quality of such interactions by implementing different factors and social skills in the design process. This, in turn, can guarantee the technology is reused in the future. To this end, a certain level of believability and trustworthiness on the part of the agent is required. One of the factors that can contribute to reaching such believability and trustworthiness is implementing emotions, namely, empathy in the agent. Empathy is a fundamental component of social interactions. It enables people to respond to others via psychological processes that make people have congruent feelings about the situations of others (Hoffman, 2001). Virtual agents that are capable of showing empathy can initiate and maintain a social interaction between users and agents (Yalçın, 2018). This is of great importance for maintaining the agents' companionship with their users for a longer time and securing further usage of the agent in the future.

To date, many studies have investigated the crucial role of empathic agents and robots in enhancing the quality of interactions with human users. For example, it has been reported that empathic agents lead to users' positive rating in the sense of trustworthiness, likeability, and caring (Brave, Nass, & Hutchinson, 2005), reducing negative feelings and the stress level of users (Prendinger & Ishizuka, 2005; Prendinger, Mori, & Ishizuka, 2005b), lowering frustration in users (Hone, 2006; Klein, Moon, & Picard, 2002), contributing to a higher perception of friendliness (Mascarenhas, Martinho, Prada, Paiva, & Leite, 2013), as well as forming a sustainable long-term interaction (Leite, Castellano, Pereira, Martinho, & Paiva, 2014). All of these studies emphasize the positive influence of implementing empathy in virtual agents to foster the efficacy of the agent and the quality of interaction between users and the agent.

In order to make the empathic virtual agent more believable and natural, the use of nonverbal behaviors can be beneficial to express the feeling of empathy in a more obvious but, nevertheless, subtle way. Nonverbal behaviors are reported to play a significant role in regulating the flow of the conversation and conveying communicative and social cues such as attentiveness, positive affect, and liking (Bickmore & Cassell, 2005). However, far too little attention has been paid to the question of which specific emotional nonverbal behavior can lead to a higher perception of empathy of an agent.

A companion agent that uses emotional nonverbal behaviors to express empathy requires users to accurately interpret the meaning of these nonverbal behaviors, and ultimately the intention of the agent that is “empathizing.” It is thus of great importance to have an understanding of users’ perception and recognition of the empathic behaviors of the agent, especially the displayed emotional nonverbal behaviors. The perceived social abilities of the agent, such as empathizing via nonverbal behaviors, can influence users’ satisfaction and acceptance of the agent. The ultimate purpose of implementing social abilities in the agents is assuring that users have more qualified human–agent interactions and greater willingness to use the agent for a longer period. Hence, it is very important to make certain that they truly understand the intention and behavior of the agent.

In the domain of human–agent interaction, it has always been assumed that the agents can communicate with all kinds of users because these agents are able to generate and understand verbal and nonverbal behaviors. Therefore, the necessity of paying attention to specific user groups is considered to be less important (Krämer, Hoffmann, & Kopp, 2010). However, as humans indeed display inter-individual differences, they might have different understandings of the agent’s (nonverbal) behaviors, especially when it concerns emotional behaviors. It is, therefore, very important to investigate whether inter-individual differences of user groups (e.g., with regard to their age and culture) influence their perception of emotional nonverbal behaviors as empathic. Also, it is crucial to examine whether users prefer to see certain types of emotional nonverbal behaviors when interacting with a companion agent.

For instance, companion agents, specifically virtual assistants, are designed and expected to be interactive partners of users of varying ages. It has, however, been reported that the group that can benefit the most from such intelligent assistants is the group of older people (Beer et al.,

2015). The intention of using virtual assistants for this group of people is due to the fact that since the past decade, the number of older people is increasing (Hubbard, Cook, Tester, & Downs, 2002; Yaghoubzadeh et al., 2013). Moreover, older people have difficulties in realizing their daily tasks, such as household activities and daily routines, for instance, remembering appointments, taking medications on time, meal-times, etc. (Yaghoubzadeh et al., 2013). It is, therefore, an aim to maintain the quality of life and to support the personal and societal needs of this group of people. Moreover, as demographic investigations predict, the number of human caretakers who can offer support to older people is decreasing (Yaghoubzadeh et al., 2013). Virtual assistants, therefore, can offer support to older people in their home environment, making it possible for them to remain autonomous for a longer time.

In this regard, many studies have been conducted in the domain of human-agent/robot interactions that have considered older adults as their user group. Most of these studies have focused on the willingness of having these assistants at home (Ezer, Fisk, & Rogers, 2009) and on a design-centered investigation of their dimension of usability and acceptability (Yaghoubzadeh et al., 2013). However, there is an absence of research exploring the effect of age on the perception of empathic behaviors, namely, empathic nonverbal behaviors of the agent. The aging process brings about changes in physical, cognitive, and perceptual abilities of people. Therefore, it is an imperative research topic to investigate how older people perceive an empathic agent, and whether they are able to accurately perceive its empathic behavior through its displayed nonverbal behaviors.

Furthermore, as mentioned earlier, virtual assistants are intended to interact with different target groups. Besides age, another factor that can have an influence on understanding empathic nonverbal behaviors is the culture of people. Whether people have individualistic or collectivistic values can potentially affect their ability of theory of mind – the ability to understand the mental states of others (Premack & Woodruff, 1978). This, in turn, can influence their perception and expectation of an interactant's emotional response (i.e., empathic response). Although there have been several attempts to consider the factor of culture in studies (Aylett et al., 2009; Vannini et al., 2011), most of them have focused on the culture of the agent and not on the effect of the users' cultural background on perceiving the empathic agent. Since the implementation of empathy is reported to be effective in qualifying the virtual agent, it is important to ensure that users are able to perceive the emotional responses of the agent. This is specifically important in

the case of emotional nonverbal behaviors since they might have different meanings in various cultures. Accordingly, people with different cultural backgrounds might evaluate an agent's nonverbal behavior differently.

And, finally, speaking of the possible role of culture in the perception of an empathic agent, there is a notable lack of studies investigating the bias of users toward the culture of the agent. More specifically, the bias toward the place in which the agent is designed only, with no visual cues indicating the ethnicity of the agent, has not been investigated so far. It has been reported that people tend to apply the social rules and norms of human-human interactions to human-agent interactions (presented as the concept of media equation theory; see Reeves & Nass, 1996). Stereotypes toward the ethnicity of others are one of the social norms that people might apply to their interaction with the agent, which consequently can affect the acceptance of the agent. Virtual agents are designed and produced in different countries to be used by people worldwide. Therefore, having knowledge about potential stereotypical attitudes of users toward agents could be useful in establishing solutions to avoid or reduce such attitudes by applying certain factors when designing virtual agents. These factors can include, for example, reflecting on the agents' neutral look or designing more culture-specific-looking agents. More importantly, studies that explore the effectiveness of certain emotional nonverbal behaviors as empathic are scarce. Therefore, one of the main objectives of this dissertation was to investigate how the aforementioned factors affect users' perception of specific nonverbal behaviors as more or less empathic.

Accordingly, this dissertation had the following purposes: (1) to examine the effect of theory of mind (which can be different in late adulthood and in different cultures) on people's perception of an emotional agent, (2) to investigate whether the assumptions of media equation theory apply to stereotypes toward the agent based on knowledge of the culture of the agent. Specifically, this dissertation aimed to test the influence of inter-individual differences of the users on the perception of emotional nonverbal behaviors of the agent (and rating the agent as empathic) regarding the following factors: the effects of users' age-related differences (Study 1), users' cultural differences (Study 2), and users' bias (more specifically, potential stereotype) toward the culture of the agent (Study 3). The findings of these studies add to a growing body of literature on human communication, by expanding theory of mind to human-agent interaction. Also, it contributes to the literature by showing whether people apply the same social rules that

they use in human–human communication to their communication with a virtual entity as suggested by media equation theory. The findings presented here can also contribute to a design guideline. Designers can consider appropriate user models and criteria for categorizing users according to their age, cultural background, and potential stereotypes they might display toward the agent.

The overall structure of this dissertation takes the form of 15 chapters. This work first addresses the theoretical background in Section II by explaining media equation theory in Chapter 1. The definition of the concept of empathy and its role in human–human interactions are presented in Chapters 2 and 2.1. Chapter 3 presents the definition of the theory of mind, its implications in social interactions, and the role it plays in the context of this dissertation. This chapter is followed by an explanation of the differences between the concepts of theory of mind and empathy in Chapter 3.1, the effect of theory of mind in late adulthood in Chapter 3.2, and how theory of mind is developed in different cultures in Chapter 3.3. Then, the dimensions of collectivism/individualism, their influence on the function of theory of mind, and how they can influence the perception of an emotional agent are discussed in Chapter 3.4. Chapter 4 begins by laying out the theoretical dimensions of the research related to stereotypes toward the agent and gives an overview of earlier work in this domain. Chapter 5 and Chapter 6 provide an overview of the research that has been conducted to date with regard to investigating the role of nonverbal behaviors in human–human and human–agent interactions. Then, they are followed by Chapter 7 that gives an overview of the studies that implemented empathy in virtual agents. Subsequently, the research approach is presented in Section III, which explains the purpose of the studies conducted for this dissertation and focuses on the research gap in the domain of human–agent interactions with regard to the following three factors: users’ age, culture, and stereotypes toward the agent. Chapters 8, 9, and 10 then present the three empirical studies that were conducted to test the hypotheses and answer the research questions of the dissertation. A general discussion is provided in Section IV that presents the synopsis of the empirical results with a reflection on the findings in Chapters 11. Chapters 12 and 13 then present theoretical and practical implications of the findings. And finally, limitations and suggestions for future studies and the conclusion are presented in Chapters 14 and 15.

II THEORETICAL BACKGROUND

1 Media Equation Theory

In the domain of research related to human–computer interactions, the important role of Clifford Nass and Byron Reeves cannot be overlooked. They developed media equation theory, claiming that our response to media, such as media technologies, smartphones, computers, robots, virtual agents, etc., is similar to our response when we communicate with real human beings (Reeves & Nass, 1996). Reeves and Nass (1996) borrowed social psychology approaches in human–human interaction, applied them to human–computer interaction, and examined whether the same social reactions exist: “Our strategy for learning about media was to go to the social science section of the library, find theories and experiments about human-human interaction –... Take out a pen, cross out ‘human’ or ‘environment’ and substitute media. When we did this, all of the predictions and experiments led to the media equation” (p. 251).

Therefore, the core concept of media equation theory revolves around the idea that computers are social animals (CASA), meaning that humans apply the same social rules of human–human interactions to human–computer interactions (Reeves & Nass, 1996). There are several studies that support this theory by providing findings that people are inclined to rely on social categories in their interactions with computers. It has been reported that what actually leads to triggering social reactions is the presence of anthropomorphic cues that have an influence on classifying humans and computers in one category (Kim, Youjeong, & Sundar, 2012). For example, social cues such as voice input can generate gender-based stereotypes (Nass, Moon, & Green, 1997) or features such as information being revealed by computers can lead to self-disclosure behaviors of the users (Moon, 2000). Other cues including the agent’s gender (Lee, 2003) and personality (Nass & Lee, 2001) can also create a human-like perception of computers and, consequently, lead to social reactions toward them.

There are also several other studies that confirmed media equation theory and social reactions toward technologies, such as studies related to emotions (Newhagen & Reeves, 1992) and to politeness and reciprocity (Nass & Moon, 2000; Nass, Moon, & Carney, 1999). Social reactions have been observed in encounters with an embodied conversational agent as well. For

example, it has been suggested that during a task performance, participants felt more anxious and showed decreased performance when the animated agent was monitoring their work (Rickenberg & Reeves, 2000). Moreover, in an explorative study of the cognitive and behavioral effects of a virtual assistant that was designed for a TV/VCR, Krämer (2005) found that in the presence of the virtual program guide, participants selected a more socially desirable TV program, such as a documentary versus an entertainment program. Comparing different kinds of interfaces such as a graphical user interface, speech output, and embodied interface agents, it was demonstrated that when the users interact with an embodied interface agent, they use more natural speech in addressing the system and a more human-like attitude and behavior toward it (Krämer, 2005).

According to several studies, even subtle social cues such as the visual appearance of the agents or subtle nonverbal behavior can trigger the social reaction. For instance, it has been shown that the use of nonverbal signals in avatars leads to a social reaction by creating a sense of social presence (Becker & Mark, 2002; Reeves & Nass, 1996). Moreover, the effect of implementing emotions in the design of a life-like character has been shown in the acceptance of the conversational agent as co-equal partners (Becker, Kopp, & Wachsmuth, 2007). This effect has also been reported to be influential in facilitating and improving learning procedure by creating a social reaction between students and emotionally animated pedagogical agents (Jaques & Viccari, 2005).

Ethnicity is also an important social category that is predicted by CASA to affect people's attitudes toward agents from a similar or different ethnicity. Studies reported that in the presence of ethnicity cues, people showed a series of stereotypes including expectations, assumptions, and attributions toward agents (Nass & Moon, 2000). For example, it has been shown that agents from similar ethnicities were perceived to be more attractive, trustworthy, and intelligent in comparison with agents from a different ethnicity (Nass & Moon, 2000).

According to the assumptions of media equation theory, people tend to form a type of social relationship with virtual entities, especially when they represent natural and human-like cues such as nonverbal behaviors. In Study 3 of this dissertation, it was attempted to test whether users' perception of empathic nonverbal behaviors of the agent varies when they are primed that the agent is from a different culture. More specifically, the purpose was to investigate whether people apply stereotypical attitudes toward an agent from a different culture, not based on visual

cues representing a specific culture, but due to having knowledge that the agent was designed in a different country. First of all, however, it is important to discuss the concept of empathy and the role it plays in social interactions, why it is important to be implemented in virtual agents, and why it is important that users accurately understand the empathic behaviors of the agent.

2 The Concept of Empathy

It is the nature of human beings to be social and to have a tendency to act toward or in response to other beings (Batson, 1990). This social tendency leads to the need to bond with others (Davis, 2018), which is why humans instinctively care about others and are keen to be cared about (Brave et al., 2005). The evolutionary point of having this need cannot be overlooked: It has been helping humans to gather together to hunt for food, to detect and escape from the predators, to court, etc. (Plutchik, 1987). Even the organization of the neural activity of humans has been shaped by the need to assess the intention and the motivation of others (Brothers, 1989).

Such “caring” in human beings can be observed through the concept of empathy (Brave et al., 2005). The word *empathy* originated from the German word *Einfühlung*, which was used by the psychologist Theodore Lipps in the 1900s to refer to interpersonal knowledge and to describe the emotional appreciation of the other people’s feeling (see Brothers, 1989). Further, E.B. Titchener coined the new word *empathy* from Greek roots to translate the word *Einfühlung* (see Brothers, 1989). Generally, this concept is defined as humans’ capacity with respect to understanding and responding to the emotional status of others (Batson, Fultz, & Schoenrade, 1987; Decety & Jackson, 2004). This understanding happens through a sense of similarity between what one person experiences as emotions and the emotional experiences expressed by others (Lamm, Batson, & Decety, 2007).

The definition of empathy is usually categorized into different components. The first one is cognitive empathy, which is concerned with the ability to understand the feelings of others through “mind reading” and inferring others’ thoughts and feelings (Ickes, 1993). It is also reported that cognitive empathy is related to having the ability of the “theory of mind” (Blair, 2005, see Chapter 3). The second component is affective empathy, which is related to the

emotional response to the affective situations of others (Hoffman, 1994). Although there is a disagreement about the precise definition of the notion of empathy, there is a general agreement on the existence of these components.

It is important to note that when referring to empathy, it should be carefully differentiated from another concept called *sympathy*. Sympathy is defined as an immediate and uncontrolled emotional reaction toward an emotional situation, inundating humans when they envision themselves in the position of other people (Ioannidou & Konstantikaki, 2008). Owing to such inundation, acting ethically and appropriately and caring about others can be suspended (Ioannidou & Konstantikaki, 2008). On the contrary, empathy is a learned skill and life attitude, which is regularly used to understand others' feelings and emotional experiences in order to communicate with people (Halpern, 2003). Moreover, depending on whether people feel more or less responsible for others, they can consider different levels of empathizing (Ickes, 1993). Also, it has been indicated that empathy, contrary to sympathy, can include both negative and positive emotional valence (Brave et al., 2005). Empathy can be expressed in the form of sorrow, joy, confusion, excitement, and by means of verbal and nonverbal cues such as simple words, head nodding, etc. (Ioannidou & Konstantikaki, 2008).

2.1 The Role of Empathy in Human-Human Interactions

In human communication, the factors that might facilitate having social relationships and connections with others can be considered as beneficial. Empathy, in this sense, has both individual and societal advantages. It can contribute to bonding and creating social connections through prosocial behaviors (Eisenberg, Spinrad, & Sadovsky, 2006) by responding appropriately in emotional situations. To have the ability to empathize with others can lead to effective and fruitful communication with others by developing ideas and solutions, problem-solving, and preventing conflicts (Ioannidou & Konstantikaki, 2008).

Empathy is a concept that has received much attention in research, ranging from philosophy and psychology to cognitive science and cognitive neuroscience. Most of these studies have emphasized the crucial role empathy plays in human communication. For example, the important impact of empathy in the therapeutic domain has been emphasized in several studies. It has been shown that it maximizes the quality of the physician–patient interaction (Di

Blasi, Harkness, Ernst, Georgiou, & Kleijnen, 2001; Finset & Ørnes, 2017). Empathy has also been reported to be helpful in the diagnosis, treatment, and recovery procedure by making the patients less stressed regarding their symptoms, increasing their self-efficacy, and, thus, helping them be more autonomous and involved in therapy (Halpern, 2003). Also, with regard to the basic caring service, when nurses respond empathically, patients show less anxiety, depression, and hostility (Madea, 1987).

The importance of having the ability to empathize with others has been reflected in other studies outside the therapeutic context as well. For instance, the impact of the lack of empathy has been shown in motivating aggression (Eisenberg et al., 2006; Miller & Eisenberg, 1988) and antisocial behaviors (Eysenck, 1981). On the contrary, being capable of empathy has been found to be associated with prosocial behaviors, such as the helpfulness of children at school (Litvack-Miller, McDougall, & Romney, 1997). Further, some cultural studies have shown that empathy increases sensitivity to cultural differences. In this regard, it has been reported that empathic people take the perspective of other cultures and respond to people with a different cultural background from their perspective (Germain, 1998). For example, at a school and in a learning environment, it has been shown that empathy in teacher–student relationships can manifest in a form of caring, which in return helps the students to show better academic motivation and performance (Goleman, 1998; Oliner & Oliner, 1995). Studies have also shown the important effect of empathic behavior, both cognitive and affective, in reducing bully behaviors at schools (Caravita, DiBlasio, & Salmivalli, 2010; Gini, Albierto, Beneli, & Altoe, 2007).

As mentioned earlier, empathy can lead to an effective and fruitful interpersonal relationship. One of the crucial issues in any form of relationship is the trust between the interlocutors. The issue of trust is specifically important in the context of this dissertation, as researchers have stated that there is a close relationship between empathic behavior and generation of trust (Comfort, 1984; Ickes, Stinson, Bissonnette, & Garcia, 1990).

All the studies that have been mentioned so far describe the role of empathy for a qualified and trustworthy interaction as prominent. Inspired by research in human–human interactions, many researchers in the domain of human–agent interactions applied empathy and implemented it in embodied conversational agents, to engender a fruitful interaction (see Chapter 7). However, considering the potential complication of human–agent interactions, it is crucial to

determine whether users' have difficulties with regard to understanding empathic responses, specifically when the agent is trying to be empathic via displaying emotional nonverbal behaviors. Understanding emotional nonverbal behaviors is important as their role in enhancing the transfer of intentions and meaning cannot be underrated. Moreover, with regard to the concept of empathy, nonverbal behaviors can contribute to facilitating the perception and judgment of an empathic response of others (Tepper & Haase, 1978).

Inter-individual differences of the users might influence their perceptions, understandings, preferences, and expectations of empathic nonverbal behaviors. This, in turn, can affect their satisfaction regarding interactions with an empathic agent. But before mentioning the related scientific work, it is important to point out the crucial ability of human beings that enable them to empathize with others and more importantly, in the context of this dissertation, to understand the empathic behaviors of other people: theory of mind. Therefore, in the following chapters, this ability and factors that have an influence on its development and function will be discussed.

3 Theory of Mind

Understanding other people is a key factor in successful and efficient interactions. In this regard, Grice (1957) stated that in order to successfully understand the utterance of others we must be able to perceive their intention. A prerequisite for such a perception and also for interpreting what others mean has been reported to be theory of mind (Krämer, Eimler, Von Der Pütten, & Payr, 2011). The term *theory of mind* (ToM) was first introduced by Premack and Woodruff (1978). It refers to the ability to see other people as intentional beings, to attribute mental states to them, and to understand that their internal mental states, such as desires, beliefs, and emotions, affect their behaviors. It has been demonstrated that this ability might also be present in adult chimpanzees besides human (Premack & Woodruff, 1978), it exists in the developmental process of children by approximately the age of 4 years (Wimmer & Perner, 1983), and is delayed or impaired in autistic children (Baron-Cohen, Leslie, & Frith, 1985). It is important to note that although researchers have used different terms to explain and study ToM, such as *intentional states* (Dennett, 1988), *mind reading* (Baron-Cohen, 1997), and *mentalizing* (Frith & Frith, 2003), they all refer to the same concept and skill.

ToM helps explain the behaviors of others and predict their behavior in similar situations in the future (Dennett, 1988). More precisely, it has been stated that ToM enables humans to see the meanings and sense in the behaviors of others, which ultimately prevents the confusion and the overwhelming feelings that result from the complex tasks of everyday life (Kraml & Michlmayr, 2002). Further, predicting how other beings would behave and act in similar situations is necessary to be able to engage in useful communication with others.

ToM has been reported to be a natural ability of human beings that is important for their social life and their understanding of surrounding social environments (Tooby & Cosmides, 1995). Accordingly, Tooby and Cosmides (1995) emphasized its evolutionary role, too. They stated that humans as intensely social, cooperative, and competitive beings evolved this ability because their life was dependent on how well they were capable of inferring what was on others' mind. The crucial evolutionary point of having the skill of ToM has also been stated by Baron-Cohen (1997):

Imagine that you are an early hominid, and that another early hominid offers to groom you and your mate. You need to reason quickly about whether you should let him approach. ... Making inferences about whether his motives are purely altruistic or whether he might be deceitful is a reasoning strategy that you can apply in time to react to a social threat. (p. 25).

Although it has been said that ToM cannot lead to a total and precise understanding and prediction of the behaviors of others, it can offer a general orientation on humans' ability to predict the factors that influence a communication (Krämer, 2008). Moreover, it is important to note that there are different explanations regarding the basis and origin of ToM and how it functions. Such different explanations have led to a distinction between two categories of ToM: theory theory and simulation theory. Having knowledge about these categorizations can contribute to a better understanding of the core concept of ToM, the factors that might affect development of this skill and, consequently, understanding the emotions of others.

According to theory theory, humans possess a folk psychology ability that empowers them to deal with their everyday life activities (Kraml, & Michlmayr, 2002). What theory theorists claim is that the basis of ToM is either natural and intrinsic or is learned and acquired

through experience and by enculturation processes (Churchland, 1989; Gopnik 1993; Wellman 1990; see Carruthers & Smith, 1996).

What simulation theorists, on the other hand, suggest is that folk psychology does not have a role in understanding others. Accordingly, humans use their own mental states and project themselves onto the position of others – in the form of simulating others – to understand and predict their behaviors (Carruthers & Smith, 1996). Although there exist several views regarding simulation theory, they all agree on one common fact, that simulation acts as a device to shape humans' prediction and explanation of the behaviors of others (Kraml & Michlmayr, 2002). Therefore, it has been stated that simulation theory is attributor dependent, given that it depends on humans' own mental apparatus (Kraml & Michlmayr, 2002).

Apart from the disagreement regarding the grounds of ToM, there are also researchers (Heal, 1986; Perner, 1996) who do not see these views as exclusive and suggest a hybrid theory and simulation theory view. Accordingly, it is suggested that all of these different approaches have one thing in common. They all agree that human beings have direct but implicit knowledge regarding other humans, and that such knowledge – which leads to a starting point of mutual understanding – can be achieved via innate or learned theory or via simulation of the mental states of others (Krämer et al., 2011).

In the context of this dissertation, an integrated theory is of great importance. Accordingly, it considers different aspects of ToM, such as the ability to attribute the mental states of others to oneself via a simulation process, and the factors that have an influence on the acquisition and the development of this ability via learning and enculturing processes. Understanding ToM and the factors that influence its function can contribute to having in-depth knowledge about people's ability to understand the empathic behaviors of others, which is specifically important in this dissertation. However, it is essential to first distinguish between the two terms *ToM* and *empathy*.

3.1 Differences Between Theory of Mind and Empathy

ToM and empathy are two very important abilities for social understanding and processing of social cognition (Bosacki & Astington, 1999; Chandler, 1987; Liddle & Nettle,

2006, see McInnis, 2014; Völlm et al., 2006). These two abilities are usually reported to be combined and sometimes assumed to be the same concept. This assumption is based on the general notion that there is a correlation between developing ToM and people's increased level of empathizing (Batson et al., 2003; Eisenberg, Fabes, & Spinrad, 2006; Eisenberg et al., 1991; Feshbach, 1978, see McInnis, 2014). Also, it has been reported that these abilities need similar prerequisites in order to be developed, such as the capacity to be able to distinguish between self and another agency (Decety & Grezes, 2006, see McInnis, 2014). Although there are overlapping constructs in the definition of ToM and empathy, there are differences that need to be taken into account.

ToM is an umbrella term (Hynes, Baird, & Grafton, 2006). It mainly refers to people's ability to understand others' mental states and in a broader sense, it includes people's ability to take on the perspective of others. Empathy is related to both identifying the emotions of others, which is defined as cognitive empathy, and sharing and matching one's own emotions to the emotions of others, defined as affective empathy (McInnis, 2014). Empathy has been explained in terms of contamination. It is a tendency for the emotions of others to spread to the witness as if the witness were being contaminated by the feelings of other people (Hynes et al., 2006). It has been demonstrated that ToM comprises cognitive empathy but not affective empathy, meaning that people can have the ability of ToM without having the ability of affective empathy (McInnis, 2014). Therefore, "theory of mind is a necessary, but not sufficient, component of empathy as a whole" (McInnis, 2014, p. 2).

Understanding empathy and ToM is important since empathy is a crucial societal skill, and accurately recognizing that other people are empathizing can contribute to a more qualified social interaction. Accordingly, it is important to consider this concept in human-agent interactions, by enabling the agent to show empathic responses to the users. However, it is even more crucial to make sure that users are capable of understanding these responses (by considering their inter-individual differences), in order to maximize the advantage of implementing empathy in the agent.

As mentioned earlier, there are factors that are capable of influencing the development and function of ToM. Accordingly, the effect of two factors such as users' age (Study 1) and users' culture (Study 2) are taken into account in this dissertation. The purpose was to investigate

whether these factors can influence the perception of emotional others, more specifically, emotional nonverbal behaviors of a virtual agent. Accordingly, it was of interest to test their impact on facilitating the realization that the agent is empathizing with them. The first factor that will be discussed in the following chapter is the effect of aging on understanding the emotions of other people in human–human interactions.

3.2 Theory of Mind in Late Adulthood

The key role of empathy in different aspects of human interaction was discussed in Chapter 3. It was also mentioned that ToM and empathy go hand in hand, and it is necessary to have ToM in order to empathize with others and to understand the empathic responses of other people as well. However, it is essential to investigate how factors such as the surrounding environment or aging affect the capacity of ToM. The potential influence of these factors on ToM can consequently affect how people understand the emotions of others, namely, empathic nonverbal behaviors of others in this context.

To have knowledge of how ToM works in late adulthood is of great importance. Difficulties in inferring mental states of others and comprehending their intentions, emotions, and beliefs can result in fewer communication skills and, therefore, less qualified social life for older adults. In this regard, it has been shown that age-related difficulties concerning ToM abilities diminish the social participation of people significantly, which is linked to loneliness and health issues in older people (Bailey, Henry, & Von Hippel, 2008).

To date, many attempts have been made to examine whether there are age differences in mental state decoding. Although many of these studies have offered useful insights, there are still contradictory findings regarding how aging indeed affects ToM. Several studies have declared that there are no changes in ToM that accompany adults in later lifespan development. Most of these researches argue that “a lifetime of accumulation of knowledge about the social world may lead to us becoming much more efficient in our social interactions, and as a result, certain aspects of social cognition may comprise a domain in which experience trumps youth” (Moran, 2013, p. 33).

Accordingly, two approaches appeared to explain the effect of age on understanding the emotions of others: *sociocognitive* approach and *neuropsychological* approach. According to the sociocognitive approach, as people get older, they have an enhanced and more stabilized understanding of the emotions of other people (Phillips, MacLean, & Allen, 2002). This is the result of the extensive experiences that older people achieve across the lifespan, leading them to have a better understanding of emotional cues. It has been suggested by Blanchard-Fields (2007) that due to these accumulated experiences, older people are able to better anticipate the emotional responses of others in interpersonal relationships and consequently, understand them in a better way. The findings from general cognition are in line with this view, indicating that older people improve the crystallized aspect of intelligence, namely, general knowledge, vocabulary, and verbal memory, which in turn assists them to master the tasks related to ToM (Moran, 2013).

Many studies have examined the effect of aging on the ability of ToM. For example, Happé, Winner, and Brownell (1998) showed no significant age-related decline in tasks related to ToM. In their study, older people and younger adults had to read stories taken from the Strange Stories test (Happé, 1994). The Strange Stories test was originally designed for the purpose of investigating ToM deficits in autistic participants. These autistic people were, in fact, successful in some tasks of ToM such as false belief – the ability to understand that others have beliefs that can be different from the reality (Perner, 1991).

In these stories, and in the context of everyday situations, people say things that are not meant literally. The task of the participants was to comprehend what the real intention of the speaker was. The findings of this study showed that compared with younger adults, older people performed better at understanding the intentions of others in this set of stories, although the two groups were shown to perform equally on the control tasks (nonrelated tasks to ToM; Happé, Winner, & Brownell, 1998). Accordingly, Happé and colleagues (1998) argued that the extensive experiences of older people regarding social skills increased their wisdom and social intelligence, which might have been the reason for their better performance in understanding the mental states of others. In line with that, several other studies have confirmed that the capacity for ToM is not subject to change in the aging process (Keightley, Winocur, Burianova, Hongwanishkul, & Grady, 2006; MacPherson, Phillips, & Della Sala, 2002; Saltzman, Strauss, Hunter, & Archibald, 2000; Slessor, Phillips, & Bull, 2007). There are, however, other researchers who presented findings that contradict the lack of decline in ToM in late adulthood.

In contrast to the sociocognitive approach, the neuropsychological approach suggests that as people get older, they have more difficulties in understanding the emotions of others (Phillips et al., 2002). This decrease is due to the general cognitive decline of people in late adulthood. One of the arguments for this view is stated as:

“Since ToM is tied to the functioning of a specific cognitive architecture, declines in that cognitive architecture (e.g., Hedden & Gabrieli, 2004) may also be associated with declines in social functioning. The weight of the evidence suggests that older adults do show marked declines in aspects of ‘fluid’ intelligence (Hedden & Gabrieli, 2004); those aspects include skills such as working memory, processing speed, and numerical ability. If older adults perform worse at ToM, one explanation is that low fluid ability simply makes the task more difficult for them, independent of any domain-specific difficulties in ToM” (Moran, 2013, p. 33).

Accordingly, in a study by Maylor and colleagues, (2002), it was found that ToM abilities diminish as a result of the aging process. In this study, the performance of older adults versus younger adults was compared in a ToM stories task. Their findings indicated that ToM was impaired in both groups, until the point where the memory demands were increased, which led to impairment of ToM for the group of older people only (Maylor, Moulson, Muncer, & Taylor, 2002). This finding was replicated in a study by Sullivan and Ruffman (2004), in which they observed that when older people had to decode the feelings and thoughts of others from watching short and dynamic videos, they showed a poor performance.

As with the studies that showed no aging effect in the ability of ToM, there are also many studies showing the opposite findings (see e.g. Bailey & Henry, 2008; Castelli et al., 2010; German & Hehman, 2006; McKinnon & Moscovitch, 2007; Phillips et al., 2011). Additionally, Greenwood (2000) demonstrated that there are also neural functioning findings that support such age-related changes by showing alterations in frontal and temporal lobes, which are reported to be associated with ToM (Apperly, Samson, Chiavarino, & Humphreys, 2004).

These contradictory results are indicated to stem from methodological issues (Slessor et al., 2007). Still, the question remains of whether the accumulated life skills of older people enable them to understand the emotions of others in a more facilitated way, or the general decline in

their cognitive skills makes them have difficulties to interpret the emotional behaviors and intentions of others. What matters, however, is that in investigations that aim to examine the perception of emotional others, it is of great importance to consider the differences age might have on people's perceptions. Considering that virtual agents will be interacting with users of varying ages, investigating the factor of age can explain whether older people are able to correctly perceive the emotions of the agent and understand that its intention is to empathize with the users. The findings can contribute to the design of virtual agents that have different purposes and roles for different target groups.

3.3 Theory of Mind in Different Cultures

The second factor that is the topic of interest in this dissertation is how understanding the emotions and intentions of others (that the agent is empathizing with the users) is affected by environmental factors, namely, people's culture. Although many developmental studies regarding ToM suggest that this ability is a universal capacity (Mc Grath, 2009), there are also researchers that stress the important role of cultural factors in the development and function of ToM.

The existing differences in the social experiences of a particular group of people such as language, social values and habits, religious values, and geographical locations contribute to the definition of the term *culture* (Chiao & Ambady, 2007). Many cultural studies have emphasized the substantial cultural differences in individual traits, for example, in personality traits, visual perceptions, and spatial reasoning (Arnett, 2008; Henrich, Heine, & Norenzayan, 2010; Hofstede, 2001; McCrae & Terracciano, 2005; Nisbett & Masuda, 2003). Hence, the importance of taking into account the possible cultural influence on ToM cannot be overlooked.

Several attempts have been made to investigate the potential effect of culture in the development of ToM. Numerous studies have proved the existence of the same trajectories of ToM development in different cultures (Avis & Harris, 1991; Callaghan et al., 2005; Sabbagh Xu, Carlson, Moses, & Lee, 2006). One of the main pieces of evidence for the universality of ToM comes from the studies measuring the development of false-belief tasks in children. For example, based on a meta-analysis conducted by Wellman and colleagues (2001), it was observed that when comparing the trajectory of false-belief performance in children in different countries, their performance gets better as they get older. However, the data showed that although there is a

similarity in the performance, the starting point of gaining this ability varies in different countries (Wellman, Cross, & Watson, 2001).

Cultural factors that revolve around language and conversation can explain the differences in ToM in different cultures (Mc Grath, 2009). In support of this claim, it has been indicated that 4-year-old children with whom their families talk about emotions more openly show a more sophisticated comprehension of false-belief tasks (Hughes & Leekam, 2004; see Mc Grath, 2009), which is not the case for late-signing deaf children with less or no conversational engagement in false-belief tasks (Peterson & Siegal, 1998; see Mc Grath, 2009). There are also findings suggesting that environmental and cultural factors such as parenting style (Vinden, 2001), the language they use while interacting with their children (Ruffman, Slade, Devitt, & Crowe, 2006), and conversations related to emotions (Dunn, Brown, & Beardsall, 1991) impact the development of ToM.

On the whole, researchers agree on the universality of having ToM, although cultural factors play a role in how and when the mental states of others are perceived (Mc Grath, 2009). It is also important to note that the findings that confirm the cultural influences on ToM are in line with the theoretical approach of theory theory, which suggests that cultural surroundings shape the way people understand the mental states of others (Hughes & Leekam, 2004). Also, all of these cultural factors stem from and can be influenced by the dimension of whether the culture is collectivistic or individualistic.

The cultural dimension of collectivism and individualism and its impact on ToM is crucial in the context of this dissertation. The meaning of emotional situations and the appropriate way of expressing emotions might vary in different societies depending on their accepted norms and values (Ryder, 2008). The enculturation process in societies with different individualistic or collectivistic values (see Chapter 3.4) can influence the verbal and nonverbal way of expressing emotions (Cheon, Mathur, & Chiao, 2010). Therefore, people's perceiving and expecting to see specific behaviors as empathic responses might vary depending on their culture.

Accordingly, one of the main research questions here is concerned with knowing how people from different cultures (viz., individualistic vs. collectivistic cultures) perceive the emotional nonverbal behaviors of a virtual agent and rate it as an empathic agent (Study 2). It is thus important to have insights into the effect that the society being individualistic or

collectivistic has on the skill ToM (which can influence understanding the emotions of others). This, in turn, can determine people's expectations and preferences with regard to certain empathic responses, due to norms and values in those cultures that shape such expectations.

3.4 Reflection of Individualism/Collectivism in Theory of Mind and Empathy

Individualism/collectivism is demonstrated to be one of the most important dimensions of cultural differences with regard to values, perceptions, and behaviors of people, and was first introduced by Hofstede (2011). When a society is said to be individualistic or collectivistic, it refers to whether people prefer loosely or tightly knit social networks (Hofstede, 2011). More precisely, it has been stated that the extent to which cultures focus on individuals' desires and preferences over groups' or vice versa is related to those cultures being individualistic or collectivistic (Matsumoto et al., 2008). In individualistic societies, the interests of individuals are valued more than the interests of the group, which is the opposite for collectivistic societies (Eid & Diener, 2009; Hofstede, 2011). In collectivistic cultures, the social behaviors of people are influenced by the values of the group, including norms, duties, and obligations, which form and guide the behaviors of people (Eid & Diener, 2009; Hofstede, 2011).

There are many studies that have concentrated on collectivism/individualism and their potential role in having the skill of ToM. For instance, in individualistic societies such as the United States and Australia, people are reported to have "independent views of personhood," which can be explained by thinking more independently, framing one's own ideas, expressing thoughts more freely, and being involved in discussions (Greenfield, Keller, Fuligni, & Maynard, 2003; Nisbett, 2007). This, in turn, can ultimately shape children's primary conceptualizations of ToM. By contrast, in a collectivistic society like China, parents tend to teach their children many values that emphasize filial respect as well as conformity to the cultural norms and traditions rather than expressing one's own opinions in an independent way and getting well-established knowledge (Greenfield et al., 2003; Nisbett, 2007). This form of nurturing children in collectivistic societies results in different conceptualizations of ToM. The focus is mainly on the skill of people being acknowledgeable versus ignorant than being determined in their possibly different opinions and beliefs (Shahaeian, Peterson, Slaughter, & Wellman, 2011).

Additionally, neural evidence has also suggested the existence of the cultural effect of individualism and collectivism on ToM. For example, in an fMRI study by Kobayashi, Glover, and Temple (2006) (see Bradford et al., 2018), specific ToM abilities (ToM reasoning, non-ToM reasoning, and reading unlinked sentences as the baseline condition) were compared between American English-speaking monolingual and Japanese-English bilingual people. Although the authors observed significant activation in the medial prefrontal cortex and anterior cingulate cortex (which have been implicated in ToM) there were differences between the two groups. When both groups were engaged in performing ToM tasks, a weaker temporo-parietal junction activity was observed in Japanese participants with no significant behavioral differences. According to Kobayashi and colleagues (2006) (see Bradford et al., 2018), these findings may reflect the cultural influences on neural bases related to ToM, which might be due to the collectivistic values of Japanese culture versus individualistic values in American culture.

Understanding empathy and the acquisition of ToM are closely connected to each other. Therefore, having knowledge about the cultural background of people and its effect on their ability to understand the emotions of others is important in studies that focus on different target groups in interactions with empathic agents. This knowledge can lead to a better understanding of how expressing empathic responses and perceiving that other people are empathizing might differ in various cultures and, also, why people with different cultural values interpret certain behaviors as being more or less empathic.

Studies have also demonstrated the possibility of cultural impact on the level of illustrated empathic behaviors and how they were evaluated by people (Chopik, O'Brien, & Konrath, 2017). It is reported that in collectivistic cultures, people value and experience empathy more than people in individualistic cultures do (Chopik et al., 2017; Heinke & Louis, 2009). Henrich and colleagues (2010) explained this difference by saying that compared with collectivists, individualists are more independent in having their own personal views. This, ultimately, makes individualists possess more positive self-views and less tendency for conformity with others. Moreover, differences in the moral education and caregiving attitudes of different cultures can influence the empathic behavior of people. The enculturation process is diverse in various societies. This, as a result, can impact how people interpret the emotions of others and their expectations regarding emotional responses in specific emotional situations. How emotions are normally expressed, appraised, and perceived might vary in different societies, and this depends

on the cultural-relevant norms dictating which emotions should or should not be expressed by people (Matsumoto et al., 2008). Since people in collectivistic cultures are more concerned with having good interpersonal relationships, it triggers them to put more emphasis on social norms and how one should behave in different social situations (Takaku, Weiner, & Ohbuchi, 2001). Therefore, people's perception and interpretation of the suffering of others can be influenced and shaped by culturally constructed meanings (Cheon et al., 2010).

Additionally, studies have also investigated the differences that may exist between the perceptual abilities of collectivists and individualists. For example, in a study by Zhou and colleagues (2008), the cultural values of individualism and collectivism, such as self-construal style, were found to affect the visual perception and experience of people at a behavioral level. In that study, the authors utilized the Framed-Line Test, which is a test to measure the capacity of both incorporating and ignoring contextual information provided in a nonsocial domain. They first presented a square frame with a printed vertical line to their Japanese and North American participants and then another frame with a similar or different size. The first task was called the *Absolute Task*, in which they asked the participants to draw a similar line to the first line in an absolute length. The next task was the *Relative Task*, in which the participants were asked to draw a similar line in proportion to the height related to the surrounding frame. Their findings indicated that Japanese participants were more accurate in the Relative Task. This shows their ability to incorporate contextual information when perceiving a focal object. On the other hand, North American participants showed more accuracy in the Absolute Task, and this accuracy suggests that they are more able to ignore contextual information in the perception of a focal object (Zhou, Gotch, Zhou, & Liu, 2008).

The findings of this study suggest that culture can play a role in the way people use cues for the perception of the (emotional) states of others. More specifically, in the process of perceiving specific nonverbal behaviors as more or less empathic, various visual perceptions can impact people's usage of different aspects and cues in order to facilitate their perception of empathic nonverbal behaviors.

Altogether, the capability to accurately perceive and interpret the emotions of others, as well as sharing an emotional response in return, can lead to appropriate actions in different social and emotional contexts. Taking into account that collectivists tend to conform more with cultural

norms and traditions, and also that caring is a value in their cultures, they might have a higher expectation of receiving an emotional response from others. However, considering the cultural norms in individualistic cultures, this might not be the case for individualists. One of the main objectives of this dissertation was to investigate the different perceptions that might exist in users in an encounter with an emotional virtual agent. Therefore, in Study 2, it was examined whether the cultural background of collectivistic and individualistic users affects their understanding of the nonverbal behaviors of an empathic agent.

So far, many studies in the domain of human–agent interactions have focused on examining cultural impacts, including the dimension of collectivism/individualism, on users’ perception. However, most of these studies have used a very classic example of Western and far-Eastern users and studies have rarely taken into account Middle-Eastern users in their investigations. Human–machine interactions are increasing in the Middle-Eastern countries as well (Alharbi & Arif, 2018) and attitudes toward robots and agents are reported to be mostly positive in these societies (Riek et al., 2010). It is, therefore, crucial to also focus on cultural features of these societies, in order to gain more knowledge of users’ experiences and the perception of human–agent interactions worldwide.

Accordingly, in Study 2, the target groups of the users were chosen from two culturally different societies: Germany and Iran. The reason for using people from these two countries as the sample groups was that (1) although Germany as a Western country has individualistic tendencies, it is less individualistic than the typical example of the United States (Hofstede, 2019c; Westerhof & Barrett, 2005) (2) Iran as a Middle-Eastern country is a collectivistic society, although with less collectivistic values compared with China (Ghorbani, Bing, Watson, Davison, & LeBreton, 2003; Hofstede, 2019a), and (3) these two countries represent different cultural values, making it feasible to consider them as samples in a culture-related study. Hence, the state of the art is extended by putting more emphasis on Western and Eastern cultures that are not at the extreme ends of the spectrum regarding the individualism–collectivism dimension.

Taking into account the cultural differences between Germany and Iran, it has been demonstrated that Germans tend to communicate with each other in a more direct way. This means Germans value honesty even if it hurts, they have relatively independent views, they emphasize their own preferences and needs than the ones related to groups, and they are less

oriented toward obedience as a primary objective of socialization (Chasiotis, Kiessling, Hofer, & Campos, 2006; Hofstede, 2019b). By contrast, Iranians are said to have collectivistic tendencies, meaning that they put more focus on interdependency with others and connectedness with social values and norms (Hofstede, 2019d), on the harmony of the group and interpersonal relationships (Triandis, 1993), and on learning from early childhood to be sensible in their response to the emotions of others (Shahaeian, Nielsen, Peterson, & Slaughter, 2014). Moreover, maintaining relationships and prioritizing the group's needs instead of one's own, as well as being sensitive to the communicative intentions of other people, are characteristics that are valued and important in Iranian culture (Harb & Smith, 2008; Shahaeian et al., 2014).

All of the aforementioned differences regarding the existing values and norms in these two societies can be representations of parenting styles and strategies used for children's upbringing. As a result, people's expectations and preferences in terms of receiving empathic responses might vary. Also, these values can influence the development and having the ability of ToM (Shahaeian et al., 2014), interpreting the emotions of others, and understanding whether they are empathizing or not (Degens et al., 2012; Shahaeian et al., 2011). Therefore, the effect of collectivistic and individualistic tendencies of Iranian and German users was investigated on perceiving an empathic agent (Study 2).

4 Stereotypes Toward Virtual Agents

As indicated earlier, users' perception of an emotional agent is a significant factor in this dissertation. Accurately understanding one's interactive partner (here the virtual agent) can alleviate the level of misunderstanding and confusion in the interaction. Hence, another factor that could be important in investigating human-agent interaction (in the presence of cultural differences between users and the agent) is the existence of users' stereotypes toward an agent from a different culture. As the media equation theory suggested, people tend to carry over social rules to their interactions with technologies, in the sense that they respond to technologies in a similar way to how they respond to another human being (Reeves & Nass, 1996). A wide range of studies have examined media equation theory with respect to, for example, politeness, association of similar or different voice with a computer, and praise and criticism (Isbister & Nass, 2000; Lee & Nass, 1998; Nass & Moon, 2000). In Study 3 of this dissertation, however, the

focus was mostly on users' applying cultural stereotypes to their interactive virtual agent. The purpose was to investigate how people's categorization of ingroup/outgroup might lead them to rate an emotional agent differently. Cultural stereotypes are defined as people's generalized expectations about the motivations, behaviors, and feelings of others and using these expectations to judge other people (Stewart, Poweli, & Chetwynd, 1979; see Langenhove & Harré, 1994). People's stereotypes toward others from different cultures and ethnicities can come from categorizing others according to whether they are members of the ingroup/outgroup.

Since the agents are designed in specific countries, and there is a high probability that the users know about the design country of the agent, their potential bias toward the agent in the sense of an ingroup/outgroup needs to be examined. It has been stated that the match between emotional expressors and perceivers can contribute to fruitful and accurate communication (Elfenbein & Ambady, 2003). One of the factors that might lead to a difference in the perception among expressors and perceivers is the existing cultural differences. An important concept in psychological studies that examine such differences in perception is called *ingroup advantage*. Ingroup advantage is defined as an information-based explanation that determines that individuals can more precisely and accurately recognize and judge the emotions of people from their own cultural group versus emotions of people from other cultures (Elfenbein, 2015). The concept of ingroup/outgroup predicts that people own different identities (such as gender and ethnicity) and they interpret others as being either similar (perceived ingroup) or dissimilar (perceived outgroup) to their identities (Vang & Fox, 2014). When people categorize others as members of the ingroup, they perceive them as people with similar values (Allen & Wilder, 1979), and it has been reported that people perceive ingroup members as more trustworthy and likable (Clark & Maass, 1988; Lee, 1993). Further empirical studies have confirmed an ingroup advantage in emotion recognition as well, indicating the existence of greater accuracy for both expression and recognition of emotion by the members of the same culture (Elfenbein & Ambady, 2002).

With regard to applying stereotypical attitudes to embodied conversational agents, many studies have indicated that real-world stereotypes such as gender (Baylor & Kim, 2003; Zambaka, Goolkasian & Hodges, 2006) and age (Khan & De Angeli, 2007) are carried over to interactions with virtual entities. The effect of bias toward the ethnicity of others, as a very prominent visual factor that can be judged more easily, has also been investigated in interactions with virtual

entities. For example, Lee and Nass (1998) demonstrated that users had a more positive perception of an agent's attractiveness and trustworthiness and tended to accept the agent's arguments more when the agent had a similar ethnicity to theirs. In a study by Dotsch and Wigboldus (2008), the attitude of native Dutch subjects was examined in an encounter with either white or Moroccan avatars. The results of their study demonstrated that people tended to maintain more distance and showed increased skin conductance levels when they were encountering Moroccan avatars compared with White avatars. Similarly, Eastwick and Gardner (2009) found out that people may apply real-world racial biases in virtual environments. In their study, they tested real-world compliance tactics in the virtual world with avatars from different races. Their findings indicated that participants tended to comply more with a request by a White avatar than an avatar with a dark skin tone.

Furthermore, regarding the concept of empathy, results from a sample of medical students illustrated that the skin tone of a virtual agent is a factor that can impact the behavior and self-perception of empathy in patient interviews conducted with synthetic characters (Rossen, Johnsen, Deladisma, Lind, & Lok, 2008). Accordingly, Rossen and colleagues (2008) demonstrated that users showed less empathy toward a character with a dark skin tone in comparison with a light-skinned character.

By contrast, there are also findings that have not confirmed the results of previous studies in applying cultural stereotypes to virtual agents. For example, Vang and Fox (2014) manipulated race and task collaboration of virtual partners' identity in the context of a game. They examined how users would perceive the virtual agents, and whether they apply racial stereotypes from the real world to the virtual world. The participants in this study were randomly assigned to conditions to play anagram game competitively (outgroup, with a Black virtual agent) or cooperatively (ingroup, with a White virtual agent). Their findings indicated that owing to sharing a task and regardless of the purpose of the task, White participants evaluated their Black virtual partner as being more positive (Vang & Fox, 2014).

Additionally, in a pedagogical context, it was found that compared with the White agents in the role of experts, the Black virtual agents had a positive effect on students' learning, concentration, and self-regulation (Baylor & Kim, 2004). These findings were explained by Baylor and Kim (2004) with the "novelty effect," suggesting that since White students perceived

the agents as more novel, they paid more attention to them. The results of this study, however, are inconsistent with findings from their previous study which suggested that the pedagogical agent that had the same ethnicity as the students was perceived to be more engaging and affable (Baylor & Kim, 2003).

An implication of the aforementioned findings is the possibility that the culture of the virtual agents affects users' perception. However, more studies need to be conducted to determine whether categorizing an agent as ingroup/outgroup member indeed has an impact on users' attitudes and actions toward it. Moreover, most of the studies that have used the factor of ethnicity and users' stereotypes toward an agent are based on clear visual cues of the agent that indicate its ethnicity. Studies, to date, have rarely considered only the knowledge of the users of the cultural background of the agent and the potential influence it might have on their perception of the agent.

Although virtual agents are designed in particular countries, they are aimed at being utilized by different users worldwide. As presumed by Crandall and Eshleman (2003), people have at least some prejudices and stereotypes, no matter what their culture, nationality, ethnicity, or gender are. This makes it important to attempt to reduce such stereotypes as much as possible in the domain of human-agent interactions, in order to reduce unexpected changes in users' behavior in response to the agent. Having knowledge about the possible stereotypes that users might apply in their interaction with a virtual agent can provide insight into finding solutions. For instance, manipulating the appearance of the agent so that they are neutral-looking, might contribute to avoiding potential negative experiences of the users. Moreover, this insight can provide more evidence with respect to the media equation theory, by confirming or rejecting the notion that the mere knowledge of the country of others affects people's attitude toward them. Accordingly, in Study 3, it was strived to determine whether users' knowledge of the design country of the agent influences their perception of its emotional nonverbal behaviors and their rating of the agent as being empathic (which is a positive characteristic).

5 Nonverbal Behaviors in Human–Human Interactions

In this dissertation, the nonverbal behaviors of the agent are specifically of interest. This is because nonverbal behaviors are crucial and essential aspects of social interactions. They are basically defined as language-independent communication, which are used to regulate relationships and convey messages through expressions such as gestures, arm and leg movements, voice, distance between the communicators and so forth (Peterson & Leonhardt, 2015; Shin-Young & Harada, 2006). It has been acknowledged that nonverbal cues can foster and facilitate conversational flow by conveying thoughts and meanings in a more clear way, and can support the expression of emotions and attitudes (Peterson & Leonhardt, 2015). Moreover, it has been said that 65% of human communication is filled with nonverbal behaviors, which makes it impossible for humans to have a complete communication without utilizing nonverbal cues (Shin-Young & Harada, 2006). Besides, adults' interpretation of messages is reported to be mostly based on nonverbal cues rather than verbal ones (Shin-Young & Harada, 2006).

There are several studies that have investigated the essential role of different nonverbal behaviors such as facial expressions (Ekman, 2004), voice tone (Ambady et al., 2002), and posture (Carney, Cuddy, & Yap, 2010) in making human–human interactions more effective. Nonverbal behaviors are also of key importance in expressing empathic behaviors in a more eminent way, and as much as empathic communication is dependent on verbal messages, it depends on the nonverbal messages as well (Tepper & Haase, 1978). Additionally, it has been acknowledged that nonverbal cues play an important role in perceiving and judging the level of empathy as well (Tepper & Haase, 1978). In this regard, one of the crucial usages of nonverbal behaviors in human–human interactions is indicated in physician/consultant–patient relationships. Utilizing nonverbal behaviors are stated to be effective in facilitating emotional and empathic communication, leading to a positive effect on patient's satisfaction (Riess & Kraft-todd, 2014).

The significance of using nonverbal behaviors has been illustrated to be specifically of great importance among older people, especially older people with dementia. By using proximity and physical contact, older people initiate, elevate, and maintain spoken conversation, which in turn enables them to be more engaged in social situations (Hubbard et al., 2002). Hubbard and colleagues (2002) stated that older people use nonverbal behaviors in an intended and meaningful way to be part of the communicative world without the usage of verbal utterances. However, when it comes to understanding the nonverbal behaviors of others, specifically emotional facial

expressions, studies both in human–human (Ruffman, Henry, Livingstone, & Phillips, 2008) and human–agent interactions indicate that older adults show lower recognition of the observed emotions (Beer et al., 2015). This indeed introduces one of the main questions in this dissertation: Are older people able to correctly perceive and recognize the emotional nonverbal behaviors that the virtual agent uses to display empathic behaviors?

Having insight regarding the effectiveness of nonverbal behaviors for older people can contribute to an understanding of why implementing nonverbal behaviors in virtual agents, which are designed to interact with different target groups including older people, could be critically important. As indicated earlier, seniors utilize nonverbal cues to amplify the meaning of verbal utterances and to enhance their social engagement. However, it raises the question of whether they get more involved in an interaction with an agent that shows nonverbal behaviors for the purpose of empathizing, or they show difficulties in understanding precisely the meaning of specific (emotional) nonverbal behaviors. In other words, it is important to investigate whether older people have an accurate perception of the observed nonverbal behaviors of a virtual agent and perceive them in a positive way, in the sense that they perceive the agent attempt to be more communicative, engaging, and emotionally supportive. Therefore, in Study 1, the factor of age is taken into account for evaluating the emotional nonverbal behaviors of the virtual agent as empathic.

There are also other factors that can affect the understanding of nonverbal cues. Difficulties in interpreting the emotional nonverbal behaviors might occur when, for example, the interlocutors are not aware of the common social, situational, and cultural meaning of special nonverbal responses. Cultural differences have been observed in the expression of empathic nonverbal behaviors such as touch (Rousseau & Blackburn, 2008), body distance (Remland, Jones, & Brinkman, 1995), and eye contact. People from Eastern cultures, for instance, are reported to prefer more flexible eye contact and aversion of gaze, while in Western cultures, people prefer and value the maintenance of eye contact (Senju et al., 2013).

Moreover, as indicated in Chapter 4, culture can also play a role in generating prejudice and cultural stereotypes toward a virtual agent from a different country. It is, however, important to investigate whether the potential cultural stereotypes toward the agent can be transmitted to evaluating and perceiving its nonverbal responses and rating it as empathic.

These factors are important to consider when investigating the effect of emotional nonverbal behaviors in the context of human–agent interactions. Virtual agents are designed to interact with different users from different age cohorts as well as people from diverse cultural backgrounds. Thus, in this dissertation, the factors of age, culture of the users, and culture of the agent (which is manipulated in a cover story to make the users primed about the agent) are taken into account. The purpose is to examine the different perceptions these factors might generate regarding the nonverbal behaviors displayed by the agent. However, it is important to first have an overview of the related work that has been done in studying nonverbal behaviors in human–agent interactions.

6 Nonverbal Behaviors in Human–Agent Interactions

In the process of designing an effective empathic virtual agent, implementing nonverbal behaviors can be of significant importance in making the believability of the agent more eminent. There are several studies that investigated whether implementing nonverbal behaviors in virtual agents/robots matters in maximizing their interaction quality with human users. For example, Krämer and colleagues (2007) found that when an agent displays even subtle nonverbal behaviors, it affects the users’ evaluation and experience of that agent. In a study by DeSteno and colleagues (2012), it was found that the expressed nonverbal behaviors of a robot significantly affect the prediction of economic behavior. The robot in their study was rated as less trustworthy and was expected to give less token in an exchange, when it displayed a crossed-arm position and leaned away. The effect of nonverbal behaviors of a virtual agent on users’ perception was also investigated with regard to dominance and cooperativity (Straßmann, von der Pütten, Yaghoubzadeh, Kaminski, & Krämer, 2016). It was shown that when the agent displays symbolic gestures such as crossing the arms, stemming the hands on the hip, or touching the neck, users perceive the agent as more dominant, and in the presence of combined expressive gestures and facial display, users perceive the agent as cooperative (Straßmann et al., 2016).

Further studies examined the effect of behaviors such as human-like gaze (Heylen, Van Es, Nijholt, & van Dijk, 2002; Rickenberg & Reeves, 2000), head position and head movements (Krämer, 2001), and facial expression (Baylor, Kim, Son, & Lee, 2005) in human–agent interactions. It has been reported that when the agent shows these nonverbal behaviors, users evaluate the agent as more positive, natural, and involved. Considering the importance of using

nonverbal behaviors in facilitating the expression of emotions and empathy, Prendinger and colleagues (2005a) have illustrated that these behaviors can play a role as a stress-reducing channel in human–agent interactions.

The fact that there are inter-individual differences in humans makes it important to know the users' attitudes and preferences with regard to a virtual agent's behavior (Shneiderman et al., 2016; Krämer, Hoffmann, & Kopp, 2010). To this end, in this dissertation, factors such as the age of the users, the culture of the users, and users' stereotypes toward the culture of the agent have been taken into account. The objective was to investigate their potential influence on users' perception, evaluation, and preferences of specific emotional nonverbal behaviors to rate an agent as empathic.

There have been several studies documenting the effective role of age in perceiving the nonverbal behaviors of a virtual agent. For instance, it has been shown that observing an agent displaying self-touching gestures influenced how attentive the older people were (Krämer et al., 2010). Further, age-related differences in emotion recognition were examined by Beer and colleagues (2015). They aimed to investigate whether the age-related differences that exist in the human–human interaction literature regarding recognition of facial expression can be transmitted to understanding the facial expression of virtual agents as well. Their findings described that there is indeed such age effect, by showing the difficulty that older users have to accurately recognize negative facial emotions. However, their results indicated that the intensity of emotions did contribute to the improvement of recognition in this user group (Beer et al., 2015).

Further, since the ultimate purpose of designing virtual agents, specifically virtual assistants, is to help the users with their daily tasks, beyond their nationality and culture, a careful design with culture-specific emotional nonverbal behaviors might be needed. To this end, it is important to investigate the cultural differences that exist among users. The impact of users' culture might be seen in the accurate perception of the emotional nonverbal behaviors of the agent and ultimately understanding that the agent is empathizing with them.

Several attempts have been made to considering culture when implementing nonverbal behaviors in virtual agents. The influence of culture-specific behaviors such as the spatial extent of gestures was examined by Endrass and colleagues (2011). They attempted to understand the cultural impact on the perception of German and Japanese users of virtual characters. Their findings indicated the significant role of users' culture in the interpretation of a virtual character's behavior. It was shown that users preferred the agent to resemble verbal as well as nonverbal

behaviors (e.g., postures, spatial extent of a gesture, and overlapping speech) that were more congruent and similar to the ones in their own cultures (Endrass, André, Rehm, Lipi, & Nakano, 2011). Furthermore, a simulation of cultural differences in behaviors such as proxemics, gaze, and overlap in turn-taking has been performed in order to reflect culture-specific behaviors between North American, Mexican, and Arabic cultures (Jan, Herrera, Martinovski, Novick, & Traum, 2007). The findings of this study suggested that users can, in fact, perceive the difference between the agent's behaviors that are related to their own culture and the behaviors related to a different one.

Moreover, as mentioned earlier, there is the possibility that people apply cultural stereotypes to an agent from a different culture and ethnicity. Investigating whether human users apply real-world stereotypes when encountering avatars from a different culture, Koda and Ishida (2006) and Koda (2007) scrutinized the cultural influences (ingroup advantage) on the interpretation of an avatar's facial expression, including emotional expressions. In their first study, Koda and Ishida (2006) compared the interpretation of users from Japan, South Korea, China, the United States, the United Kingdom, France, Germany, and Mexico of the facial expressions of an avatar, designed by a Japanese designer. Their findings indicated the existence of cultural differences and ingroup advantage specifically within Japanese and between South Korean and Japanese users. Following up on this study, Koda and colleagues (2008) conducted another experiment to investigate whether an ingroup advantage appears in the interpretation of Western-designed (by French, British, and American designers) avatars' facial expressions in participants from the United States, France, Germany, and Japan. Consistent with the results of the previous study, they observed tendencies of ingroup advantage among the Western countries in the interpretation of Western-designed avatars (Koda, Rehm, & André, 2008). However, in these studies, the cultural differences of the avatars were represented in the design of the facial expressions only. Furthermore, the look of the agent made it more explicit for the users that the avatars had a different ethnicity. In Study 3 however, it was intended to determine whether people apply ingroup advantage in understanding an agent from a different country, when they have no apparent cues indicating such differences, but merely the knowledge that the agent is designed in a different country. Additionally, other nonverbal behaviors, besides facial expressions, were presented to the users to be evaluated.

The studies mentioned so far indeed offer important insights into considering cultural factors when implementing nonverbal behaviors in virtual agents. However, there is a remarkable

lack in the literature of any attempt to understand how the cultural background of the users (not the agent) affects their perception and preferences regarding specific nonverbal behaviors as empathic. This gap in the literature includes the target group of older people as well: It is not yet totally clear whether there are age-related differences regarding accurately and precisely perceiving that the agent is empathizing with the users, and which specific nonverbal behaviors are perceived as more or less empathic by older people. These are the questions that are tackled in Study 1 (with a focus on age) and Study 2 (with a focus on the users' cultural background). Additionally, previous studies investigated the role of users' culture in the perception of nonverbal behaviors, however, studies that examine the perception of emotional nonverbal behaviors are scarce.

Furthermore, although virtual agents are designed in different countries, they are intended to interact with users from all over the world. In order to achieve effective communication between the users and the agent, it is thus important to consider whether the knowledge of the users of the country of the agent makes them biased toward correctly understanding the empathic nonverbal responses of the agent. This assumption that participants might be biased toward the culture of the agent is tested in the third study of this dissertation.

7 Empathy in Human–Agent Interactions

As indicated in Chapter 2, empathy plays a central role in human social interactions. Therefore, enabling virtual agents to act empathically (inspired by the way humans might act in social interactions) can elevate the users' experience of interacting with the agents. Empathic agents are defined as agents with the potential to put themselves in the shoes of users (or another agent) and respond to their emotional situation appropriately (Paiva et al., 2017). It has long been discussed that empathic agents can contribute to enhancing a more qualified human–agent interaction. For example, in a study by Prendinger et al. (2005b), a virtual character was used as a quizmaster in the context of a mathematical game. In addition to questionnaire data, they used physiological information of the participants to accurately associate their autonomic nervous system activity when interacting with an empathic agent. The results demonstrated that when the virtual quizmaster responded to the users in an empathic way, it had a significant effect on the decline of users' stress. This suggests that empathic behavior can have a positive influence on how users perceive the difficulty of the task (Prendinger et al., 2005b). Using the same method,

Prendinger and Ishizuka (2005) observed consistent results. They reported that in the context of a job interview, when the agent (in the role of an interviewer) gives empathic feedback to the participants, it has a positive effect on reducing the interviewees' negative feelings and stress level.

Moreover, an investigation on users' perception of an empathic agent showed that in the presence of empathic responses, the users perceive the agent as likable, trustworthy, and caring (Brave et al., 2005). Also, in a learning environment, it has been reported that when the agent shows more emotional intelligence, the users tend to perceive the agent as more trustworthy (Maldonado et al., 2005).

Besides the beneficial usage of an interaction with technology, there are unpleasant side effects such as the triggering of negative emotions, confusion, frustration, anxiety, and anger in users (Klein et al., 2002). To reduce such negative side effects, Klein et al. (2002) utilized an empathic agent that provided the users with support to manage their negative emotions. What the findings of their study indicated is that the empathic agent indeed reduced users' frustration. Also, they found that users continued to interact with the empathic agent, even though it was the agent that caused the frustration in the first place. These findings were further supported by Hone (2006).

Furthermore, since one of the goals of designing artificial entities such as assistant agents and robots is to be in daily interactions with the users, it is important to assure the users re-use the technology after their primary interaction. Hence, implementation of empathic behaviors is stated to lead to the users' higher perception of friendliness of the specific artificial entity (Leite et al., 2013), which can ultimately generate a sustainable long-term interaction with it (Leite et al., 2014).

Many studies have also focused on investigating cultural differences regarding empathic behaviors of the agent, as well as creating an environment in which the users could be triggered to empathize with the agent. In this regard, Degens and colleagues (2012) indicated that in the design process, culture is one of the requirements to create a conceptual model for the social behaviors of an empathic agent. According to their model, culture is remarkably important in impacting the behaviors of the agent through understanding social norms and moral circles.

Further, a computer application called FearNot was presented by Vannini and colleagues (2011), in which empathy is at the center of the interaction between the virtual characters and users. The objective of designing this application was to tackle the issue of bullying in schools.

They provided a school-based virtual learning environment, in which different virtual characters played various bullying scenarios. Importantly, in the design of this application, cultural differences in two societies (the United Kingdom and Germany) were taken into account in order to create social agents that the users could feel empathic toward.

Additionally, Aylett and colleagues (2009) presented a semi-immersive graphical environment, ORIENT (Overcoming Refugee Integration with Empathic Novel Technology). In this environment, a believable agent was developed for educational purposes to improve intercultural empathy among the users. It is reported that utilizing this simulated environment is effective in personal and social education, which takes the beliefs, attitudes, and behaviors of students into account for qualified and enhanced educational purposes (Aylett et al., 2009). In this simulation, the users observed an imaginary foreign culture in which a believable educational application was used for the improvement of intercultural empathy, cultural awareness of the users, and integration of students with refugee or immigration background. Importantly, they adapted Hofstede's cultural dimensions for simulation of various cultures, and also adapted the culture of the agent to facilitate the process of learning cultural differences for the users (Aylett et al., 2009).

The aforementioned studies offer useful insights into understanding the positive effect of considering empathy and the factor of culture in human-agent interactions. The concern and focus of these studies were mostly on modeling and simulating various cultural environments, in order to improve empathic behaviors of the users with regard to specific purposes. The aim of this dissertation (Study 2), however, is not triggering empathy in the users by considering their culture, but by investigating its effect on their interpretation of the empathic (nonverbal) behaviors of an agent.

III RESEARCH APPROACH

Previous chapters presented the theoretical background regarding the concept of empathy and the ability to understand that others are empathizing, as the basis of the research in this dissertation. To accurately understand the empathic responses of a virtual assistant, it is important that users be aware of the intention of the agent (i.e., being assistive), trust the agent, and maintain their social relationships with it in order to reuse it and benefit from such technology for a longer period.

In previous chapters, the positive role of implementing empathy in virtual agents was outlined. While the literature shows the important effect of empathic responses of the agent on users' evaluation of the agent, the impact of user variables such as age and culture on the perception of an empathic agent is not yet clear. Although virtual agents have the potential to be used by all kinds of users, the group of older people can benefit the most from virtual agents, namely, virtual assistants. This is due to seniors' difficulties in handling daily life routines such as recalling appointments or taking medications on time. However, there is a lack of research investigating whether older people have difficulties to accurately recognize the empathic responses of the agent. More importantly, utilizing emotional nonverbal behaviors can facilitate the agent's expression of empathy. However, no previous study has investigated how the user's age affects the perception of the nonverbal behaviors of the agent, and whether older users perceive certain emotional nonverbal behaviors of the agent as more or less empathic. The aim of Study 1, therefore, is to determine whether older people and younger adults show different perceptions of an empathic agent based on the agent's displayed emotional nonverbal responses.

Furthermore, the second study of this dissertation is concerned with another user variable, culture. Although several attempts have been made to consider users' culture in the implementation of nonverbal behaviors of the agent (e.g., Endrass et al., 2011; Jan et al., 2007), studies that investigate the effect of the cultural background of users on their perception of *emotional* nonverbal behaviors (that can lead to the perception of empathy) are scarce. Also, several studies have taken the *culture of the agent* into account with the purpose of changing behaviors in users by triggering empathic responses in them (e.g., Vannini et al., 2011). However, the question remains open regarding which emotional nonverbal behaviors of the agent,

independent of which culture it represents, trigger the perception of empathy in users who have a diverse cultural background.

Finally, in Study 3, the aim is to examine how the knowledge about the culture of the agent can make users biased toward it and lead them to perceive the agent that is designed in a different culture as less empathic. Cultural stereotypes toward agents have been investigated in several studies (e.g., Lee & Nass, 1998; Pratt, Hauser, Ugray, & Patterson 2007). However, in these studies, the ethnicity of the agent had an obvious representation in the appearance of the agent. It is important to note that far too little attention was paid to investigating the effect of users' mere knowledge of the culture of the agent on their perception of empathic agents. Furthermore, studies that have investigated the effect of ingroup/outgroup advantage in understanding the emotional nonverbal behaviors of the agent (e.g., Koda & Ishida, 2006; Koda, 2007) mostly focused on emotional facial expressions only. Other nonverbal behaviors of agents and nonverbal behaviors with different intensities are, therefore, rarely taken into account. Also, in these studies, as in others, the cultural orientations of the agents were manifested in the agent's appearance, which plays a role in making the bias more prominent. Furthermore, in their studies of understanding emotional nonverbal behaviors, Koda (2007) and Koda and colleagues (2006; 2008) used static pictures of agents to investigate users' perceptions, and studies that use dynamic video stimuli are rare. Accordingly, in Study 3, several emotional nonverbal behaviors were presented to the users as responses to emotional situations in sequences of dynamic videos.

Having knowledge about the possible cultural stereotypes of the users toward the agent could provide a basis for further investigations. Accordingly, it can be investigated whether agents should have a culture-specific or neutral-looking appearance in order to reduce users' dissatisfaction with their interactions with agents designed in different cultures. This would, in turn, lead to less miss-interpretation and miscommunication between human users and virtual agents, and therefore enhance the possibility of reusing the agents in the future.

In the next chapter, empirical studies that have been conducted in this dissertation to investigate the effect of the aforementioned factors on users' perception of emotional nonverbal behaviors of the agent will be presented.

IV EMPIRICAL STUDIES

The goal of this dissertation is to provide a basis for research that is concerned with designing virtual agents for different target groups, by considering the factors that can influence users' perception of an empathic agent. In all three studies of this dissertation, the focus was on the perception of the users, and on examining which specific nonverbal behaviors are perceived as more empathic. Therefore, it was strived to use the same materials (with minor changes such as translating the material and adding extra questionnaires) and design of the study in order to avoid other potential factors that can affect such perceptions.

The goal of the following chapters is, therefore, to present the empirical studies that investigated the factors that can potentially influence users' perception of emotional nonverbal behaviors of an empathic agent. Accordingly, Study 1 is focused on investigating the factor of users' age, Study 2 is focused on users' cultural background, and Study 3 is concerned with users' stereotypical attitudes toward the agent.

8 Study 1: Empathy for Everyone? The Effect of Age When Evaluating a Virtual Agent

The aim of the first study was to investigate the role of age in the perception of emotional nonverbal behaviors of a virtual assistant and rating the agent as empathic. To date, it has not been investigated which specific nonverbal behaviors lead to the higher perception of empathy in different users from different age cohorts. Therefore, in this study, several emotional nonverbal behaviors with different intensities were presented to the participants of varying ages. These nonverbal behaviors were presented as emotional responses to imaginary emotional situations. It was, thus, intended to scrutinize which emotional nonverbal behaviors are rated as more empathic by older people and younger adults. Further, it was aimed to investigate whether after watching the responses; the participants rate the agent differently regarding friendliness, intelligence, trustworthiness, and helpfulness.

8.1 Hypotheses and Research Questions

As mentioned in chapter 3.2, studies have investigated the role of age in understanding the mental states of others or ToM. Some of these studies proved that ToM abilities increase with aging due to accumulation of knowledge and experience across lifespan - as advocates of sociocognitive approach demonstrated - (Blanchard-Fields, 2007; Keightley et al., 2006; MacPherson et al., 2002; Saltzman, et al., 2000; Slessor et al, 2007). By contrast, other studies suggested a decrease in this ability due to general cognitive decline of older people - as advocates of neuropsychological approach argued (Bailey & Henry, 2008; Castelli et al., 2010; German & Hehman, 2006; Maylor et al., 2002; McKinnon & Moscovitch, 2007). Even though findings of these studies indicated a mixed pattern in the ability of ToM in older adults, they all have one issue in common: the probability that there is an age effect regarding understanding other people's minds and their emotional responses in either a better or worse way.

Accordingly, in the first study of this dissertation, this probability is examined by investigating the perception of older and younger users of emotional nonverbal behaviors of a virtual assistant. The first hypothesis, thus, indicated that:

H1: There is an age difference with regard to the perception of the agent as empathic.

It is further assumed that perceiving the agent as empathic is based on the observed emotional nonverbal behaviors in all participants. This is due to the findings of the studies (e.g. (Peterson & Leonhardt, 2015) suggesting that nonverbal behaviors increase transferring the messages to interlocutors:

H1.1: The perception of the agent as empathic is based on the observed nonverbal behaviors.

Furthermore, an interaction effect between the age of the participants and the observed nonverbal behaviors is predicted, indicating that different nonverbal behaviors lead to a different perception of empathy in the users:

H1.2: There is an interaction effect of age \times nonverbal behaviors in the perception of empathy.

Moreover, the specific focus of this study was to examine if older people and younger adults perceive specific emotional nonverbal behaviors as more or less empathic. Therefore, the first research question of the study asked:

RQ1: In case of the existence of a significant between-group effect, which emotional nonverbal behaviors are rated as more empathic by older and younger participants?

Besides empathy, other items such as friendliness, intelligence, trustworthiness, and helpfulness were presented to the participant to evaluate the agent. The intention of presenting these items was to avoid attracting the attention of the participants to the main purpose of the study (which was evaluating the item empathy). Therefore, ratings these items are considered as the overall evaluation of the agent. Accordingly, the second research question of the study asked:

RQ2: Do older people and younger adults rate the agent differently regarding the agent's friendliness, intelligence, trustworthiness, and helpfulness in the presence of emotional nonverbal responses?

8.2 Method

The design of the study was a mixed factorial design with a 2 between-subjects factors (age: older vs younger participants) \times 2 within-subjects factors (situations: sad vs happy) \times 4

within-subjects factors (nonverbal behaviors for each situation: 3 emotional vs 1 neutral). The purpose was to examine whether the age of the participants affects their perception of an empathic agent in the presence of the emotional nonverbal behaviors of the agent. It was also strived to test how the overall rating of the participants might differ depending on the age of the participants with regard to friendliness, intelligence, trustworthiness, and helpfulness.

8.2.1 Material

The virtual agent which was used in this study was an embodied conversational agent called Billie (Social Cognitive System Group, CITEC, Bielefeld, Germany¹), which is a humanoid agent and has a child-like look (see Figure 2).

A total number of 25 videos were created and implemented in the virtual assistant, Billie. The task of the agent was to review the imaginary monthly schedule of the participants. The first video was an introductory video in which Billie introduced itself and explained its role. The following 24 videos included emotional situations (as imaginary appointments and events) which were expressed by Billie. There were six emotional scenarios in total, including three happy and three sad situations. These situations were selected out of 12 emotional sentences which their emotional valences were evaluated in an online pre-test. Fifteen German participants rated the positive and negative values of the sentences on a 5-point Likert-scale from “I totally agree” (highest score = 5) to “I totally disagree” (lowest score = 1). In the end, six sentences with the highest emotional scores were selected as the emotional situations to be presented in the main study (see Table 1). All the materials in this study were in German.

¹ <https://scs.techfak.uni-bielefeld.de/>

Table 1
Emotional sentences uttered by the agent

Situations	<i>M</i> and <i>SD</i> of the Ratings
<p>Happy: On Thursday, July 11th, at 19:00, you will go to the concert of your favorite singer.</p>	<p>Positive Rate: $M = 4.33, SD = 1.04$</p>
<p>Happy: On Wednesday, July 24th, at 17:00, you will have a reunion with your old friend.</p>	<p>$M = 4.20, SD = .56$</p>
<p>Happy: On Saturday, July 27th, at 11:00, you will go on a picnic with your best friends.</p>	<p>$M = 4.27, SD = .59$</p>
<p>Sad: On Tuesday, July 9th, at 14:00, you will visit your friend in the hospital who suffers from cancer.</p>	<p>Negative Rate: $M = 3.87, SD = 1.35$</p>
<p>Sad: On Tuesday, July 16th, at 11:00, you will have a doctor appointment for your upcoming operation.</p>	<p>$M = 4.13, SD = .83$</p>
<p>Sad: On Sunday, July 28th, at 19:00, you will hold a speech at the funeral of your friend.</p>	<p>$M = 4.47, SD = .64$</p>

These emotional sentences were implemented in the agent and expressed by Billie. After uttering each of these sentences, Billie displayed a nonverbal behavior at the end of the videos as an empathic response. Using Key-frame editor, several nonverbal behaviors were created to display emotional responses. Key-frame editor (Figure 1) is a program (Social Cognitive System Group, CITEC, Bielefeld, Germany) that is designed for editing nonverbal behaviors of the agent, which includes a list of joints and facial expressions. Giving different values to each of these lists, different bodily postures and facial expressions can be made with the desirable duration.

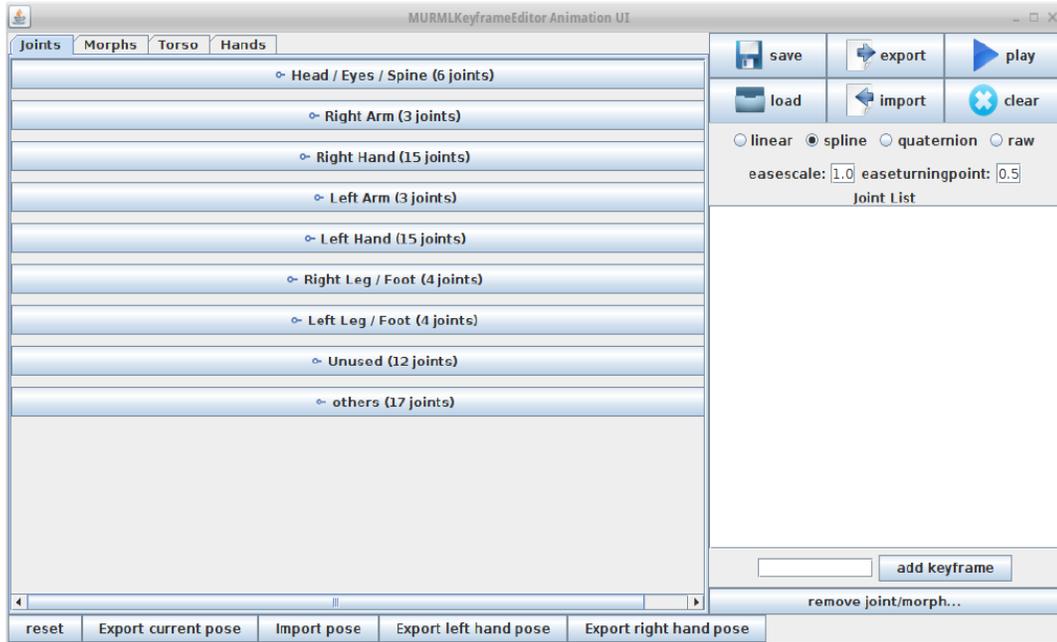


Figure 1: Key-frame editor and the list of body parts which can be manipulated

First, an initial position was created in which the agent had its hands rested in front of the stomach while the fingers were tangled. The nonverbal responses to happy situations were named as Smile, Head Nod, and Head Nod plus Smile. The nonverbal behaviors in response to sad situations were named as Sad Face, Head Down, and Dropping the Arms plus Sad Face. The details of the nonverbal behaviors and their duration are presented in Table 2. Developing the emotional facial expressions of the agent was inspired by facial expressions of the avatars in Study by Koda and colleagues (2008). However, in order to investigate whether different intensity of the displayed nonverbal behaviors influences users' perception, different body and head movements were also created and implemented in the agent.

After showing each of these nonverbal behaviors for some seconds, the agent's gesture and facial expression went back to the initial position. Moreover, a neutral behavior, as the control behavior - in which the agent showed no emotional nonverbal behavior - was created as well. The purpose of creating the control behavior was to test whether participants rate the agent differently in the presence or absence of the emotional nonverbal behaviors.

Table 2

Emotional nonverbal behaviors and their duration

Happy Nonverbal Behaviors	Sad Nonverbal Behaviors
<p>Smile: the sides of the mouth rise, making the cheeks be pushed up, to display happiness. Duration: 3 seconds.</p>	<p>Sad Face: the inner corners of the eyebrows move upward, and the upper eyelids are drooped. At the same time, the corners of the lips move downward, making the chin looks raised to display sadness. Duration:4 seconds.</p>
<p>Head Nod: the head is tilted in alternating up and down arcs for one time, to display acknowledgment and interest to what has been said. Duration: 2 seconds.</p>	<p>Head Down: the agent lowers its head and moves its eye gaze downward, stays in this position for four seconds, and moves it up again. Total duration: 5 seconds.</p>
<p>Head Nod plus Smile: The combination of one-time head-nodding accompanied by the smiling face, to show a more intensive acknowledgment and bonding signal. Duration: 4 seconds.</p>	<p>Dropping the Arms plus Sad Face: the agent shows the sad face and drops its arms to its sides simultaneously to display sadness. It stays in the dropped-arms position for four seconds and then goes back to the initial pose. Total duration: 5 seconds.</p>



Figure 2: Examples of Billie showing emotional nonverbal behaviors from left to right: Initial position (neutral), Smile, Sad Face, Head Down, Dropping the Arms plus Sad Face

In order to assign all emotional nonverbal behaviors to the related emotional context, the agent repeated the situations randomly and showed different context-related emotional responses (the assignment of nonverbal behaviors is shown in Figure 3 & Figure 4). Therefore, all happy nonverbal behaviors and the neutral behavior were assigned to happy situations, and all sad nonverbal behaviors and the neutral behavior were assigned to sad situations.

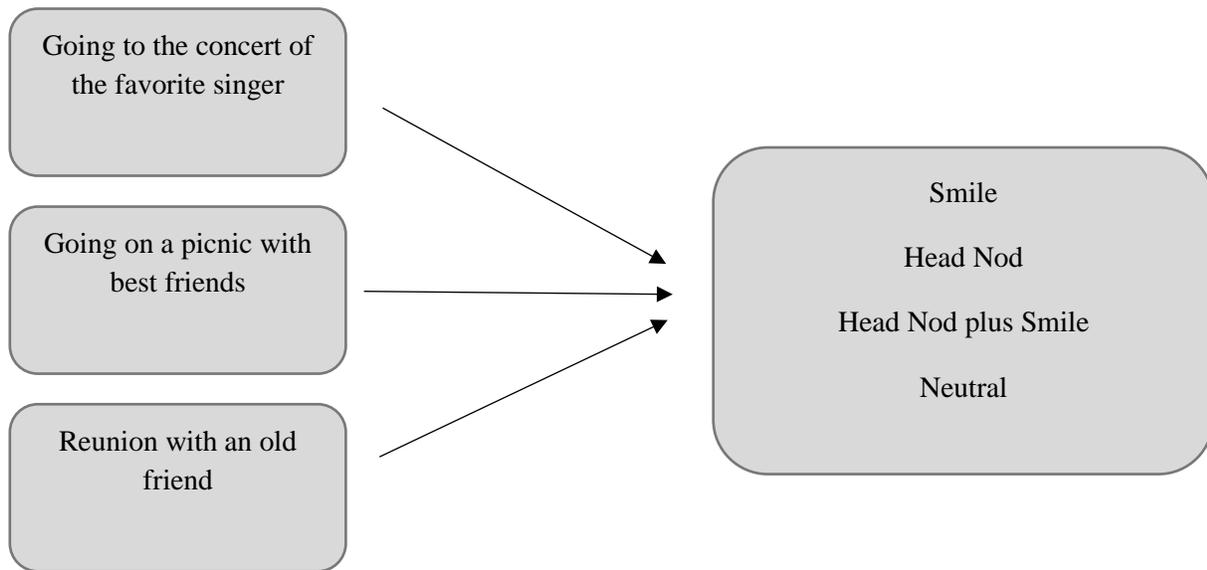


Figure 3: Assignment of Nonverbal Behaviors to the Happy Situations

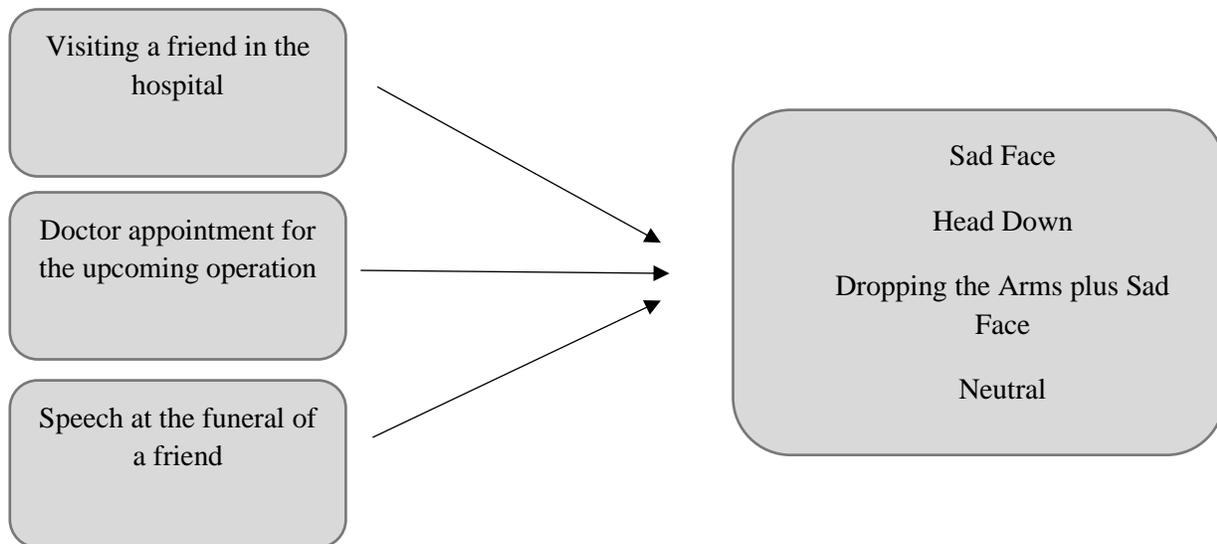


Figure 4: Assignment of Nonverbal Behaviors to the Sad Situation

8.2.2 Sample

A total number of 60 participants, 30 older people (8 male, 22 female) and 30 younger adults (10 male, 20 female) were recruited to take part in the study. On average, the older people were 66 years old ($M = 66.43$, $SD = 10.33$; range from 53 to 96 years old). The average age of the younger adults was 24 years old ($M = 24.36$, $SD = 5.56$; range from 18 to 39 years old). The younger participants were recruited online and in order to avoid any problem with regard to encountering technology, the older participants were recruited in the lab. The stimulus materials were presented to the seniors on a 15.6-inch laptop screen. In advertising for recruiting participants, it was recommended to open the survey on a laptop or Personal Computer, and not on a mobile screen. Prior to commencing the study, ethical clearance was sought from the ethics committee of the University of Duisburg-Essen. Participation in the study was voluntary and anonymous. The seniors received 5 € for taking part in the study and the younger adults had a chance to participate in a raffle of 8 Amazon vouchers.

8.2.3 Procedure

The first page of the survey presented to the participant was the welcoming page that included the briefing of the study as well. Then, the participants were asked to fill in their demography, including gender and age. Afterward, the introduction video was presented to the participants. In this video, Billie introduced itself and its role as a virtual assistant to review the imaginary monthly schedule of the participants. Further, Billie instructed the participant to click on “enter” whenever they were ready to proceed with the study. Then, the participants watched the following 24 videos in which Billie was expressing the emotional situations (see Table 1).

Moreover, in order to avoid the sequence effect, three sets of video orders were created, and the participants were randomly assigned to each of these orders. After expressing each sentence, Billie showed either an emotional nonverbal response or the neutral response. Then, immediately after watching each video (the duration on average was 12 seconds) the participants needed to rate the agent by filling in the questionnaires. Finally, on the last page of the study, the participants were debriefed on the purpose of the study.

8.2.4 Measures

Independent Variables: *Age group.* In order to investigate the effect of users' age on the perception of an empathic agent, two groups of older people and younger adults participated in the study. All participants watched the same videos of Billie expressing emotional sentences (situations) and displaying nonverbal behaviors.

Dependent variables: *Questionnaires with Five Items (friendly, intelligent, empathy, trustworthy, and helpful).* In order to test the participants rating of the emotional virtual agent, five items were presented to the participants to rate ("How do you rate Billie?"). On a 5-point Likert scale (from 1 = *I totally disagree* to 5 = *I totally agree*), the participants had to rate the agent: "I feel Billie is:" friendly, intelligent, empathic, trustworthy, and helpful.

8.3 Results

To test the hypotheses and the research questions of the study, multiple mixed ANOVAs using IBM SPSS version 23 were conducted separately for each of the five items (friendliness, intelligence, empathy, trustworthiness, and helpfulness). Furthermore, to analyze the interaction effect between the within and between-factors, separate repeated ANOVAs were conducted for each target group. The evaluation of the scores was calculated as follows: the average of each emotional nonverbal behavior and the average of all neutral behaviors, i.e. the average of all "Smile" behaviors (which were 3), the average of all "Head nod plus Smile" behaviors (which were 3) etc., and the average of neutral behaviors (which were 6). Post-hoc test with Bonferroni correction was used in the analysis of all cases. When the assumption of sphericity was violated, Greenhouse-Geisser correction was reported.

H1 stated that there is an age difference in perceiving the agent as empathic. Results indicated a significant main effect of age on rating the item empathy $F(1,58) = 5.01, p = .029, \eta^2 = .08$. Across all videos, the older people rated the agent as more empathic compared to younger people, supporting H1. Descriptive values of the two age groups are shown in Figure 5.

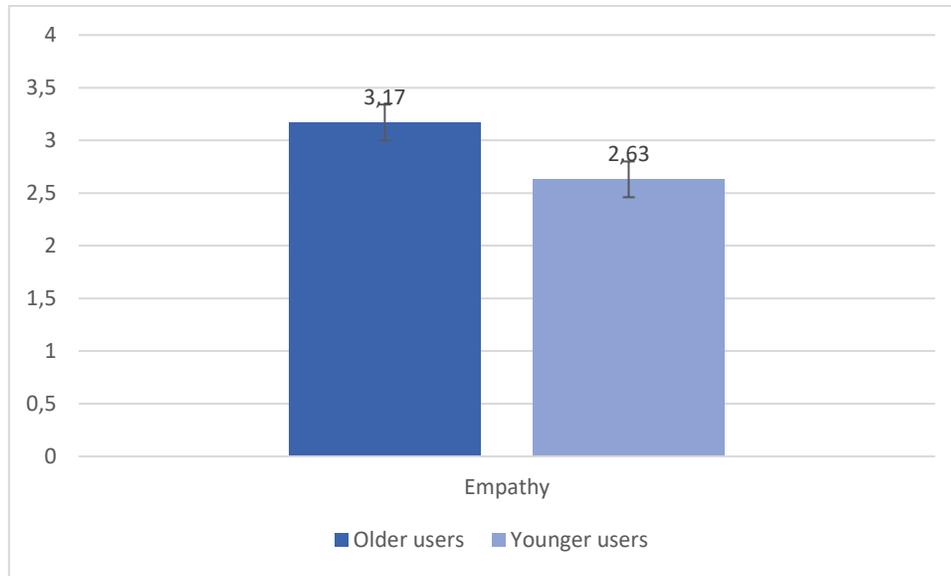


Figure 5: Descriptive values of rating the item empathy for older and younger users

H1.1 assumed that users perceive the agent as empathic based on its displayed nonverbal behaviors. The results revealed a significant effect of the nonverbal behaviors on rating the agent as empathic, $F(3,174) = 9.24$, $p < .001$, $\eta^2 = .14$. This supports H1.1, indicating that observing the nonverbal behaviors affected the perception of empathy across all participants. Contrasts revealed that all emotional nonverbal behaviors except for the Head Nod affected participants' perception of empathy (independent of the age) compared with the neutral behavior. The results are shown in Table 3 and Table 4.

Table 3

Statistical values of the contrasts for the perceived nonverbal behaviors – independent of age (empathy)

Nonverbal behavior	<i>F</i>	<i>df</i>	<i>p</i>	η^2
Dropping the Arms plus Sad	18.16***	1,59	<.001	.23
Head Down	18.50***	1,59	<.001	.23
Sad Face	6.29*	1,59	.011	.10
Head Nod plus Smile	10.30**	1,59	.002	.15
Smile	5.36 *	1,59	.024	.08
Head Nod	.09	1.59	.755	.00

* $p < .05$. ** $p < .01$. *** $p < .001$

Table 4

Descriptive values of the perceived nonverbal behaviors – independent of age (empathy)

Nonverbal behavior	<i>M</i>	<i>SD</i>
Dropping the Arms plus Sad	3.11	1.00
Head Down	3.13	1.08
Sad Face	2.95	1.08
Head Nod plus Smile	2.97	1.07
Smile	2.90	1.03
Head Nod	2.72	1.02
Neutral	2.70	1.04

Based on the assumption of H1.2, there is an interaction effect of age and nonverbal behaviors. The results supported the hypothesis by indicating a significant interaction effect between the nonverbal behaviors and the age of the participants, $F(3,174) = 4.33$, $p = .012$, $\eta^2 = .07$. This demonstrates that the perception of empathy by older and younger people was affected differently when they observed different nonverbal behaviors.

Furthermore, RQ1 asked which nonverbal behaviors are perceived as more empathic by the older and younger users. To answer this, contrasts were performed to compare each emotional nonverbal behavior with the neutral behavior for the two groups separately. The results revealed that compared with the neutral behavior, the older users rated the following nonverbal behaviors as more empathic: Dropping the Arms plus Sad Face, Head Down, Sad Face, Head Nod plus Smile, and Smile. Statistical and descriptive values are presented in Table 5 and Table 7.

Table 5

Statistical values of the contrasts for the older users (empathy)

Nonverbal behavior	<i>F</i>	<i>df</i>	<i>p</i>	η^2
Dropping the Arms plus Sad	10.93**	1,29	.003	.27
Head Down	8.71**	1,29	.006	.23
Sad Face	7.31*	1,29	.011	.20
Head Nod plus Smile	6.80*	1,29	.014	.19
Smile	5.77*	1,29	.023	.16

Head Nod	1.02	1,29	.321	.03
----------	------	------	------	-----

p <.05. ***p* <.01. ****p* <.001

Additionally, contrasts revealed that compared with the neutral behavior the younger participants rated the following nonverbal behaviors as more empathic: Dropping the Arms plus Sad Face, and Head Down. Table 6 and Table 7 presents the results and descriptive values obtained from the performed contrasts for the younger users.

Table 6

Statistical values of the contrast for the younger users (empathy)

Nonverbal behavior	<i>F</i>	<i>df</i>	<i>p</i>	η^2
Dropping the Arms plus Sad	9.62**	1,29	.004	.25
Head Down	9.05**	1,29	.005	.24
Sad Face	.63	1,29	.433	.02
Head Nod plus Smile	3.56	1,29	.069	.11
Smile	.515	1,29	.479	.02
Head Nod	3.91	1,29	.057	.12

p*<.05. *p*<.01. ****p*<.001

Table 7

Descriptive values of the perceived nonverbal behaviors for the older and younger users (empathy)

Nonverbal behavior	Old		Young	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Dropping the Arms plus Sad	3.48	.95	2.73	.92
Head Down	3.43	1.03	2.93	1.11
Sad Face	3.34	.94	2.56	1.08
Head Nod plus Smile	3.24	.96	2.70	1.11
Smile	3.24	1.03	2.55	.92
Head Nod	2.83	1.09	2.62	.96
Neutral	2.93	1.11	2.48	.93

RQ 2 was concerned with the overall evaluation of the participants of the agent, by rating the agent’s friendliness, intelligence, trustworthiness, and helpfulness. The results showed no significant effect of age on rating the agent’s friendliness, $F(1, 58) = 1.96, p = .166, \eta^2 = .03$, intelligence, $F(1, 58) = 2.66, p = .108, \eta^2 = .04$, and helpfulness, $F(1, 58) = .74, p = .391, \eta^2 = .01$. This suggests that there were generally similar interpretations among the two groups of how friendly, intelligent, and helpful the agent was. Table 8 compares the descriptive values of older and younger users’ evaluation of the agent’s with regard to these three items.

Table 8

Descriptive values of the perceived nonverbal behaviors for the older and younger users (friendly, intelligent, and helpful)

Items	Old		Young	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Friendly	3.48	.17	3.14	.17
Intelligent	3.21	.17	2.81	.17
Helpful	3.44	.17	3.23	.17

However, the age of the users was found to have a significant effect on rating the agent as trustworthy, $F(1, 58) = 4.65, p = .035, \eta^2 = .07$. Accordingly, older people rated the agent as more trustworthy than younger adults. Figure 6 presents the descriptive values of the two groups on rating the agent’s trustworthiness.

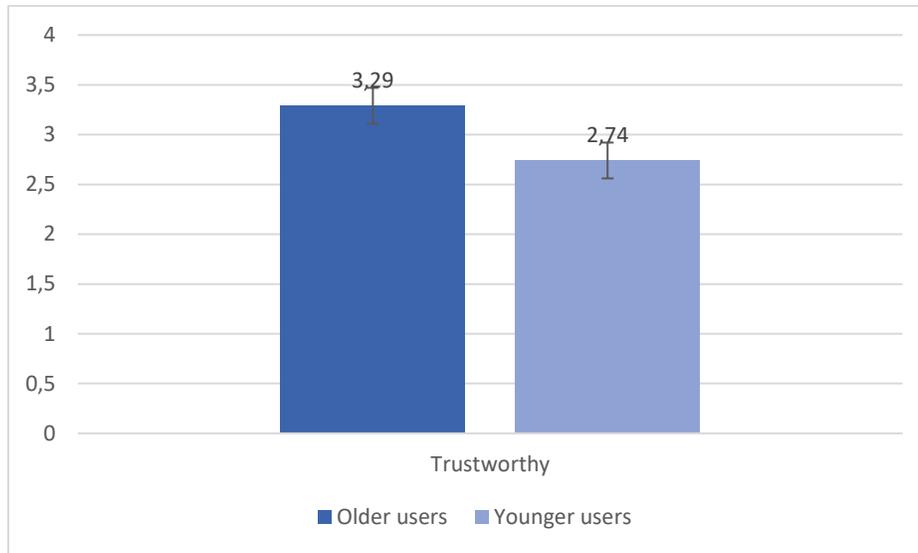


Figure 6: Descriptive values of rating the item trustworthy for older and younger users

Moreover, there was a significant effect of nonverbal behaviors on the rating of all participants (independent of age) on the trustworthiness of the agent $F(3,174) = 2.97, p = .043, \eta^2 = .05$. This means that the participants perceived the agent as trustworthy when they observed it showing nonverbal behaviors. Moreover, the data revealed a significant interaction effect of age and nonverbal behavior on rating the trustworthiness of the agent, $F(3,174) = 4.53, p = .008, \eta^2 = .07$. This indicates that older people and younger adults rated different nonverbal behaviors as trustworthy.

8.4 Additional Analyses

In order to gain in-depth knowledge of the effect of the nonverbal behaviors on the users' perception of the trustworthiness, additional analyses were conducted to test which emotional nonverbal behaviors influenced users' rating of this item. Accordingly, contrasts were performed to compare each emotional nonverbal behavior with the neutral behavior for the item trustworthy. The results showed no significant effect of any specific nonverbal behavior that affected the perception of trustworthiness across all participants. Statistical and descriptive values are shown in Table 9 and Table 10.

Table 9

Statistical values of the contrasts for the perceived nonverbal behaviors – independent of age (trustworthy)

Nonverbal Behaviors	<i>F</i>	<i>df</i>	<i>p</i>	η^2
Dropping the Arms plus Sad Face	3.23	1,59	.077	.05
Head Down	3.10	1,59	.083	.05
Sad Face	.18	1,59	.669	.01
Head Nod plus Smile	.32	1,59	.571	.01
Smile	.50	1,59	.479	.01
Head Nod	2.40	1,59	.126	.04

Table 10

Descriptive values of the perceived nonverbal behaviors – independent of age (trustworthy)

Nonverbal Behaviors	<i>M</i>	<i>SD</i>
Dropping the Arms plus Sad Face	3.10	1.09
Head Down	3.10	1.10
Sad Face	3.00	1.09
Head Nod plus Smile	3.02	1.10
Smile	3.02	1.05
Head Nod	2.91	1.05
Neutral	2.98	1.01

Furthermore, as mentioned earlier, a significant interaction effect was observed between the age group and nonverbal behaviors on perceiving the item trustworthy. To break down this interaction, contrasts were performed to compare each emotional nonverbal behavior with the neutral behavior for each group separately. The results demonstrated that older people perceived the agent as more trustworthy when it displayed Dropping the Arms plus Sad Face, and Head Down. Statistical and descriptive values for older users are shown in Table 11 and Table 13.

Table 11

Statistical values of the contrast for the older users (trustworthy)

Nonverbal Behaviors	F	df	p	η^2
Dropping the Arms + Sad Face	5.24*	1,29	.030	.15
Head Down	6.58*	1,29	.016	.18
Sad Face	3.91	1,29	.057	.12
Head Nod + Smile	2.49	1,29	.125	.08
Smile	3.52	1,29	.071	.10
Head Nod	2.26	1,29	.143	.07

* $p < .05$. ** $p < .01$. *** $p < .001$

Performing the same contrasts for younger adults indicated that they rated the agent as more trustworthy when they observed the neutral behavior of the agent compared with the Sad Face. The results are presented in Table 12 and Table 13.

Table 12

Statistical values of the contrasts for the younger users (trustworthy)

Nonverbal Behaviors	F	Df	p	η^2
Dropping the Arms plus Sad Face	.13	1,29	.712	.01
Head Down	.00	1,29	1.00	.00
Sad Face	5.11*	1,29	.031	.15
Head Nod plus Smile	.91	1,29	.348	.03
Smile	1.76	1,29	.194	.06
Head Nod	.38	1,29	.541	.01

* $p < .05$. ** $p < .01$. *** $p < .001$

Table 13

Descriptive values of the perceived nonverbal behaviors for the older and younger users (trustworthy)

Nonverbal behavior	Old		Young	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Dropping the Arms plus Sad	3.43	1.04	2.76	1.06
Head Down	3.41	1.07	2.78	1.05
Sad Face	3.35	1.02	2.65	1.07
Head Nod plus Smile	3.34	1.04	2.70	1.09
Smile	3.36	1.02	2.68	1.00
Head Nod	3.07	1.02	2.75	1.08
Neutral	3.17	1.00	2.78	.99

Additionally, in order to test whether there is a correlation between older people's higher rating of empathy and their higher rating of trustworthy, a linear regression analysis was conducted. The results indicated that older people's rating of trustworthiness can be predicted by their rating of the item empathy by 82%, $R^2 = .82$, $F(1, 29) = 13.15$, $p < .001$, $\beta = .91$, $p < .001$. This suggests that the empathic responses of the agent can indeed increase the older users' perception of the agent's trustworthiness.

8.5 Discussion

The role of empathy in making the human-agent interaction more believable and lifelike is of great importance. Therefore, having knowledge of how empathic agents would be perceived by different groups of users can contribute to enhancing the quality of users' experience of a virtual agent. ToM is indicated to be important in understanding empathy and its development can be affected by factors such as age (Moran, 2013). Therefore, in Study 1, two target groups of older people and younger adults were selected to investigate whether their understanding of an empathic agent differs according to their age.

The focus of this study was specifically on the nonverbal behaviors of the virtual agent and how they can contribute to showing that the agent is empathizing with the users. Therefore, it

was aimed to compare the perception of older and younger users of the emotional nonverbal behaviors of a virtual assistant and examine which behaviors elicit a higher perception of empathy for these two groups.

The findings indicated that when the agent showed emotional nonverbal behaviors, the older people perceived the agent as more empathic compared with younger adults, supporting H1. It was, however, important to see whether the perception of the agent was based on the observed nonverbal behaviors or was due to the overall positive rating of the agent. An analysis of the data showed the significant effect of the nonverbal behaviors on all participants' perceptions, and, also, the significant interaction effect between age and nonverbal behaviors. These indicate that (1) all participants rated the agent based on the nonverbal behaviors and (2) this rating was different between the older and young people (confirming H1.1 and H1.2). This suggests that the seniors indeed noticed the differences between the nonverbal behaviors. That is why they rated the agent as more empathic when it showed appropriate nonverbal behavior as a response to the emotional situation.

The ability to notice the differences, specifically the emotional facial expressions (which are more subtle nonverbal behaviors compared with the body movement of the agent) and perceive them as empathic responses might be due to seniors' accumulated experience of emotion recognition, as proposed by the sociocognitive approach (see Phillips et al., 2002). These findings suggest that seniors had a more understanding that the agent was empathizing with the users, compared with younger adults. This is in line with the findings of studies that reject the assumption of changing the ability of ToM and understanding emotions in late adulthood (Keightley et al., 2006; MacPherson et al., 2002; Saltzman et al., 2000; Slessor et al., 2007).

With regard to RQ1, it aimed to investigate which specific nonverbal behaviors are perceived as more empathic by the users. The findings illustrated that compared with the neutral behavior, the older users perceived Dropping the Arms plus Sad Face, Head Down, Sad Face, Head Nod plus Smile, and Smile as more empathic. The most empathically perceived nonverbal behaviors for the younger participants were Dropping the Arms plus Sad Face, and Head Down. One of the possible explanations for these findings might be that the context-related emotional responses were more associated with the feeling of empathy for seniors. This might come from their broader experience of seeing emotional responses, which can lead to a more precise

interpretation of emotions. Another plausible explanation can be that since young people probably have more experience with technology, they might have felt the need to see more advanced and sophisticated emotional nonverbal behaviors to rate the agent as more empathic. However, this assumption needs to be further investigated by controlling for users' experience with technology.

Furthermore, RQ2 was concerned with knowing whether the overall evaluation of the agent differs between older people and younger adults in the presence of emotional responses. Therefore, the participants' ratings of the items friendly, intelligent, and trustworthy were analyzed. Although the age of the participants was not found to have a significant effect on perceiving the agent as friendly, intelligent, and helpful, in the presence of emotional nonverbal behaviors seniors rated the agent as more trustworthy compared with their younger peers.

Additional calculations indicated a significant correlation effect between the perception of empathy and perceiving the agent as trustworthy. This makes the findings of this study consistent with previous findings, suggesting that emotional agents engender the perception of trustworthiness in users (Brave et al., 2005). However, in this study, it was found that this perception is remarkably higher for older users only, and younger adults preferred a more neutral response. It seems possible that these results are due to seniors' lower experience with technology. In this regard, their probably lower experience of using and interacting with technology could have led them to feel more insecure when interacting with a virtual agent. Therefore, observing the agent showing emotional responses probably made them feel less patronized, and, as a result, they rated the agent as more trustworthy. By contrast, younger people preferred the neutral behavior in order to perceive the agent as trustworthy. This might be explained by their more extensive experience with technology, making them have less stress in an encounter with a virtual agent. Also, as mentioned in the case of empathy, there is a possibility that younger people expected a more sophisticated display of emotional responses by the agent.

As stated earlier, creating the impression of empathy is especially important in human-agent interactions, since previous findings suggest that it increases the satisfaction of the users (Prendinger et al., 2005b) and the feeling that the agent is "caring" (Brave et al., 2005). The findings of this study can contribute to a better understanding of the users' preferences and the expectations they have of the empathic responses of a virtual agent. The usage of virtual

assistants can be especially beneficial for seniors in order to enable them to be autonomous and independent in their home environment for as long as possible. Therefore, the studies that take these users into account can provide insights into investigating which specific issues should be increased or decreased for this target group in an interaction with a virtual assistant. In this respect, the results of this study suggest that in order to make the impression of empathy stronger for older users, and to enhance their feeling of trust in an agent, implementing emotional nonverbal behaviors in the agent can be taken into account.

8.6 Conclusion

Study 1 was designed to determine the effect of age on users' perception of a virtual assistant based on its emotional nonverbal behaviors. It was shown that age indeed has an influence on users' perception of the intention of the agent to be empathic. Moreover, it was found that certain nonverbal behaviors with varying intensities affect the perception of older people and younger adults differently. Furthermore, the results of this study indicated that an agent that shows emotional responses nonverbally leads to a higher perception of trustworthiness in older people, and this can be a side effect of perceiving the agent as empathic. Engendering trust in seniors is of great importance in interactions with a virtual assistant. This kind of agent will be involved in the everyday lives of these users, who probably do not have much experience with technology. Therefore, making them trust the agent can maximize the probability of further usage of the agent for this target group.

8.7 Limitations and Future Work

The generalizability of these results is subject to certain limitations. The focus of this study was on the perception of nonverbal behaviors. However, in order to create the feeling that the agent is empathizing with the users, the use of both verbal and nonverbal responses could be more effective. It is thus suggested that future studies investigate the combination of these two factors on users' perception of empathy.

Additionally, users' experience with technology and their confidence in dealing with novel technologies were not controlled for in this study. It is therefore recommended that further

investigations take this factor into account. And, finally, there is a possibility that older people need more attention and concentration to perceive subtle nonverbal behaviors when encountering an agent in an interactive setting. Further studies that take interactive scenarios into account will need to be conducted.

9 Study 2: Lost in Interpretation? The Role of Culture on Rating the Emotional Nonverbal Behaviors of a Virtual Agent

The first study of this dissertation investigated the potential differences between people of diverse age groups in perceiving the empathic virtual agent. However, in that study, cultural differences that might affect users' perception of an empathic agent have not been taken into account. Culture and sociocultural dimensions of individualism/collectivism have the potential to influence the ability of ToM in people (Greenfield et al., 2003; Nisbett, 2007; & Shahaieian et al., 2011), and hence, understanding the emotional states of others. Therefore, Study 2 examined whether the cultural background of the users from individualistic and collectivistic countries (though not on the extreme spectrum of individualism and collectivism) affects their perception of an emotional agent. Accordingly, young participants from Germany and Iran, as representatives of individualistic and collectivistic societies were selected in this study. Same as in Study 1, it was attempted to examine which presented emotional nonverbal behavior engenders the perception of empathy in the two groups. The between-group difference in rating the other four items of friendliness, intelligence, trustworthiness, and helpfulness have again been examined as the overall evaluation of the agent.

9.1 Hypotheses and Research Questions

As stated in Chapter 7, many attempts have been made to test the cultural effects when implementing nonverbal behaviors in virtual agents (e.g. Endrass et al., 2011 & Jan et al., 2007). However, no previous study has investigated the effect of users' culture on perceiving an empathic agent based on its nonverbal behaviors. Considering the cultural differences in Germany and Iran (see Chasiotis et al., 2006; Hofstede, 2019a & b; Shahaieian et al., 2014; Triandis, 1993), and the effect of individualistic and collectivistic values on the development and function of ToM, it is assumed that people from these two cultures perceive the emotions of others differently. Hence, in an encounter with an emotional virtual agent, different user perception of the agent is predicted:

H1: The perception of the empathic agent is different between Iranian and German users.

Furthermore, nonverbal behaviors are reported to be effective in expressing the meaning and the purpose of the interlocutor (Peterson & Leonhardt, 2015). Also, perception and judgment of others' level of empathy can be facilitated by observing nonverbal behaviors (Tepper & Haase, 1978). Therefore, it is assumed that the agent's displayed nonverbal behaviors affect users' rating of the item empathy:

H1.1: Perceiving the agent as empathic is due to observing the agent showing emotional nonverbal behaviors as responses to emotional situations.

It is also presumed that cultural differences between Iranian and German users make them rate different nonverbal behaviors as more or less empathic:

H1.2: There is an interaction effect of the culture of the users \times the nonverbal behaviors in perceiving the virtual agent as empathic.

Additionally, one of the goals of this study was to determine the preferences of the two groups regarding specific nonverbal behaviors. Having knowledge about the differences could contribute to the design of (if needed) culture-specific agents that utilize certain emotional nonverbal behaviors in interaction with users from diverse cultures:

RQ1: In case of the existence of a between-subject effect, which emotional nonverbal behaviors are perceived as more empathic by German and Iranian participants?

Finally, same as in Study 1, in order to distract the attention of the participants to the main purpose of the study (perception of agent's empathic behavior), they were presented with four

extra items to rate: agent's friendliness, intelligence, trustworthiness, and helpfulness. The rating of these items is considered as the overall evaluation of the participants of an emotional agent:

RQ2: Do Iranian and German users' ratings of the item friendly, intelligent, trustworthy, and helpful differ, when they observe the emotional nonverbal behaviors of the agent?

9.2 Method

The design of the study was a 2 between-subject factors (culture: Iranians vs Germans) × 2 within-subject factors (situations: sad situations vs happy situations) × 4 within-subject factors (nonverbal behaviors for each situation: 3 emotional vs 1 neutral) mixed factorial design. It was attempted to examine whether users' evaluation of an emotional agent is influenced by their different cultural background.

9.2.1 Material

The same emotional utterances and the nonverbal behaviors that were used in Study 1 (see Chapter 8.2.1) were utilized in this study as well. Since the materials were in German in Study 1, in this study they were translated into Persian for the Iranian participants. To ensure the consistency of all translations, the materials were also back-translated to German.

The virtual agent Billie, which is a cartoon-like character with a neutral appearance, has a white skin tone, dark hair, and no specific features of the race (see Figure 2). Therefore, it was ensured that neither German nor Iranian participants could have the impression that the agent has a similar ethnicity. Moreover, the participants had no prior encounter with the agent, as well as a knowledge of where the agent was designed and produced in.

The videos in which the agent introduce itself and review the imaginary monthly schedule of the participants were dubbed into Persian by a native Persian speaker. In order to adjust the dubbed voice to the German voice of Billie and to manipulate the voice pitch, frequency, and

speech rate of the dubbed utterances, version 2.3.3 of Audacity(R) recording and editing software² was utilized

9.2.2 Sample

Two hundred participants (Iranians $N = 100$, 46 males and 54 females, Germans $N = 100$, 45 males and 55 females) were recruited online for this study. The Iranian participants were between the age of 17 to 39 years ($M = 29$, $SD = 4.0$) and the German participants were between the age of 18 to 43 years ($M = 24$, $SD = 5.6$). Prior to commencing the study, ethical clearance was sought from the ethics committee of the University of Duisburg-Essen. Participation in this study was voluntary and anonymous.

9.2.3 Procedure

On the first page of the survey, the participants were briefed on the study and instructed with regard to the procedure. The first video, similar to Study 1, was an introductory video by the agent, in which it introduced itself and its role. Afterward, the other 24 videos including emotional utterances and nonverbal responses were presented to the participants. The duration of each video on average was 12 seconds. Immediately after watching each video, the participants were asked to rate the agent by filling in the questionnaires. At the end of the study, the participants were debriefed on the purpose of the study and thanked for the participation. Then, they were given an option to take part in a raffle of Amazon voucher.

9.2.4 Measure

Independent Variables: *Groups with cultural differences.* To investigate the effect of users' culture on perceiving the agent as empathic two groups of young people from Iran and Germany were recruited in this study. All participants were presented with similar videos of Billie uttering emotional situations and responding with emotional or neutral behaviors. However, to avoid sequence effect, the order of the videos was randomized and each participant was randomly assigned to each of these video orders.

² Audacity® software is copyright © 1999-2019 Audacity Team.
The name Audacity® is a registered trademark of Dominic Mazzoni.

Dependent variables: *Questionnaires with Five Items (friendly, intelligent, empathy, trustworthy, and helpful).* To test whether displaying nonverbal behaviors affect the perception of the agent and to investigate which emotional nonverbal behaviors are perceived as more or less empathic, five items were presented to the participants (“How do you rate Billie?”). The rating was on a 5-point Likert Scale from “I totally disagree” (score = 1) to “I totally agree” (score =5). The participants needed to rate that “I feel Billie is:” friendly, intelligent, empathic, trustworthy, and helpful.

9.3 Results

To analyze the users’ ratings of the agents’ friendliness, intelligence, empathy, trustworthiness, and helpfulness multiple mixed ANOVAs were conducted for each item separately. Furthermore, in order to analyze the interaction effects between the between and within-factors, repeated ANOVAs were conducted for each target group separately. Also, Post-hoc test using Bonferroni correction was used in all cases. When the assumption of sphericity was violated, Greenhouse-Geisser correction was reported.

Based on the assumption of H1, when the agent shows emotional responses, Iranian and German users who have different cultures perceive the agent differently with regard to being empathic. The results of the study supported this hypothesis, by showing a significant main effect of the culture on rating the agent as empathic, $F(1, 198) = 22.85$ $p < .001$, $\eta^2 = .10$. Descriptive values indicated that Iranian participants rated the agent as more empathic compare with their German peers (see Figure 7).

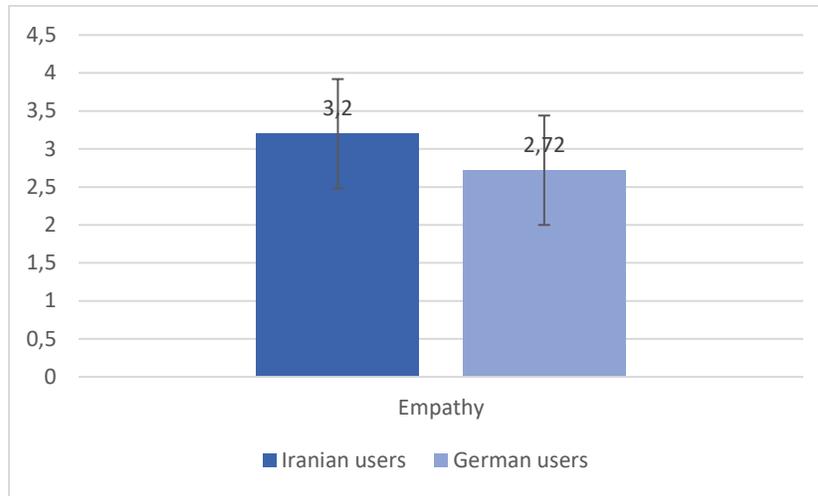


Figure 7: Descriptive values of rating the item empathy for Iranian and German users

As H1.1. stated, nonverbal behaviors of the agent are assumed to have a role in users' rating of the agent as empathic. Results of the study indicated a significant effect of the nonverbal behaviors on rating the item empathy, $F(2.42, 479.4) = 29.22, p < .001, \eta^2 = .13$, confirming H1.1. This suggests that perceiving the agent as empathic was affected by observing the nonverbal behaviors across all participants. Additionally, contrasts were conducted to compare each emotional behavior with the neutral behavior for all participants. The purpose was to test which nonverbal behaviors were perceived as more empathic by all users. The results showed that compared with the neutral behavior, all emotional nonverbal behaviors affected the perception of empathy (see Table 14 and Table 15).

Table 14

Statistical values of the contrasts for the perceived nonverbal behaviors – independent of culture (empathy)

Nonverbal behavior	<i>F</i>	<i>df</i>	<i>p</i>	η^2
Smile	17.14***	1,199	<.001	.08
Head Nod plus Smile	37.04***	1,199	<.001	.16
Head Nod	8.74*	1,199	.003	.04
Sad Face	18.92***	1,199	<.001	.09
Head Down	37.54***	1,199	<.001	.16
Dropping the Arms plus Sad	36.66***	1,199	<.001	.16

* $p < .05$. ** $p < .01$. *** $p < .001$

Table 15

Descriptive values of the perceived nonverbal behaviors - independent of culture (empathy)

Nonverbal behavior	<i>M</i>	<i>SD</i>
Dropping the Arms plus Sad	3.12	.97
Head Down	3.15	.96
Sad Face	3.01	1.0
Head Nod plus Smile	3.08	.93
Smile	2.97	.83
Head Nod	2.87	.87
Neutral	2.75	.80

As Hypothesis 1.2. stated, an interaction effect of culture and nonverbal behaviors is expected. This is due to the influence that culture might have on understanding the emotions of others and the interpretation of specific nonverbal behaviors. Results revealed a significant interaction effect of culture and nonverbal behaviors, $F(2.42, 479.4) = 28.77, p < .001, \eta^2 = .13$, supporting H1.2. To break down this interaction and answer the RQ1 (Which emotional nonverbal behaviors are perceived as more empathic by Iranian and German users?) contrasts were performed to compare each emotional behavior with the neutral behavior. It was observed that all emotional nonverbal behaviors influenced the Iranian users' rating of the agent as empathic. The statistical and descriptive values are shown in Table 16 and Table 18.

Table 16

Statistical values of the contrasts for the Iranian users (empathy)

Nonverbal behavior	<i>F</i>	<i>df</i>	<i>p</i>	η^2
Smile	53.26**	1,99	.001	.35
Head Down and Smile	54.26**	1,99	.001	.36
Head Nod	4.31*	1,99	.04	.36
Sad Face	50.96**	1,99	.001	.34
Head Down	71.89***	1,99	<.001	.42
Dropping the Arms plus Sad	58.87***	1,99	<.001	.37

* $p < .05$. ** $p < .01$. *** $p < .001$

However, performed contrasts for German participants indicated the significant effect of only one nonverbal behavior; Head Nod on rating the item empathy (see Table 17 and Table 18).

Table 17

Statistical values of the contrasts for the German users (empathy)

Nonverbal behavior	<i>F</i>	<i>df</i>	<i>p</i>	η^2
Smile	2.38	1,99	.126	.02
Head Down and Smile	.94	1,99	.335	.01
Head Nod	4.46*	1,99	.037	.04
Sad Face	2.62	1,99	.109	.03
Head Down	.81	1,99	.369	.09
Dropping the Arms plus Sad	.08	1,99	.779	.01

* $p < .05$. ** $p < .01$. *** $p < .001$

Table 18

Descriptive values of the perceived nonverbal behaviors for the Iranian and German users (empathy)

Nonverbal behavior	Iranians		Germans	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Dropping the Arms plus Sad	3.50	.93	2.73	.86
Head Down	3.51	.85	2.79	.94
Sad Face	3.41	.95	2.61	.88
Head Nod plus Smile	3.40	.83	2.77	.93
Smile	3.33	.71	2.62	.78
Head Nod	2.91	.82	2.82	.93
Neutral	2.79	.76	2.71	.85

Furthermore, RQ2 asked whether there is a difference between Iranian and German users with regard to the overall evaluation of the agent. Accordingly, the two groups' rating of the items friendly, intelligent, trustworthy, and helpful were compared. Results indicated no significant between-group effect on rating the item friendly, $F(1, 198) = 3.11$, $p = .08$, $\eta^2 = .01$ ($M_{\text{Iranians}} = 3.45$, $SE_{\text{Iranians}} = .08$, $M_{\text{Germans}} = 3.26$, $SE_{\text{Germans}} = .08$) and intelligent, $F(1, 198) = 3.29$, $p = 3.298$, $\eta^2 = .07$ ($M_{\text{Iranians}} = 3.04$, $SE_{\text{Iranians}} = .08$, $M_{\text{Germans}} = 2.84$, $SE_{\text{Germans}} = .08$).

However, the results showed a significant main effect of users' culture on the perception of agent's trustworthiness, $F(1, 198) = 18.58, p < .001, \eta^2 = .09$, and helpfulness, $F(1, 198) = 7.46, p = .007, \eta^2 = .04$. Figure 8 compares the descriptive values of the ratings of these two items.

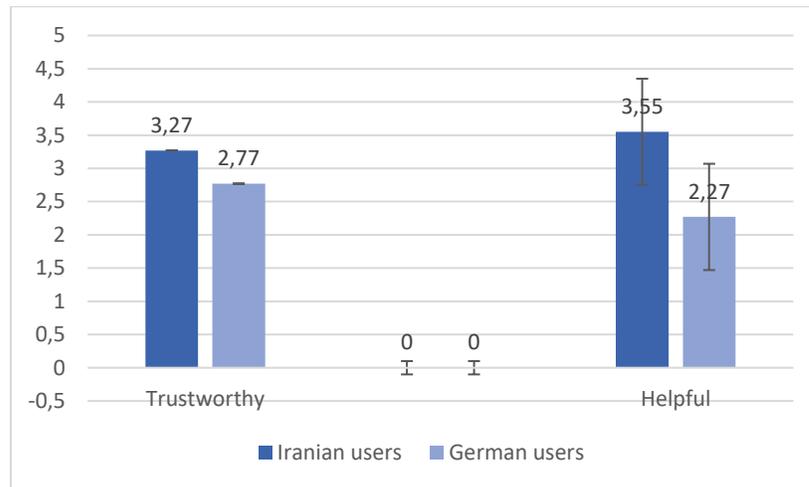


Figure 8: Descriptive values of rating the items trustworthy and helpful for Iranian and German users

It is apparent from Figure 8 that there is a difference in the overall rating of the agent between the two groups. Accordingly, Iranian users evaluated the agent as more trustworthy and helpful than German users. Moreover, the data revealed that across all participants, the nonverbal behaviors had a significant effect on the perception of agent's trustworthiness, $F(2.73, 540.51) = 2.74, p = .048, \eta^2 = .01$, and helpfulness, $F(2.81, 557.29) = 2.74, p = .022, \eta^2 = .02$. This suggests that perceiving the agent as trustworthy and helpful was the result of observing the nonverbal behaviors.

However, the interaction effect of culture and nonverbal behaviors was not found to have a significant effect on users' rating of the item of trustworthy, $F(2.73, 540.51) = .69, p = .54, \eta^2 = .00$. This suggests that the nonverbal behaviors did not affect Iranian and German users' ratings of the agent's trustworthiness differently. By contrast, it was observed that there is a significant interaction effect of culture and nonverbal behaviors on the perception of helpfulness,

$F(2.81, 557.29) = 5.68, p = .001, \eta^2 = .03$. This indicates that the influence of nonverbal behaviors on rating the item helpful was different between Iranian and German participants.

9.4 Additional Analyses

To gain deeper knowledge regarding the effectiveness of the nonverbal behaviors, further analyses were conducted to test which emotional nonverbal behaviors contributed to the perception of the agent’s trustworthiness and helpfulness. These analyses are important since creating a feeling in the users that the agent is trustworthy and is intended to be assistive is one of the main reasons to implement emotional responses in the agent.

Therefore, first, contrasts were performed for all participants and for each item separately, comparing each emotional nonverbal behavior with the neutral behavior. Table 19 provides the results of the performed contrasts for both groups and Table 20 compares the descriptive values of the perceived nonverbal behaviors across all participants for item trustworthy.

Table 19

Statistical values of the contrasts for the perceived nonverbal behaviors – independent of culture (trustworthy)

Nonverbal behavior	<i>F</i>	<i>df</i>	<i>p</i>	η^2
Smile	.00	1,199	1.00	.00
Head Down and Smile	.31	1,199	.576	.01
Head Nod	.19	1,199	.657	.01
Sad Face	1.15	1,199	.283	.01
Head Down	7.50**	1,199	.007	.04
Dropping the Arms plus Sad	.66	1,199	.417	.01

* $p < .05$. ** $p < .01$. *** $p < .001$

Table 20

Descriptive values of the perceived nonverbal behaviors – independent of culture (trustworthy)

Nonverbal behavior	<i>M</i>	<i>SD</i>
Smile	3.01	.91
Head Down and Smile	3.03	.98
Head Nod	3.02	1.0
Sad Face	2.97	.95
Head Down	3.10	.92
Dropping the Arms plus Sad	3.04	.95
Neutral	3.01	.87

As demonstrated in Table 19 and Table 20, when the agent showed the nonverbal behavior Head Down, the participants perceived that the agent is trustworthy. This means that head movement, in the sense of sadness, might attract the attention across all users, and could contribute to the perception of the trustworthiness of the agent.

Additionally, to test which nonverbal behaviors led to the perception of helpfulness across all participants each emotional nonverbal behavior was compared with the neutral behavior by performing contrasts. Yielded data indicated that Smile and Head Down could engender the feeling of helpfulness in the users (see Table 21 and Table 22).

Table 21

Statistical values of the contrasts for the perceived nonverbal behaviors – independent of culture (helpful)

<i>Nonverbal behavior</i>	<i>F</i>	<i>df</i>	<i>p</i>	η^2
Smile	10.20**	1,199	.002	.05
Head Down and Smile	3.40	1,199	.067	.02
Head Nod	.77	1,199	.379	.01
Sad Face	10.06	1,199	.304	.01
Head Down	3.96*	1,199	.048	.02
Dropping the Arms plus Sad	.24	1,199	.624	.01

* $p < .05$. ** $p < .01$. *** $p < .001$

Table 22

Descriptive values of the perceived nonverbal behaviors – independent of culture (helpful)

Nonverbal behavior	<i>M</i>	<i>SD</i>
Smile	3.48	.81
Head Down and Smile	3.44	.84
Head Nod	3.40	.87
Sad Face	3.41	.83
Head Down	3.44	.85
Dropping the Arms plus Sad	3.39	.85
Neutral	3.37	.80

Furthermore, as indicated in Chapter 9.3, there was a significant interaction effect between the culture of the users and rating the item helpful. Therefore, contrasts were performed comparing each emotional nonverbal behavior with the neutral behavior for this item. The purpose was to gain deeper knowledge about nonverbal behaviors which can create the perception of helpfulness in the two groups. Results demonstrated that Iranian users rated the agent as more helpful when it showed Smile, Head Nod plus Smile, Sad Face, and Head Down (see Table 23 and Table 25). However, performing the same contrasts for German participants showed no significant effect of a specific nonverbal behavior on the perception of helpfulness (see Table 24 and Table 25). This suggests that the nonverbal behaviors created for the purpose of showing emotional responses were not effective enough to express the feeling that the agent is helpful for German participants.

Table 23

Statistical values of the contrasts for the Iranian users (helpful)

Nonverbal behavior	<i>F</i>	<i>df</i>	<i>p</i>	η^2
Smile	16.37***	1.99	.001	.14
Head Nod plus Smile	6.88*	1.99	.010	.06
Head Nod	.15	1.99	.697	.01
Sad Face	4.94*	1.99	.028	.05
Head Down	7.60**	1.99	.007	.07
Dropping the Arms plus Sad	1.74	1.99	.189	.02

* $p < .05$. ** $p < .01$. *** $p < .001$

Table 24

Statistical values of the contrasts for the German users (helpful)

Nonverbal behavior	<i>F</i>	<i>df</i>	<i>p</i>	η^2
Smile	.01	1.99	.910	.00
Head Nod plus Smile	.15	1.99	.695	.01
Head Nod	.76	1.99	.385	.01
Sad Face	1.74	1.99	.190	.02
Head Down	.11	1.99	.739	.01
Dropping the Arms plus Sad	.98	1.99	.323	.01

Table 25

Descriptive values of the perceived nonverbal behaviors for the Iranian and German users (helpful)

Nonverbal behavior	Iranians		Germans	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Dropping the Arms plus Sad	3.54	.81	3.23	.86
Head Down	3.62	.80	3.26	.86
Sad Face	3.59	.76	3.22	.85
Head Nod plus Smile	3.62	.81	3.26	.84
Smile	3.69	.75	3.28	.82
Head Nod	3.48	.85	3.32	.89
Neutral	3.46	.83	3.28	.76

Additionally, in order to test whether Iranian users' perceiving that agent is empathic is correlated with their higher rating of the item trustworthy and helpful, a linear regression analysis was performed. This investigation is crucial to see whether in line with the findings in Study 1 and consistent with the findings of previous studies (Brave et al., 2005), empathic agents can engender the feeling of trustworthiness in the users. Furthermore, it could provide useful knowledge that whether users' understanding of the empathizing behaviors of the agent leads to the feeling that the agent is indeed assisting them. This is specifically important in the case of a virtual assistant which will be involved with the daily life of the users. Therefore, users might have a high expectation of agents' assisting level and a need to trust the agent with their daily tasks.

The results indicated a significant correlation between higher rating of empathy and the perception of trustworthiness, indicating that the rating of the item trustworthy in Iranian users could be predicted by their rating of the item empathy by 21%, $R^2 = .21$, $F(1,98) = 26$, $p < .001$, $\beta = .46$, $p < .001$. Moreover, there was also a significant correlation between the higher rating of empathy and the higher rating of helpfulness for Iranian participants, $F(1,98) = 44.17$. $P < .001$, $\beta = .56$, $p < .001$. This suggests that the rating of helpfulness could be predicted by 31%, $R^2 = .31$ by Iranian users' rating of the item empathy.

9.5 Discussion

The usage of supporting technologies such as virtual assistants is accelerating, bringing with it its own hurdles and challenges. Most of these assistants are designed to have social communications to provide more sufficiency and values to the lives of users, such as assisting them with different tasks. This means that there is a high probability that these assistants will be the daily companions of users and be involved in their everyday life. In this regard, coping with users' differences such as their different cultures is specifically of great importance, as the quality of human-agent interactions is dependent on users' understanding of the intention of the agents.

As studies suggest, empathic agents have positive roles in users' evaluations and feeling toward the agent (see e.g., Brave et al., 2005; Prendinger & Ishizuka, 2005b). Understanding empathy and the ability of ToM are closely connected, and cultural differences as stated in individualism and collectivism can influence understanding the intentions of others. Therefore, it is important that researchers conduct cultural comparisons in order to gain insights into users' differences in perceiving and accepting an emotional agent.

Accordingly, Study 2 had two primary aims: (1) to investigate whether there are cultural differences with respect to the perception of a virtual assistant as empathic, and (2) to examine which specific displayed emotional nonverbal behaviors are perceived as more empathic by users from different cultures (with collectivistic and individualistic tendencies). For these purposes, participants from two countries, Germany and Iran, were recruited.

As the findings of this study demonstrated, there is indeed a cultural difference with regard to the perception of an empathic agent, supporting H1. Accordingly, it was shown that when the agent displayed emotional nonverbal behaviors in responses to the uttered emotional situations, Iranian users rated it as more empathic compared with German users. These results also confirmed H1.1 of the study, which stated that participants' rating of the agent as empathic is indeed due to their observation of the emotional nonverbal behaviors of the agent.

The observed differences between the perceptions of the Iranian and German participants could be attributed to the cultural values that exist in Iranian upbringing and norms. As mentioned in Chapter 3.4, Iranian culture promotes interpersonal similarities and harmony and values respect and politeness. These values are transferred to people by means of their upbringing and by teaching them from early childhood to be highly sensitive to the emotions of others in

order to maintain harmony in relationships (Degens et al., 2012). Accordingly, when an interlocutor acts non-responsively, especially in an emotional context, Iranian people might interpret the behavior as impolite, violating their expectation of a harmonious interaction.

Another plausible explanation for these findings may be that the relatively collectivistic values of Iranian users could have an influence on their tendency to care about the needs and emotions of others. Therefore, this might have led them to expect empathic responses from the agent after it expressed the emotional situations. It is thus possible that due to this expectation of Iranian users, they paid more attention to the agent showing emotional nonverbal behaviors, therefore, understood its intention (of empathizing) accurately. However, these data must be interpreted with caution and further studies will need to more explicitly examine whether the existing differences can be explained by the more collectivistic tendency of Iranian users.

Moreover, as the data showed, German participants rated the agent as less empathic, which might be a reflection of the potential individualistic values of German users. As mentioned in Chapter 3.4, members of individualistic cultures have more independent views of others, and less other-oriented sense of well-being (Henrich et al., 2010; Hofstede, 2019b). This might be a possible explanation for the fact that German users did not feel the need for emotional responses from the agent. Also, it is plausible that due to the lack of expectations to receive an emotional response, they might have not understood the intention of the agent accurately. Therefore, German users did not rate the agent as more empathic. However, these assumptions need to be investigated in future studies, by taking into account, for instance, the users' expectations of emotional responses of an agent, and also their inclination to care for other people's needs.

The other purpose of this study was to examine which emotional nonverbal behaviors of the agent engender a greater perception of empathy in Iranian and German users (RQ1). Therefore, the different ratings of the users of the displayed nonverbal behaviors were assessed. As shown in the results, Iranian users rated all emotional nonverbal behaviors as more empathic compared with the neutral behavior (Smile, Head Nod plus Smile, Head Nod, Sad Face, Head Down, and Dropping the Arms plus Sad Face). However, the results of the German users revealed that compared with the neutral behavior, only the nonverbal behavior Head Nod was perceived as more empathic. This means that while Iranian users easily perceived the emotional nonverbal behaviors as empathic, German users rated Head Nod merely as more empathic, which is a nonverbal behavior that could be defined as the most rational behavior for them. Therefore, it

can be stated that the intensity of displaying nonverbal behaviors matters, meaning that users' preferences in this regard might vary depending on their cultural background.

The secondary objective of this study was to answer whether, in the presence of the emotional responses, the evaluation of users from different cultures differ with regard to the agents' friendliness, intelligence, and trustworthiness (RQ2). As the results demonstrated, users did not differ in rating the agent as friendly and intelligent. However, a difference between the users' perception of trustworthiness and helpfulness, based on the displayed nonverbal behaviors of the agent, was observed. On the basis of the findings, Iranian participants evaluated the agent as being more trustworthy and helpful compared with their German peers.

Further analysis indicated that although the two groups did not differ in how they perceived certain nonverbal behaviors as more trustworthy, the following nonverbal behaviors were rated as more helpful by Iranian users: Smile, Head Nod plus Smile, Sad Face, and Head Down. Having knowledge about specific nonverbal behaviors that can trigger the feeling of helpfulness in the users can contribute to designing virtual assistants for users with different cultures.

Furthermore, additional analysis was conducted to test whether there is a correlation between perceiving the agent as empathic and the higher rating by Iranian participants that the agent is trustworthy and helpful. The results indicated a significant correlation effect, meaning that perception of trustworthiness and helpfulness can be interpreted as a side effect of perceiving the agent as empathic. This could be explained by the fact that Iranian users rated the agent as more empathic and thereby as more trustworthy and helpful because they perceived the agent as "caring." However, it should be taken into account that there was no significant interaction effect between culture and the nonverbal behaviors in the rating of trustworthiness, making it important to interpret this explanation with more caution.

In sum, it can be stated that taking into account the culture of the target groups can provide insight into possible differences in understanding emotions and intentions of the agent. Accordingly, it can be demonstrated that perception of certain emotional nonverbal behaviors as more empathic is sensitive to the culture of the users, rendering the factor of culture important in studies that aim to provide a precise and effective design guideline. Considering the enculturation process of different users worldwide, and its influence on users' preferences and expectations of

an agent, designing a culture-specific virtual agent that acts in line with the cultural expectations of its users can lead to higher satisfaction and better experience on the part of the users.

9.6 Conclusion

This study set out to determine how users' cultural background influences their ability to understand others' minds, which, in turn, can impact their perception of whether a virtual agent is empathizing with them. More specifically, in this investigation, the aim was to assess whether the culture of the users affects the perception of certain emotional nonverbal behaviors of a virtual agent as empathic. The findings support the idea that people's culture is connected with the evaluation of emotional nonverbal behaviors. It has been shown that Iranian users, who are reported to have rather collectivistic values, compared with their German peers who have relatively individualistic values, rated the virtual agent as being more empathic. These evaluations were based on the displayed emotional nonverbal behaviors of the agent, making the importance of implementing these behaviors in the agent more relevant. Also, Iranian users' overall evaluation of the agent was more positive in terms of perceiving the agent as more trustworthy and helpful. Taken together, this study has shown that the cultural background of the users matters when confronted with an emotional agent, hence providing insights for designing a culture-specific virtual agent.

9.7 Limitations and Future Work

The findings of this study are subject to some limitations. The users' experience with technology, their prior knowledge of artificial entities, and their confidence with regard to interacting with such supporting technologies have not been controlled for in this study. Further work is required to test the effect of these factors and investigate whether more experience with technology leads to higher expectations of seeing sophisticated nonverbal behaviors.

Moreover, it is important to note the existence of microcultures within one big culture, which can deviate from its overall individualistic and collectivistic values. It is, therefore, crucial that future cultural studies take these variables into account as well. Furthermore, since a combination of verbal and nonverbal responses can lead to a higher perception of empathy in the users, further studies with a focus on these two factors together are therefore suggested.

Additionally, as the design of this study was not based on an interactive setting, it is recommended to investigate whether users' perceptions of empathy differ in an interactive scenario, in which a higher level of attention is required.

It is also important to note that although the appearance of the virtual agent in this study was neutral, looking neither German nor Iranian, its potential effect on users' evaluation of the agent could be considered in future research. And finally, it is recommended that future studies specifically test the skill of ToM in different cultures as well, in order to conclude more precisely that this ability influences users' perception of the empathic agent.

10 Study 3: Investigating Stereotypical Attitudes of Users Toward a Virtual Agent Designed in a Different Country

The first two studies of this dissertation asked the questions about the role of factors such as users' age and culture on their perception of an empathic agent. In the third study, it was attempted to examine how priming the users about the culture of the agent, without any apparent feature related to a specific culture in the agent's appearance, could possibly affect their perception of an empathic agent.

In the first two studies, the effect of factors such as users' experience with technology, their own level empathy, and mood in an encounter with an emotional agent have not been controlled for. Therefore, in Study 3, the moderating effect of these factors on users' perception has been investigated as well. Controlling these factors is crucial since it has been reported that users' prior experience with (new) technologies can influence how they handle encountering and accepting them and can affect their motivation for further usage of those specific technologies (Agarwal & Prasad, 1999; Wilkowska & Ziefle, 2009). Moreover, studies have shown that individual differences with respect to one's own empathic behaviors are associated with the accurate recognition of other's people's mental states and emotions (Baron-Cohen et al., 2001; Dimberg et al., 2011; Svetieva & Frank, 2016). Therefore, in Study 3, it is asked whether a higher level of empathy in participants has a moderating effect on correctly perceiving that the agent is empathizing with them.

And finally, people's mood has also been reported to influence their perception of emotions. Positive or negative mood of people can govern the usage of ToM and inferences about the mental states of others and, therefore, can impact people's judgment of other's emotion and intention (Converse, Lin, Keysar, & Epley, 2008; Juckel, Heinisch, Welpinghus, & Brüne, 2018). The virtual agent in Study 3 was aimed to display emotional responses to look empathic. It was thus important to control whether users' understanding of these responses and more importantly, their recognition of the intention of the agent (empathizing), is influenced by their current mood, besides their bias toward the agent.

Additionally, in order to test whether users tend to rate the agent negatively, unlike the first two studies, attributes of warm/cold, incompetent/competent, inactive/active were presented

to the participants as well. Also, the participants were asked about the appropriateness of the agent's behaviors.

Finally, in contrast to the first two studies, in Study 3, open-ended questions were presented to the participants, in which they could state their opinion of the agent in their own words. The purpose of presenting open-ended questions was to obtain in-depth knowledge regarding the evaluation of the nonverbal behaviors of the agent.

10.1 Hypotheses and Research Questions

As mentioned earlier in chapter 4, in interaction with others, people tend to categorize others as ingroup or outgroup members (Nass et al., 2000). This categorization is usually based on observed cues, such as physical cues indicating the ethnicity of others (Nass et al., 2000). Considering that people tend to apply the same social rules from human-human interactions to human-agent interactions (Reeves & Nass, 1996), users have been reported to show a different attitude and action when interacting with a computer agent with similar ethnicity than an agent from a different one (Pratt et al., 2007). Moreover, it has been shown that people perceive the agent as more positive and acceptable when the agent has a similar ethnicity to their own (Lee & Nass, 1998).

In Study 3, it was strived to investigate whether people also apply the same stereotypes to an agent when they know that the agent is from a different culture (with no physical cues). This is specifically important as the role of unconscious application of stereotypes regarding the culture of others has been indicated to form empathy between users and virtual agents (Aylett et al., 2009; Rossen et al., 2008). Therefore, it is predicted that user' stereotypes of the agent based on their knowledge of its culture can impact their rating of specific emotional nonverbal behaviors of the agent and perceiving it as empathic:

H1: Priming the participants about the country that the agent was designed in leads to their different perceptions of the virtual agent as being empathic.

The nonverbal behaviors are assumed to enhance understanding the intention of others (see e.g., Peterson & Leonhardt, 2015). Hence, it could conceivably be hypothesized that the users perceive the intention of the agent based on its displayed emotional nonverbal behavior:

H1.1: Perceiving the agent as empathic is based on the observed nonverbal behaviors.

Furthermore, considering users' expectations, ingroup advantage, and cultural stereotypes, it can be assumed that they show different perceptions of the nonverbal behaviors of an agent from a different culture. Therefore, an interaction effect between the groups and nonverbal behaviors is expected, meaning that they rate specific nonverbal behaviors as more or less empathic:

H1.2: There is an interaction effect of the groups \times nonverbal behaviors, suggesting that different nonverbal behaviors lead to different perceptions of empathy in the two primed groups.

Additionally, based on the assumed differences between the two groups' perception of an empathic agent, it was aimed to investigate which specific nonverbal behaviors are perceived as more or less empathic by the two primed groups. Therefore, the first research question of this study is concerned with knowing:

RQ1: In case there is a significant between-group difference, which emotional nonverbal behaviors are perceived as more empathic by the two primed groups?

Similar to the first two studies of this dissertation, extra items were presented to the participants in order to avoid attracting their attention to the main purpose of the study (evaluating empathy). These items in this study were the agent's unfriendliness and reactivity.

Rating these items was considered as the overall evaluation of the agent. Therefore, the second research question in this study asked:

RQ2: Do the two primed groups perceive the agent as unfriendly and reactive differently, when the agent shows emotional nonverbal behaviors?

In order to test whether priming the groups affects the perception of users regarding other emotional aspects, attributes of warmhearted/cold, incompetent/competent, and inactive/active were presented to the participants to rate. Using attributes contribute to the understanding of whether participants tend to rate the agent positively or negatively when they have the knowledge of its design country. It is thus asked whether users' ratings of these attributes change based on knowing that the agent is from a different culture:

RQ3: How do the primed groups perceive the agent regarding the attributes of warmhearted/cold, incompetent/competent, and inactive/active in the presence of the emotional nonverbal behaviors?

Moreover, since the agent shows different emotional nonverbal behaviors in response to different emotional situations, participants who have been primed about the culture of the agent might find the agent's behavior as more or less appropriate. Therefore, the hypothesis that will be tested is:

H2: There is a difference between the two primed groups' ratings of the appropriateness of the agent's behavior.

Additionally, users' evaluation of the appropriateness is predicted to be due to the observed nonverbal behaviors, similar to the evaluation of the aforementioned items and attributes:

H2.1: It is predicted that users' evaluation of whether the agent's behavior is appropriate or not is based on the presence of different nonverbal behaviors of the agent.

Furthermore, the interaction effect between the nonverbal behaviors and the primed group is predicted. This means that there is a different rating of specific nonverbal behaviors as more or less appropriate between the two groups:

H2.2: There is an interaction effect of the primed groups \times the nonverbal behaviors in rating the appropriateness of the agent's behavior.

As mentioned earlier in Chapter 10, there is a possibility that extra factors, besides priming the groups, have a moderating effect on users' perception of the virtual agent as empathic. The factors that are included in this study are the mood of the participants (positive or negative affect), their experience with technology, and also their own level of empathy. Accordingly, it was aimed to test whether high or low empathic users perceive the agent as more empathic when it shows emotional nonverbal behaviors. Therefore, the fourth research question of the study asked:

RQ4: Is there a moderating effect of the participants' mood, experience with technology, and their level of empathy on rating the agent as empathic when they are primed about the design country of the agent?

Furthermore, in order to test the participants' own opinion of the agent, before presenting them the items and the attributes to rate, they were asked to give their opinion and impression of the agent in an open-ended question. The answers to this question could contribute to the knowledge that whether the users have indeed the impression that the agent is empathizing with them. Therefore, the fifth research question of the study asked:

RQ5: (Qualitative) What are the users' evaluation and understanding of the agent in their own words when they observe the agents' emotional nonverbal responses?

Future usage of the agent is an important goal in designing virtual assistants for different user groups. It is thus crucial to investigate whether knowing that the agent is from a different culture affects the users' decision of utilizing that agent. Based on the assumed ingroup advantage and the stereotypes that the users might have toward an agent from a different culture (see e.g., Clark & Maass, 1988; Lee, 1993) it is hypothesized:

H3: Priming the groups about the design country of the agent affects their rating of the further usage of the agent.

And finally, in order to test whether the answers are based on knowing that the agent is from a different culture or other reasons, in an open-ended question the participants were asked to explain their given answers:

RQ 6: (Qualitative) What are the participants' reasons for either using or not using the agent in the future?

10.2 Method

In Study 3, a laboratory experiment with a 2 between-subjects factors (primed groups: the agent is designed Germany vs the agent is designed in Malaysia) \times 2 within-subjects factors (situations: sad vs happy) \times 4 within-subjects factors (nonverbal behaviors for each situation: 3 emotional vs 1 neutral) was conducted. The aim of this study was to test the effect of priming the groups about the design country of the agent on the perception of the agent. The users' evaluations of the agent were measured in the presence and the absence of the emotional nonverbal behaviors

10.2.1 Material

The materials of Study 1 -including videos of the virtual agent Billie expressing emotional situations and displaying nonverbal responses (see Chapter 8.2.1) were used in this study.

10.2.2 Sample

A total number of 108 native Germans (females = 74, males = 34) between the age of 18 to 40 years old ($M = 22$, $SD = 3.64$) participated in this study, in a lab of the University of Duisburg-Essen. In order to avoid any minor cultural effect on the data, only people who declared that are native German were included in the measurement. All participants agreed to participate in the study by reading the consent page at the beginning of the measurement. The study was approved by the Ethics Committee of the University of Duisburg-Essen. Most of the participants were students. The duration of the study was between 35 to 45 minutes and as a "thank you" for their participation, the participants were offered the option of choosing either 5 Euro or test person certification which was required for their studies.

10.2.3 Procedure

The first page of the survey was the welcoming page which included the briefing of the study as well. On this page, the participants were informed about the upcoming videos of the virtual agent, Billie, in which the agent would act as an assistant to review their imaginary calendar for upcoming appointments and events. They were further instructed to rate this agent's

behavior. Then, the participants needed to fill in their social demography. Afterward, the cover stories of the study were randomly presented to the participants (see Measures). Then, the participants watched the introductory video by Billie (in which Billie introduced itself and explained its role). Next, the remaining 24 videos and questionnaires were presented. Afterward, all participants had to fill in the QCAE, PANAS, and STC questionnaires (see Measures). At the end of the study, the participants were thanked for the participation in the study and the debriefing page appeared that explained the aim of the study.

10.2.4 Measures

Independent variables:

Primed groups. In order to prime the participants, two cover stories were created. The cover story 1 stated that:

*“Billie is a virtual agent that is designed and produced in **Malaysia**. This agent is designed to act as a virtual assistant and is widely accepted by **Malaysian users**. Furthermore, scientists are working on enabling this virtual agent to act independently. In this study, we would like to know your opinion on the emotional nonverbal behaviors of the virtual agent Billie.”*

The cover story 2 stated that:

*“Billie is a virtual agent that is designed and produced in **Germany**. This agent is designed to act as a virtual assistant and is widely accepted by **German users**. Furthermore, scientists are working on enabling this virtual agent to act independently. In this study, we would like to know your opinion on the emotional nonverbal behaviors of the virtual agent Billie.”*

The participants were randomly assigned to one of the groups (German designed agent, $N = 53$ or Malaysian designed agent, $N = 55$).

Dependent variables:

Evaluating the items (unfriendly, empathy, and reactive)

In order to test users' perception of the virtual agent regarding its unfriendliness, empathy, and reactivity, after watching each video, the participants were asked to fill in a blank box. In this box, they had to type their impression of the agent ("What do you think that Billie is trying to display?"). Next, there was a questionnaire with a 5-point Likert scale to rate three different items ("How do you rate Billie?"): "I feel Billie is:" Unfriendly, Empathic, and Reactive. The scales to rate the items were ranging from "I totally disagree" (lowest score = 1) to "I totally agree" (highest score = 5) for empathy and reactive. For the item unfriendly the range was from "I totally disagree" (highest score = 5) to "I totally agree" (lowest score = 1).

Evaluating the attributes (warmhearted/cold, incompetent/competent, and inactive/active)

On the next page of the study, participants were asked to rate Billie on attribute scales: "I have the feeling that Billie is:" cold (score = 1)/warmhearted (score = 5), incompetent (score = 1)/competent (score = 5), and inactive (score = 1)/active (score = 5).

Evaluating the appropriateness of the agent's behavior

Then, the next questionnaire appeared in which the participants were asked to evaluate the appropriateness of Billie's behavior ("Do you find Billie's behavior appropriate?") on a scale from "totally inappropriate" (score = 1) to "totally appropriate" (score = 5).

Evaluating the probability of further usage of the agent

After watching the last video and filling in the questionnaires, the participants were presented with a question about the further usage of the agent ("Can you imagine using this virtual agent as a personal assistant?"). The participants were asked to answer this on a scale from "most likely" (score = 1) to "not likely at all" (score = 5). Additionally, they needed to explain their answers in an open-ended question.

Moderator: Questionnaire of Cognitive and Affective Empathy (QCAE)

To measure the empathy level of the participants, the Questionnaire of Cognitive and Affective Empathy (QCAE) as a multidimensional empathy questionnaire was used (Reniers, Corcoran, Drake, Shryane, & Völlm, 2011). QACE consists of five subscales, and each of them is rated on a 4-point Likert Scale from “totally disagree” (score = 1) to “totally agree” (score = 4). The subscales relevant to assessing cognitive empathy include measuring *Perspective Taking* and *Online Simulation*. The subscales related to measuring affective empathy are *Emotion Contagion*, *Peripheral Responsivity*, and *Proximal Responsivity* (Cronbach’s $\alpha = 0.84$).

Moderator: Positive and Negative Affect Schedule (PANAS)

To assess the mood of participants, Positive and Negative Affect Schedule (PANAS) (Watson, Clark, & Tellegen, 1988) was presented to them which consists of 10 mood states for positive affect and 10 for negative affect. In this questionnaire, the participants were asked to what degree they have felt these mood states in the past week on a 5-point scale. The range was between very slightly or not at all = 1, a little = 2, moderately = 3, quite a bit = 4, extremely = 5 (Cronbach’s α negative affect = 0.86, Cronbach’s α positive affect = 0.83).

Moderator: Subjective Technical Confidence (STC)

Subjective Technical Confidence (STC) (Beier, 1999) was utilized to measure the experience and the confidence of the users when interacting with technology. This questionnaire assesses the subjective confidence of people when it comes to their ability to solve technical problems. The short version of this questionnaire was used in this study which includes eight items to be rated on a 6-point scale from “strongly disagree = 1” to “strongly agree = 6”, (Cronbach’s $\alpha = 0.80$).

10.3 Results

Using SPSS version 23, several mixed ANOVAs were conducted for each of these items and attributes separately. Moreover, contrasts using repeated ANOVAs were conducted to test which nonverbal behaviors were rated as more empathic. In all analyses, Post-hoc test using

Bonferroni correction was used. Furthermore, Greenhouse-Geisser correction was reported whenever the assumption of sphericity was violated.

Perception of the items (empathy, unfriendliness, and reactivity)

H1 assumed that priming the participants about the design country of the agent affects users' perceptions of the agent's empathy. The results indicated no significant effect of priming the groups on the agent perception for the item empathy, $F(1, 106) = .63$, $p = .426$, $\eta^2 = .00$ ($M_{\text{Malaysian}} = 2.85$, $SE_{\text{Malaysian}} = .09$, $M_{\text{German}} = 2.96$, $SE_{\text{German}} = .09$), rejecting H1.

According to H1.1, the perception of empathy is based on the observed nonverbal behaviors. The results support this hypothesis by indicating a significant effect of the nonverbal behaviors on perceiving the item empathy across participants, $F(3, 318) = 121.34$, $p < .001$, $\eta^2 = .53$.

Additionally, it was hypothesized that there is an interaction effect between the groups and nonverbal behaviors (H1.2) in rating the item empathy. The results, however, did not provide support for this hypothesis, $F(3, 280.37) = .80$, $p = .479$, $\eta^2 = .00$. This indicates that the perceiving the agent as empathic was not different across the two groups.

RQ1 was concerned with knowing which nonverbal behavior had an influence on the perception of empathy in the two groups. However, as no between-group effect was found between priming the groups and perceiving the agent as empathic, contrasts were performed to compare each emotional nonverbal behavior to the neutral behavior for all participants. The purpose was to gain deeper knowledge about the influence of the nonverbal behaviors in all users. The results are shown in Table 26.

Table 26

Statistical values of the contrasts for the perceived nonverbal behaviors – both primed groups (empathy)

Nonverbal behaviors	<i>F</i>	<i>df</i>	<i>p</i>	η^2	<i>M/SD</i>	
Dropping the Arms plus Sad Face	249.50***	1,107	<.001	.70	<i>M</i> = 3.57	<i>SD</i> = .92
Head Down	225.34***	1,107	<.001	.67	<i>M</i> = 3.44	<i>SD</i> = .94
Sad Face	169.35***	1,107	<.001	.61	<i>M</i> = 3.25	<i>SD</i> = .95
Head Nod + Smile	295.43***	1,107	<.001	.73	<i>M</i> = 3.32	<i>SD</i> = .89
Smile	248.60***	1,107	<.001	.69	<i>M</i> = 3.23	<i>SD</i> = .87
Head Nod	.68	1,107	.409	.01	<i>M</i> = 2.17	<i>SD</i> = .77

Neutral behavior in **item empathy** (*M* = 2.13, *SD* = .70).

p* < .05. *p* < .01. ****p* < .001

As can be seen from Table 26, compared with the neutral behavior, all nonverbal behaviors, except for the Head Nod, led to the perception of empathy in the participants. The descriptive values stated that Dropping the Arms plus Sad Face, Head Down, Head Nod plus Smile, Sad Face, and Smile affected the higher perception of empathy respectively.

RQ2 was concerned with knowing whether, in the presence of the nonverbal responses, the two primed groups perceive the agent as unfriendly and reactive differently. The results indicated no between-group effect on perceiving the agent as unfriendly, $F(1, 106) = .49$, $p = .485$, $\eta^2 = .00$, ($M_{\text{Malaysian}} = 2.00$, $SE_{\text{Malaysian}} = .09$, $M_{\text{German}} = 1.91$, $SE_{\text{German}} = .09$), and reactive, $F(1, 106) = .10$, $p = .305$, $\eta^2 = .01$ ($M_{\text{Malaysian}} = 2.97$, $SE_{\text{Malaysian}} = .09$, $M_{\text{German}} = 3.10$, $SE_{\text{German}} = .09$). This suggests that priming the groups about the country of the agent does not affect their perception of agent's unfriendliness and reactivity in a different way.

However, the results demonstrated a significant effect of the nonverbal behaviors on perceiving the agent as unfriendly, $F(3, 318) = 16.01$, $p < .001$, $\eta^2 = .27$, and reactive: $F(3, 318) = 82.54$, $p < .001$, $\eta^2 = .29$. These findings indicate that the participants' rating of the agent was

based on the nonverbal behaviors they observed and not due to the general impression they received of the agent.

Moreover, the yielded data indicated no significant interaction effect of the groups \times the nonverbal behaviors for rating the items unfriendly and reactive. Accordingly, it can be stated that the perception of the agent with regard to rating these items was not different across the two groups: unfriendly, $F(3, 279.45) = .21, p = .862, \eta^2 = .00$, and reactive, $F(3, 207.56) = .22, p = .797, \eta^2 = .00$.

Perception of the attributes (warmhearted/cold, incompetent/competent, and inactive/active)

RQ3 asked about the effect of priming the groups on users' perceptions of the agent regarding the following attributes: warmhearted/cold, incompetent/competent, and inactive/active. The results revealed that priming the participants about the culture of the agent did not significantly affect their perception of the agent as being warmhearted/cold, $F(1, 106) = .65, p = .420, \eta^2 = .00$ ($M_{\text{Malaysian}} = 3.03, SE_{\text{Malaysian}} = .09, M_{\text{German}} = 3.13, SE_{\text{German}} = .09$), incompetent/competent, $F(1, 106) = .02, p = .894, \eta^2 = .00$ ($M_{\text{Malaysian}} = 3.55, SE_{\text{Malaysian}} = .10, M_{\text{German}} = 3.53, SE_{\text{German}} = .10$), and active/inactive, $F(1, 106) = 1.17, p = .282, \eta^2 = .01$ ($M_{\text{Malaysian}} = 3.24, SE_{\text{Malaysian}} = .08, M_{\text{German}} = 3.36, SE_{\text{German}} = .08$).

However, it is important to note that a significant effect of the nonverbal behavior was found on rating the attribute of warmhearted/cold, $F(3, 318) = 198.06, p < .001, \eta^2 = .65$, incompetent/competent, $F(3, 318) = 3.28, p = .041, \eta^2 = .03$, and inactive/active, $F(3, 318) = 113.5, p < .001, \eta^2 = .52$. These effects suggest that the agent's displayed nonverbal responses after uttering the emotional situations had an influence on the perception of the agent regarding these attributes. However, the results did not show any interaction effect of the groups \times nonverbal behaviors for perceiving the attributes. Therefore, it can be stated that different nonverbal behaviors did not lead to different perceptions in the users: warmhearted/cold, $F(3, 257.92) = 1.07, p = .350, \eta^2 = .01$, incompetent/competent, $F(3, 203.79) = .46, p = .622, \eta^2 = .00$, and inactive/active, $F(3, 205.42) = 1.07, p = .342, \eta^2 = .01$.

Perceived appropriateness of the agent's behavior

According to the assumption of H2, there is a difference between the primed groups' rating of the appropriateness of the agent's behavior. The results, however, did not reveal a significant between-group effect on this rating, $F(1, 106) = 1.18, p = .280, \eta^2 = .01$, rejecting H2.

It is important to note that the data showed a significant within-group effect (nonverbal behaviors) on rating the appropriateness $F(3, 318) = 51.82, p < .001, \eta^2 = .33$. This suggests that the rating of the participants of the appropriateness of the agent's behavior was based on the agent's displayed nonverbal behaviors, supporting H2.1. By contrast, H 2.2, which assumed that there is an interaction effect of the groups \times the nonverbal behaviors on rating the appropriateness of the behaviors, was not supported, $F(3, 224.59) = .30, p = .752, \eta^2 = .00$. This indicates that these ratings were not different across the two groups.

Moderated regression analyses (mood, experience with technology, and level of empathy)

Multiple moderated regression analyses were conducted to test whether participant's mood, experience with technology, and empathic level have a moderating effect on predicting their ratings of the agent's empathy (RQ4) when they are primed about the country of the agent. The results of the regression analyses for these factors are shown in Tables 27-30.

Table 27

Moderated regression analysis for item empathy (positive affect)

Predictors	β	t	p
Groups	-.710	-1.209	.229
Positive Affect	.003	.246	.805
Groups \times Positive Affect	.029	1.538	.127

Model Summary, $F(3,104) = 2.80, p = .043, R^2 = .07$

The results of the regression to test the effect of positive affect on the perception of empathy indicated a significant effect of the model. However, it can be seen that none of the

variables are predictive enough on their own to be significant and affect the perception of empathy.

Table 28

Moderated regression analysis for item empathy (negative affect)

Predictor	β	<i>t</i>	<i>p</i>
Groups	-.386	-1.098	.274
Negative Affect	-.013	-.871	.385
Groups \times Negative Affect	.038	1.695	.092

Model Summary $F(3,104) = 1.62, p = .188, R^2 = .04$

It can be seen from the data in Table 28 that there was no significant effect of negative affect on participants' perception of empathy. Accordingly, it can be suggested that negative affect could neither independently nor in interaction with the group significantly explain the evaluation of empathy.

Table 29

Moderated regression analysis for item empathy (experience with technology)

Predictor	β	<i>t</i>	<i>p</i>
Groups	.641	1.167	.245
Experience	-.065	-.773	.440
Groups \times Experience	-.107	-.879	.381

Model Summary, $F(3,104) = 2.11, p = .103, R^2 = .05$.

As shown in Table 29, moderated regression analysis which was used to test the effect of participants' experience with technology on their perception of empathy demonstrated no

significant effect. This suggests that users' experience with technology could not independently or in interaction with the primed groups explain the evaluation of empathy.

Table 30

Moderated regression analysis for item empathy (level of empathy)

Predictor	β	t	p
Groups	-.125	-.103	.917
Level of Empathy	-.066	-.258	.796
Groups \times Level of Empathy	.106	.246	.806

Model Summary, $F(3,104) = .64, p = .591, R^2 = .02$.

And finally, the participants' level of empathy was not found to have a significant moderating effect on their evaluation of the agent's empathy (Table 30). This demonstrates that none of the predictors could neither independently nor in interaction with each other explain the dependent variable (perception of empathy).

Further usage of the agent

H3 predicted that priming the groups about the design country of the agent influences their rating of further usage of the agent. One-way ANOVA was conducted to test this hypothesis. Results demonstrated no significant effect of the primed groups on further usage of the agent, $F(3,107) = .65, p = .419$ ($M_{\text{Malaysian}} = 3.93, SE_{\text{Malaysian}} .96, M_{\text{German}} = 3.75, SE_{\text{Malaysian}} = 1.24$), rejecting H3. Descriptive values of the responses across all participants showed a medium probability of using the agent as their assistant in the future, with a small tendency to "not likely at all" of imagining to use the agent as their assistant ($M = 3.84, SD = 1.10$).

Afterward, the participants were asked to explain their responses with regard to the probability of the usage of the agent in an open-ended question (RQ6). Comparing their answers revealed that:

- 53 % of the participants indicated that they considered the agent as a little innovative, redundant and not necessary at all as they are able to easily read their calendar entries on their own.
- 33 % of the participants stated that they perceive the agent's emotions as unreal and unclear to interpret.
- 14 % of the participants indicated that they had the impression that the agent's emotions were appropriate, sensitive and interesting, and the agent is practical and fulfilling its purpose.

Qualitative Analyses

In order to evaluate the perception of the users in their own words, at the beginning of each questionnaire, the participants were first asked to give their opinion on their impression of the agent in an open-ended question. Accordingly, RQ5 was concerned with the answers of the participants in this regard (“What are users’ perceptions of the agent in the presence of emotional nonverbal behaviors in their own words?”). The participants were asked: “What do you believe that the agent is trying to display?”. To compare the answers in these sections, three categories were created:

1. Neutral responses, including answers such as calendar, agent, nothing, objective, assistant, a boy, a girl (coded with 0)
2. Happy emotional responses, including answers such as smile, friendliness, joy, excitement, confirming, and competent (coded with 1)
3. Sad emotional responses, including answers such as sorrow, empathy, sympathy, care (coded with 2).

Since all nonverbal behaviors were presented three times, the mean scores for each nonverbal behavior were counted and divided into $K=324$ ($3 * 108$, $N=108$). The results are shown in Tables 31-33.

Table 31

Perception of positive emotional responses regarding the happy nonverbal behaviors

Happy Nonverbal Behaviors	Neutral (0)	Happy (1)	Sad (2)
Head Nod	91	6.71	1.85
Head Nod + Smile	20	79.61	0.30
Smile	22.22	77.77	-

As shown in Table 31, when the agent displayed Head Nod plus Smile and Smile the participants perceived that the agent was responding with positive emotions. The response of the agent when it showed Head Nod was highly perceived as a neutral response.

Table 32

Perception of negative emotional responses regarding the sad nonverbal behaviors

Sad Nonverbal Behaviors	Neutral (0)	Happy (1)	Sad (2)
Head Down	14.81	1.85	83.33
Dropping the Arms + Sad Face	13.58	0.9	85.5
Sad Face	16.36	0.92	82.72

Table 32 indicates that all sad nonverbal behaviors were perceived as sad emotional responses. The nonverbal behavior Dropping the Arms plus Sad Face was perceived as the most emotional response of the agent.

Table 33

Perception of the neutral responses of the agent

Neutral Behaviors	Neutral (0)	Happy (1)	Sad (2)
Neutral Happy	96.6	3.3	-
Neutral Sad	95.37	-	4.23

The neutral behaviors were perceived almost in all cases as a neutral response. This means that observing emotional nonverbal behaviors indeed had an effect on the perception of the agent as empathic and users could differentiate between the emotional and neutral responses (Table 33).

10.4 Additional Analyses

The results in Chapter 10.3 indicated no significant effect of priming the groups on perceiving the agent as unfriendly and reactive. Also, rating the attributes and rating the appropriateness of the agent's behaviors showed no significant effect of the primed groups on the perception of the agent. However, in order to gain knowledge about the evaluations of the nonverbal behaviors across all users, besides for the item empathy, further analyses have been conducted. These analyses contribute to understanding which nonverbal behaviors influenced users perception with regard to different items and attributes.

The effect of nonverbal behaviors on rating the items (unfriendliness and reactivity)

To test which nonverbal behaviors led to the perception of unfriendliness and reactivity contrasts were performed for each item separately. Accordingly, each emotional behavior was compared with the neutral behavior. Table 34 shows statistical and descriptive values of users' rating of these two items.

Table 34

Statistical values of the contrasts for the perceived nonverbal behaviors – both primed groups (unfriendly and reactive)

Nonverbal behaviors	Items	<i>F</i>	<i>df</i>	<i>p</i>	η^2
Dropping the Arms + Sad Face	<i>Unfriendly</i> (<i>M</i> = 1.81, <i>SD</i> = .85)	25.92***	1,107	<.001	.76
	<i>Reactive</i> (<i>M</i> = 3.65, <i>SD</i> = .85)	202.67***	1,107	<.001	.19
Head Down	<i>Unfriendly</i> (<i>M</i> = 1.85, <i>SD</i> = .84)	21.87***	1,107	<.001	.17
	<i>Reactive</i> (<i>M</i> = 3.49, <i>SD</i> = .87)	176.05***	1,107	<.001	.62
Sad Face	<i>Unfriendly</i> (<i>M</i> = 1.96, <i>SD</i> = .82)	13.74***	1,107	<.001	.11
	<i>Reactive</i> (<i>M</i> = 3.41, <i>SD</i> = .77)	186.97***	1,107	<.001	.63
Head Nod + Smile	<i>Unfriendly</i> (<i>M</i> = 1.65, <i>SD</i> = .72)	57.47***	1,107	<.001	.35
	<i>Reactive</i> (<i>M</i> = 3.50, <i>SD</i> = .90)	214.14***	1,107	<.001	.66
Smile	<i>Unfriendly</i> (<i>M</i> = 1.70, <i>SD</i> = .70)	69.38***	1,107	<.001	.39
	<i>Reactive</i> (<i>M</i> = 3.37, <i>SD</i> = .80)	208.50***	1,107	<.001	.66
Head Nod	<i>Unfriendly</i> (<i>M</i> = 2.20, <i>SD</i> = .90)	.43	1,107	.513	.00
	<i>Reactive</i> (<i>M</i> = 2.63, <i>SD</i> = .90)	59.69***	1,107	<.001	.36

Neutral behavior in **item unfriendly** (*M* = 2.23, *SD* = .86).

Neutral behavior in **item reactive** (*M* = 2.12, *SD* = .84).

p* <.05. *p* <.01. ****p* <.001

Table 34 indicates that all nonverbal behaviors except Head Nod, affected the perception of unfriendliness in the users. This means that the users perceived the agent as less unfriendly when it displayed emotional nonverbal responses. Descriptive values indicated that compared with the neutral behaviors, observing the agent displaying Head Nod plus Smile, Smile, Dropping the Arms plus Sad Face, Head Down, and Sad Face made the participant rate the agent as less unfriendly respectively.

Moreover, for the item reactive, results indicated that compared with the neutral behavior all emotional nonverbal behaviors affected the participant's perception of the agent as being reactive. This suggests that the participants perceived that the agent was intending to be

emotionally responsive with regard to the uttered emotional situations. Results showed that Dropping the Arms plus Sad Face, Head Nod plus Smile, Head Down, Sad Face, Smile, and Head Nod led to a higher perception of reactivity of the agent respectively (Table 34).

The effect of nonverbal behaviors on rating the attributes (warmhearted/cold, incompetent/competent, and inactive/active)

Priming the groups did not affect the perception of the agent with respect to rating the attributes of warmhearted/cold, incompetent/competent, and inactive/active as well. However, having insight about the participants' (independent of the primed groups) ratings of certain nonverbal behaviors in these attributes could contribute to a better understanding of the nonverbal behaviors. Additionally, it can provide insight into the design of further nonverbal behaviors for future research. Therefore, contrasts were performed for all participants and for each attribute separately to compare each emotional nonverbal behavior with the neutral behavior. The results are indicated in Table 35.

Table 35

Statistical values of the contrasts for the perceived nonverbal behaviors – both primed groups
(warmhearted/cold, incompetent/competent, inactive/active)

Nonverbal behaviors	Items	<i>F</i>	<i>df</i>	<i>p</i>	η^2
Dropping the Arms + Sad Face	<i>Warmhearted/cold</i> (<i>M</i> = 3.64, <i>SD</i> = .86)	243.36***	1,107	<.001	.69
	<i>Incompetent/competent</i> (<i>M</i> = 3.54, <i>SD</i> = .94)	.47	1,107	.493	.00
Head Down	<i>Inactive/active</i> (<i>M</i> = 3.80, <i>SD</i> = .71)	136.36***	1,107	<.001	.56
	<i>Warmhearted/cold</i> (<i>M</i> = 3.44, <i>SD</i> = .94)	147.52***	1,107	<.001	.58
	<i>Incompetent/competent</i> (<i>M</i> = 3.49, <i>SD</i> = .80)	.04	1,107	.839	.00
Sad Face	<i>Inactive/active</i> (<i>M</i> = 3.57, <i>SD</i> = .78)	102.20***	1,107	<.001	.49
	<i>Warmhearted/cold</i> (<i>M</i> = 3.36, <i>SD</i> = .87)	175.83***	1,107	<.001	.62
	<i>Incompetent/competent</i> (<i>M</i> = 3.42, <i>SD</i> = .90)	.28	1,107	.592	.00
Head Nod + Smile	<i>Inactive/active</i> (<i>M</i> = 3.55, <i>SD</i> = .71)	124.88***	1,107	<.001	.54
	<i>Warmhearted/cold</i> (<i>M</i> = 3.60, <i>SD</i> = .76)	321.01***	1,107	<.001	.75
	<i>Incompetent/competent</i> (<i>M</i> = 3.75, <i>SD</i> = .84)	12.10**	1,107	.001	.10
Smile	<i>Inactive/active</i> (<i>M</i> = 3.76, <i>SD</i> = .70)	170.11***	1,107	<.001	.61
	<i>Warmhearted/cold</i> (<i>M</i> = 3.52, <i>SD</i> = .80)	290.48***	1,107	<.001	.73
	<i>Incompetent/competent</i> (<i>M</i> = 3.71, <i>SD</i> = .76)	12.50**	1,107	.001	.10
Head Nod	<i>Inactive/active</i> (<i>M</i> = 3.60, <i>SD</i> = .70)	137.93***	1,107	<.001	.56
	<i>Warmhearted/cold</i> (<i>M</i> = 2.51, <i>SD</i> = .76)	21.75	1,107	.409	.17
	<i>Incompetent/competent</i> (<i>M</i> = 3.50, <i>SD</i> = .90)	.26	1,107	.611	.00
	<i>Inactive/active</i> (<i>M</i> = 3.00, <i>SD</i> = .78)	41.20***	1,107	<.001	.28

Neutral behavior in **warmhearted/cold** (*M* = 2.28, *SD* = .77).

Neutral behavior in **incompetent/competent** (*M* = 3.47, *SD* = .88).

Neutral behavior in **inactive/active** (*M* = 2.58, *SD* = .89).

p* <.05. *p* <.01. ****p* <.001

It is apparent from Table 35 that compared with neutral behavior, all nonverbal behaviors except Head Nod affected the perception of the agent as being warm or cold. Comparing the Descriptive values demonstrated that when the agent showed Dropping the Arms plus Sad Face,

Head Nod plus smile, Smile, Head Down, and Sad Face respectively, the participants had a tendency to rate the agent as more warmhearted.

For the attribute of incompetent/competent, the results showed that when the agent displayed Head Nod plus Smile, and Smile, compared with the neutral behavior, the participants tended to rate the agent as more competent. Comparing the descriptive values showed Head Nod plus Smile and Smile led to a higher perception of competency respectively.

And finally, the results demonstrated that compared with the neutral behavior, all emotional nonverbal behaviors affected the perception of the agent as being inactive/active. Descriptive values indicated that Dropping the Arms plus Sad Face, Head Nod plus Smile, Smile, Head Down, Sad Face, and Head Nod led to a higher perception of the agent as being active. Therefore, based on the findings of this study, the participants had an overall tendency to rate the agent based on its displayed nonverbal behaviors as positive.

The effect of nonverbal behaviors on rating the appropriateness of the behaviors

In order to have more precise knowledge regarding the nonverbal behaviors, it was strived to test which nonverbal behaviors were rated as more appropriate by all participants. Therefore, contrasts were performed to compare each nonverbal behavior with the neutral behavior for all participants. Results are indicated in Table 36.

Table 36

Statistical values of the contrasts for the perceived nonverbal behaviors – both primed groups (appropriateness)

Nonverbal Behaviors	<i>F</i>	<i>df</i>	<i>p</i>	η^2	<i>M</i> and <i>SD</i>
Dropping the Arms + Sad Face	12.06*	(1,107)	<.001	.10	<i>M</i> = 3.54 <i>SD</i> = .91
Head Down	25.67***	(1,107)	<.001	.19	<i>M</i> = 3.60 <i>SD</i> = .81
Sad Face	5.80*	(1,107)	.018	.05	<i>M</i> = 3.38 <i>SD</i> = .85
Head Nod + Smile	103.15***	(1,107)	<.001	.49	<i>M</i> = 3.97 <i>SD</i> = .69
Smile	165.33***	(1,107)	<.001	.60	<i>M</i> = 4.02 <i>SD</i> = .61
Head Nod	1.45	(1,107)	.231	.01	<i>M</i> = 3.21 <i>SD</i> = .81

Neutral behavior in **inappropriate/appropriate** (*M* = 3.14, *SD* = .77).

p* <.05. *p* <.01. ****p* <.001

As can be seen from the table above all emotional nonverbal behaviors except for the Head Nod affected participant's perception of the agent's behavior as inappropriate/appropriate. The participants rated Smile, Head Nod plus Smile, Head Down, Dropping the Arms plus Sad Face, and Sad Face respectively as most appropriate nonverbal responses to the emotional contexts.

10.5 Follow-up study (with Pictures)

A follow-up online study was further conducted to test whether the results of the Study 3 was merely based on users' knowledge of the design country of the agent. The findings of this study could also assure that the users were not under the impression of other factors such as the look of the agent. Therefore, in order to control the potential effect of the agent's appearance, 2

conditions out of 4 included no picture of Billie, in which the participants needed to imagine working with an agent from either their own culture or a different one. Fifty-two native German participants (38 females and 14 males), with the age range from 18 to 60 years old ($M = 27$, $SD = 8.5$) took part in this study. The participants did not have a prior encounter with the virtual agent Billie. The total duration of the study was approximately 5-6 minutes.

In this study, no videos with emotional situations and nonverbal responses were presented to the participants. Instead, four different conditions were created: (1) the agent was designed in Germany and with a picture of the agent, (2) the agent was designed in Germany without a picture of the agent, (3) the agent was designed in Malaysia and with a picture of the agent, and (4) the agent was designed in Malaysia without a picture of the agent.

In the cover stories, similar to the cover stories in the main study, participants were informed that the agent is designed in Germany or Malaysia and is well accepted by German or Malaysian users. However, in this study, they were asked to *imagine* working with this virtual agent and give their impression of it. The participants were randomly assigned to each of these four conditions (German_Pic, $N = 17$, German_NoPic $N = 9$, Malaysian_Pic $N = 18$, Malaysian_NoPic $N = 8$). Afterward, the same questionnaires as in the main study were presented to the participants. First, the participants needed to fill in a box to express their impression of Billie (“what would be your attitude toward Billie?”). Then they were asked to rate the agent regarding the three items, “I feel Billie is:” unfriendly, empathic, and reactive. The participants had to rate Billie on a 5-point Likert scale from “I totally agree” to “I totally disagree” (highest score = 5, lowest score = 1 for empathy and reactive) and from “I totally agree” to “I totally disagree” (highest score = 1, lowest score = 5 for item unfriendly).

The second questionnaire was related to rating the agent on different attributes on a 5-point Likert Scale from “I totally agree” (highest score) to “I totally disagree” (lowest score). The participants were asked to rate “that Billie would be:” cold (score = 1)/warmhearted (score = 5), incompetent (score = 1)/competent (score = 5), and inactive (score = 1)/active (score = 5).

Furthermore, the questionnaire related to appropriateness of Billie’s behavior was presented to the participants. They needed to rate the appropriateness on a scale from “totally appropriate” (lowest score = 1) to “totally appropriate” (highest score = 5). Finally, the participants were asked to rate if they could imagine using this virtual agent as their personal

assistant on a scale (1 = *most likely* to 5 = *not likely at all*). They further needed to give an explanation to their answers by filling in a box.

Additionally, to test whether users' mood, experience with technology, and level of empathy have a moderating effect on their evaluations, the participants were asked to fill in the questionnaires regarding measuring these factors (PANAS, STC, and, QCAE).

Evaluation of the items (unfriendly, empathic, and reactive)

In order to test the between-group effect of the users' evaluation of the virtual agent Billie, a MANOVA was conducted. The four conditions were chosen as independent variables and users' perception of empathy, unfriendly, and reactive were chosen as dependent variables. Results indicated no significant effect of priming the groups (with or without pictures) on the evaluation and imagination of empathy, $F(3, 48) = .974, p = .41, \eta^2 = .05$, unfriendly, $F(3, 48) = .63, p = .599, \eta^2 = .04$, and reactive, $F(3, 48) = 2.22, p = .09, \eta^2 = .12$. Figure 9 compares the descriptive values of rating these three items.

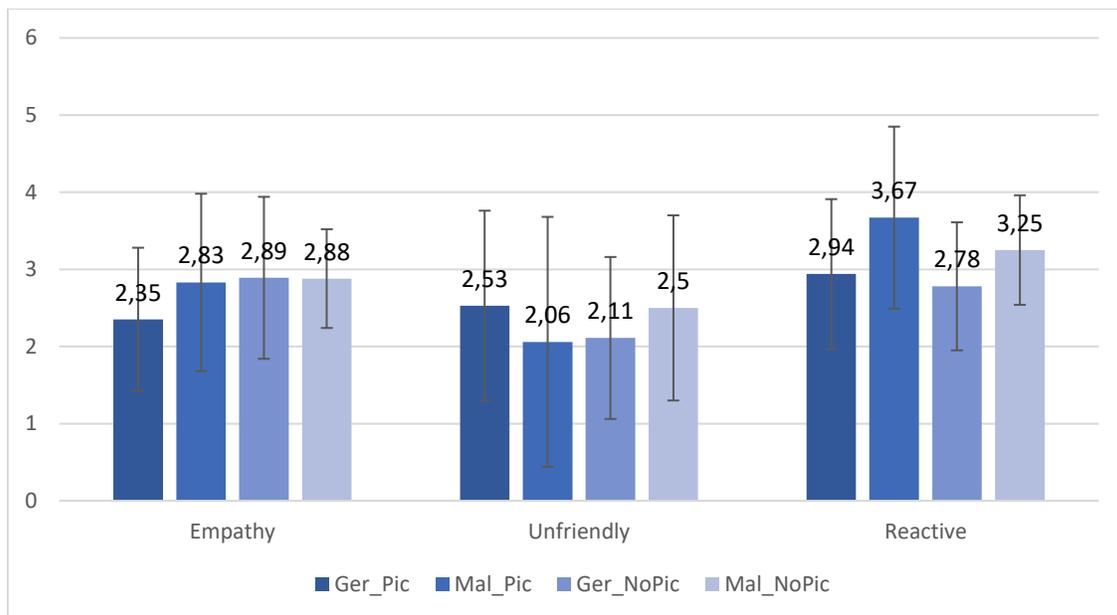


Figure 9: Descriptive values of the four primed groups' ratings (empathy, unfriendly, and reactive)

Evaluation of the attributes (warmhearted/cold, incompetent/competent, and inactive/active)

To examine the effect of primed groups on evaluating the agent of the attributes of warmhearted/cold, incompetent/competent, and inactive/active, a MANOVA was conducted with groups as independent variable and the attributes as dependent variables. The results demonstrated that priming the groups had no significant effect on rating these attributes: warmhearted/cold, $F(3, 48) = .73, p = .539, \eta^2 = .04$, incompetent/competent, $F(3, 48) = .57, p = .634, \eta^2 = .03$, inactive/active, $F(3, 48) = .95, p = .423, \eta^2 = .05$.

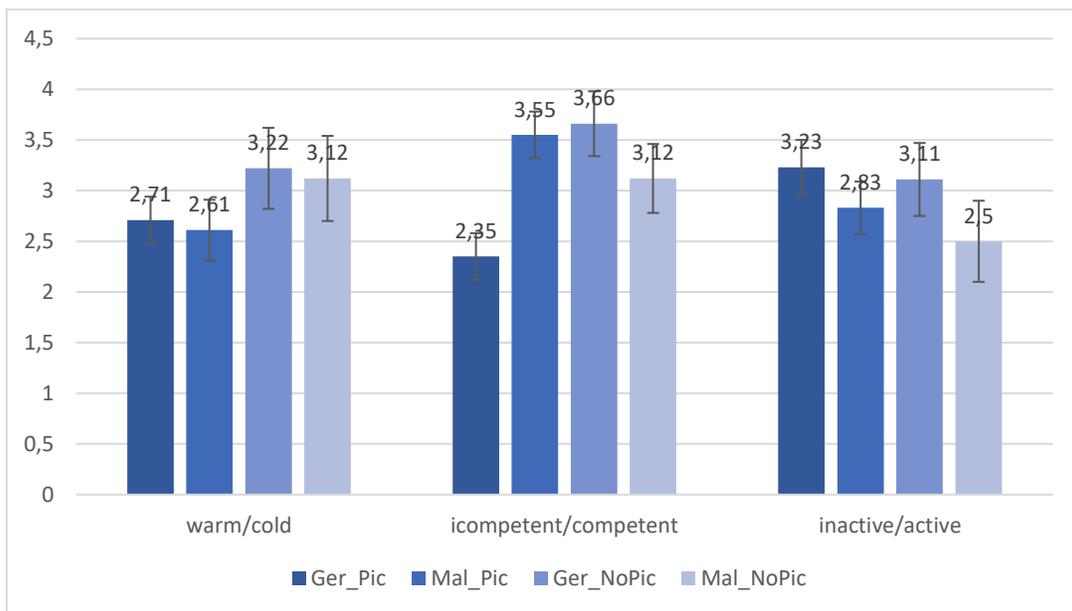


Figure 10: Descriptive values of the four primed groups' ratings (warmhearted/cold, incompetent/competent, and inactive/active)

As the descriptive values indicated, the participants who saw a picture of Billie tended to rate it as cold in contrast to the people who were not presented with a picture. However, in general, it can be seen that most of the participants had a tendency to rate the agent as competent and slightly as active.

Evaluation of appropriateness of the agent's behavior

In order to test whether there is a group effect on rating the appropriateness of the agent's behavior, one-way ANOVA was conducted with the groups as the independent variable and participants' rating of the attribute of appropriateness as the dependent variable. Results showed no significant effect of the primed groups on rating the potential appropriateness of the agent's behavior, $F(3, 48) = 19, p = .903$, ($M_{Ger_Pic} = 3.00, SD_{Ger_Pic} = .80, M_{Ger_NoPic} = 3.22, SD_{Ger_NoPic} = .67, M_{Mal_Pic} = 3.22, SD_{Mal_Pic} = 1.26, M_{Mal_NoPic} = 3.13, SD_{Mal_NoPic} = .64$).

As shown in the descriptive values, all participants tended to imagine that the agent would behave appropriately. ($M = 3.13, SD = .92$).

The probability of future usage of the agent

One-way ANOVA was conducted to test the effect of the primed groups on participants' rating of possible usage of the agent in the future. Results demonstrated no significant between-group effect on the rating of further usage, $F(3, 48) = .43, p = .729$, ($M_{Ger_Pic} = 3.76, SD_{Ger_Pic} = 1.25, M_{Ger_NoPic} = 3.90, SD_{Ger_NoPic} = .78, M_{Mal_Pic} = 3.40, SD_{Mal_Pic} = 1.37, M_{Mal_NoPic} = 3.63, SD_{Mal_NoPic} = 1.19$). According to the descriptive values, it can be seen that the participants showed a medium probability, with a very small tendency to "not likely at all" of imagining to use the agent as an assistant ($M = 3.63, SD = 1.20$).

Moderated regression analyses (mood, experience with technology, and level of empathy)

Multiple moderated regression analyses were conducted to test the moderating effect of participants' mood, experience with technology, and the level of empathy on their perception of a possible empathic agent. Results are indicated in the tables below:

Table 37

Moderated regression analysis for item empathy in the follow-up study (positive affect)

Predictors	β	t	p
Groups	-.468	-.947	.348
Positive Affect	-.142	-.568	.572
Groups \times Positive Affect	.228	1.384	.172

$\alpha = 0.87$

Model Summary, $F(3, 48) = 1.42, p = .246, R^2 = .08$

The results of the regression, as indicated in Table 37, showed no significant effect of positive affect on the users' evaluation of the possibility that the agent would be empathic. This means that the positive mood of the participants could neither independently nor in an interaction with the primed groups influence their evaluation of the agent.

Table 38

Moderated regression analysis for item empathy in the follow-up study (negative affect)

Predictors	β	t	p
Groups	.021	.062	.950
Negative Affect	-.024	-.067	.946
Group \times Negative affect	.078	.431	.668

$\alpha = 0.91$

Model Summary, $F(3, 48) = .84, p = .478, R^2 = .05$

Similarly, the results of moderated regression indicated no significant effect of the negative affect of the users on their evaluation that the agent can be empathic, suggesting that none of the predictors could explain the participants' evaluation (Table 38).

Table 39

Moderated regression analysis for item empathy in the follow-up study (experience with technology)

Predictors	β	t	p
Groups	.631	.986	.328
Experience	.134	.546	.587
Groups \times Experience	-.106	-.715	.477

$\alpha = 0.81$

Model Summary, $F(3, 48) = .848, p = .474, R^2 = .05$

As demonstrated in Table 39, users' experience with technology did not have a moderating effect on evaluating the agent as empathic. This suggests that none of the predictors could predict the participant's evaluation of the item empathy.

Table 40

Moderated regression analysis for item empathy in the follow-up study (level of empathy)

Predictors	β	t	p
Groups	-.876	-.770	.444
Level of Empathy	-.314	-.476	.636
Groups \times Level of Empathy	.386	.939	.352

$\alpha = 0.85$

Model Summary, $F(3, 48) = .995, p = .403, R^2 = .06$

And finally, the yielded data indicated no significant effect of users' level of empathy on their rating that the agent could possibly be an empathic agent. This suggests that none of the predictors could explain their evaluation, neither independently nor in interaction with each other (Table 40).

Qualitative Analysis

In order to compare the answers to the open-ended questions, categories were formed of the given answers. In these questions the users stated their impression of Billie after reading the cover stories (and seeing a picture of it in 2 out of 4 conditions). Based on the answers, three categories were made:

1. Neutral coded with 0: answers such as, neutral, no idea, I do not have information, I need more information.
2. Positive coded with 1: answers such as nice boy, humane, pleasant, life-like, helpful, intelligent, friendly, good, interesting, open, and positive.
3. Negative coded with 2: answers such as tense, boring, artificial, bad, skeptical, not sympathetic, too young, disproportionate, inexperienced, casual, and strange.

The results showed that:

- 16 % of the participants rated the agent as neutral.
- 39 % of the participants rated the agent as positive.
- 35 % of the participants rated the agent as negative.

Furthermore, the qualitative analysis of the users' explanations regarding their the probability of future usage of the agent demonstrated that:

- 23 % of the positive responses were: because the agent seems friendly, helpful, reliable, positive and supportive, fun looking, practical, and easy.
- 48 % of negative responses were: the agent is not sympathetic and approachable, not professional, is childish, annoying, spying, I do not like the design and layout, I do not need such an agent, and the agent's clothing is not appropriate.
- 29 % of the neutral responses were: the agent is neutral, I need more information of the agent, my acceptance depends on the context, I would rather read more reviews of the agent, and I need more experience with the agent.

The findings of the follow-up study further support the findings of the main study (Study 3). Accordingly, it can be stated that users do not evaluate the agent differently, when they are informed about the country of the agent.

10.6 Discussion

Media equation theory indicated that users apply social rules from human–human interactions to human–computer interactions (Reeves & Nass, 1996). In this respect, many studies provided evidence that users’ responses to a computer are based on social rules and norms they use when they interact with a human (Isbister & Nass, 2000; Nass & Moon, 2000). One of these social rules that has been investigated in the domain of human–computer interaction is stereotypical attributions of users by categorizing computer agents as ingroup or outgroup members. When others are categorized as ingroup members, they are perceived to have the same values (Allen & Wilder, 1979), and to be more likable and trustworthy (Clark & Maass, 1988; Lee, 1993). Many studies have investigated users’ stereotypical attitude toward agents’ culture (see e.g., Baylor, 2005; Nass et al., 2000). However, as mentioned earlier in Chapter 4, there have been mixed findings in this regard. Therefore, it is still important to test whether cultural stereotypes play a role in how users perceive agents and act in an interaction with them. Furthermore, most of the studies to date have investigated the effect of users’ stereotype based on apparent visual cues in the agent representing its ethnicity. Therefore, in this study, the purpose was to examine how cultural stereotypes, which can be triggered by knowing that the agent is from a different country (without obvious cultural features), influence users’ perception of an emotional agent and their rating it as empathic.

Contrary to H1, which assumed that priming the users about the design of the agent influences their perceptions of an emotional agent as empathic, results revealed no difference in the ratings of the agent between the two groups. Although users’ ratings were based on observed nonverbal behaviors (H1.1), they did not differ in perceiving the agent as more or less empathic. This suggests that having knowledge only of the country that the agent was designed in, without any obvious features in the appearance of the agent, is not sufficient to create stereotypes in users to rate empathy. This may be explained as: to have a different opinion of the emotional agent from a different country, obvious cultural features might also have to be present in the agent,

such as in its appearance or in the accent. This explanation could further be supported by the nonsignificant interaction effect of the groups and the nonverbal behaviors (H1.2). This suggests that the two groups did not differ in the way they rated specific nonverbal behaviors as more or less empathic.

As the analysis showed, apart from the Head Nod, all the other emotional nonverbal behaviors were rated as empathic by all participants. One possible explanation could be that the intensity of the emotion expression mattered. In this regard, the more body movements and facial expressions could lead the participants to perceive that the agent was empathizing with them.

Additionally, there were two other research questions in this study: whether there are differences between the two primed groups in how they perceive the agent as unfriendly and reactive, (RQ2) and whether they have different ratings regarding attributes such as the agent is warmhearted/cold, incompetent/competent, and inactive/active (RQ3).

The results revealed that users' knowledge regarding the country of the agent does not affect their overall rating of the agent with regard to these items and attributes. However, the data revealed a significant effect of nonverbal behaviors for all of these ratings, meaning that it was indeed the observation of the nonverbal behaviors that affected users' evaluation of the agent. It is important to note that the data indicated no significant interaction effect of the groups and the nonverbal behaviors for neither rating the items nor the attributes. Therefore, it can be suggested that users' perception of the agent was similar across the two groups. It seems plausible that these results are, as for the item empathy, due to the absence of the cultural features in the agent. Another possible explanation could be that the participants who were primed that the agent was designed in Malaysia overlooked the cultural distance and stereotypes of the agent because they observed that the agent was striving to empathize with them. However, since the manipulation check failed in this study and the users were not asked at the end of the study about the country of the agent, this assumption needs to be considered with caution.

There are, however, other possible explanations as well. It can be stated that there is a probability that the neutral look of the agent (neither German nor Malaysian look) reduced the ingroup/outgroup bias toward the agent. Moreover, in the study of Vang and Fox (2014), which was concerned with White users' evaluation of Black virtual agents, sharing a task, regardless of the purpose, contributed to a more positive evaluation of the Black agents. Although the

mechanism of this study is different from the study by Vang and Fox (2014), and the agents in that study had obvious ethnicity differences in their appearances, it is possible that the ingroup/outgroup bias is reduced due to the agent's empathizing with the users. This is in line with the findings of previous studies suggesting that when the agent shows empathic responses, the quality of human-agent interactions is maximized (Brave et al., 2005; Mascarenhas et al., 2013; Paiva et al., 2017; Prendinger & Ishizuka, 2005b). Therefore, it can be stated that the users, under the impression of the agent's emotional responses, overlooked that the agent was from a different country. However, as mentioned earlier, it should be kept in mind that it is plausible that the absence of stereotypes might be a result of not paying attention to the cover story. Therefore, this needs to be checked in further studies.

In order to have a deeper insight into the nonverbal behaviors created, analyses were conducted to test which nonverbal behaviors were rated positively or negatively. Based on the data, all emotional nonverbal behaviors except Head Nod engendered a feeling in the users that the agent is empathic and less unfriendly (therefore more positive). For the item reactive, however, all nonverbal behaviors made the users rate the agent as reactive. For rating the items empathy and unfriendly, Head Nod did not offer an emotional intensity for the users to differentiate it significantly from the neutral behavior. However, its effect on the rating of reactivity might be due to the movement of the head, which influenced the participants' evaluation of the agent as being responsive.

Additionally, all sad and happy emotional nonverbal behaviors were perceived as active, which can be explained by the body movement and facial expression of the agent, making the users perceive it as responsive. The sad emotional nonverbal behaviors were not effective in influencing the users to rate the agent's competence, but Head Nod plus Smile and Smile as happy responses led to a higher perception of competency. One possible explanation might be that to evaluate the competency of the agent, other features of the agent might be needed. For example, the users might need to know how successful the agent is in assisting them in a task or showing verbal responses in order to make the users aware of its assistive intentions.

And finally, all nonverbal behaviors except Head Nod influenced users' perception of the agent as being warmhearted. The fact that Head Nod did not influence the emotional perception of the agent, but merely its reactivity and activeness, can be considered when designing agents

in which the purpose is only to show a reaction (confirmation) to the situation. Considering this nonverbal behavior as an empathic response does not seem to be practical, at least for the target group of this study.

Additionally, in this study, it was of interest to test whether the two groups evaluate the agent's behavior as appropriate in a different way (H2). The results again indicated no effect of priming the groups on rating the appropriateness, thus rejecting H2. The significant within-group effect (nonverbal behaviors), however, supported H2.1, indicating that the users rated the appropriateness of the agents' behavior based on its nonverbal behaviors. However, the insignificant interaction of the groups and the nonverbal behaviors did not provide support for H2.2, which assumed that users of the different groups perceived the appropriateness of the behaviors differently.

Similar to users' other ratings in this study, it is plausible that the agent's difference was ignored because of its emotional behavior and lack of apparent visual features indicating that the agent was from a different country. However, because it was important to test whether the nonverbal behaviors could properly display that the agent is empathizing with the users or responding emotionally with a positive intention, the effect of each emotional nonverbal behavior was analyzed for all participants. The data indicated that in comparison with the neutral behavior all nonverbal behaviors except Head Nod were rated as appropriate responses. This is consistent with the findings from rating the item empathy, in which participants rated Head Nod as the least empathic. Thus, the intensity of the nonverbal behaviors indeed influenced users' evaluation of the agent, which should be taken into account when aiming to create certain perceptions in the users.

To ensure that users' evaluations are based on the knowledge of the country of the empathic agent only, this study also set out to assess whether other factors such as the users' level of empathy, mood, and experience with technology influence their rating of empathy.

Data revealed that neither the level of empathy nor the users' experience with technology had a moderating effect on perceiving the agent as empathic. But it was shown that when the users have a positive mood, they rate the agent as being more empathic. However, as the positive affect did not influence the users' perception of empathy independently (as there was just the significant effect of the model), the results need to be interpreted with caution.

Additionally, qualitative analysis was used in this study to gain insights into participants' perception of an emotional agent in their own words, without biasing them with the presented items and attributes in the questionnaires. In this regard, participants were asked to give their opinion regarding the understanding of the agent, after watching each video (and before filling in the questionnaires). As the results indicated, most of the happy and sad nonverbal behaviors were accurately perceived as happy and sad responses by the users. It is important to note that there was a common view among the participants, in that they used the words *empathy*, *sympathy*, and *caring* mostly for when the agent showed sad emotional nonverbal behaviors. This can be explained by the negative emotional valence of the sad situations, in which there is a high probability that people link the emotional responses as caring responses. Furthermore, as in the quantitative results, while a minority mentioned that Head Nod is a happy or sad response, most of the participants agreed that it is a neutral behavior. This finding provides knowledge for designing nonverbal behaviors for different purposes. It suggests that if the objective of the agent is to act empathically toward the user, it is plausible that head nodding will not be perceived as an emotional response by the users.

Furthermore, the probability of future usage of the agent was tested by asking the participants directly whether they could imagine themselves using this agent as their assistant. H3 of the study stated that priming the two groups about the design country of the agent affects users' responses to this question. The results, however, did not provide support for this hypothesis, as there was no significant effect of priming the groups on willingness to use the agent in future. Furthermore, on average, the participants answered that they were inclined not to use the agent in future. Looking at their answers indicates that most of them did not feel the need to use such an agent as they were capable of reading their calendar and that the agent looked unreal. One possible explanation for these findings might be that the majority of the participants in this study were young users (mostly students), who did not feel connected to the aim of designing a virtual assistant for daily usage. Perhaps with the group of older people, especially older people in need of assistance, qualitative data will yield different results. Therefore, this needs to be investigated in future studies.

Additionally, in order to ensure that users' rating of the agent was not influenced by the agent's appearance and look (as mentioned in some of their comments), a short follow-up online study was conducted. In this study the same cover stories were presented to German participants.

In some conditions, the participants saw a picture of the agent only (with no emotional situations uttered and no nonverbal behaviors displayed), and in other conditions they had to imagine interacting with a virtual assistant from either Germany or Malaysia.

The data indicated that as in the main study, there was no significant effect of priming the groups on rating the agent as empathic, unfriendly, reactive, warmhearted/cold, incompetent/competent, and inactive/active. Further, there was no significant difference between the groups with regard to imagining that the agent's behavior could be appropriate or inappropriate. Moreover, no moderating effect of the participants' level of empathy, mood, and experience with technology was observed on their rating of the agent. These findings provide support for the results of the main study, that informing the users about the country in which the agent was designed did not engender stereotypical attitudes toward that agent. This suggests that it is plausible to avoid stereotypical attitudes toward the agent by designing a neutral-looking agent, no matter where the agents are designed and whether the users are aware of the agent's production country. This assumption can be further supported by the findings of the qualitative data, in which the majority of the participants tended to rate the virtual agent as positive.

Comparing the probability of the future usage of the agent, again, did not indicate a significant difference between users' answers in different groups. However, their comments provide insights into the look of the specific agent, such as the agent is childish, it does not look professional, and that its clothing style is inappropriate. This provides information that should be taken into account for future design, that is, there is a possibility that a child-like agent does not give the impression that it is a professional assistant. However, this assumption needs to be tested with different age groups and with users from different cultures.

In sum, Study 3 (and the follow-up study) make a contribution to research with the focus on the effect of cultural stereotypes on users' perception of a virtual agent. The findings demonstrate that having mere knowledge about the design and production country of the agent does not bias people to evaluate neutral-looking agents as empathic (and positive in general). Accordingly, it is important to investigate whether users' perception of the agent differs in perceiving its empathic responses if they can discern in the appearance of the agent that it has a dissimilar ethnicity. If the findings indicate that there is no effect of users' knowledge (of the culture of the agent) on evaluating the agent, but there is an effect of its physical appearance,

designers can take this into account to produce agents with a more neutral appearance. This way, there is a higher probability to avoid any stereotypical attitudes of the users toward the agent, which can ultimately increase their satisfaction with the agent.

10.7 Limitations and Future Research

The generalizability of these results is subject to certain limitations. For instance, all of the participants in this study were young German participants. In order to test the stereotypical attitudes of the users toward the agents from different countries, participants from different cultures (Malaysian, e.g., in the case of this study) should be taken into account. Furthermore, an interactive scenario, in which the probability that the agent responds incorrectly might lead to users' frustration, and thus increase their stereotypical attitudes. Therefore, testing an emotional agent in such an interactive scenario is recommended. The manipulation check needs to be met at the end of the study, to make sure that the participants have paid attention to information presented to them at the beginning of the study in the cover story. And finally, as mentioned earlier, it is important to examine whether users' perception differs when besides having knowledge of the agent's design country, they see representations of a specific culture of the agent in its look.

10.8 Conclusion

The main goal of this study was to determine whether in an encounter with a virtual assistant, priming people about the country that the agent has been designed and produced in makes them have a stereotypical attitude toward the agent. In this regard, it was assumed that users' perception would differ in perceiving the agent as empathic based on its emotional nonverbal behaviors and, also, users' overall rating as either more positive or negative. Results demonstrated that merely knowing the agent's country of origin, as long as the agent does not have any physical representation of that country, does not affect users' evaluations of the agent.

V General Discussion

In the following chapters, first, the main findings of the empirical studies are summarized, interpreted, and discussed in relation to the theoretical background. Afterward, the contribution of this dissertation with regard to theoretical and practical implications is presented. Subsequently, limitations and areas for further research are identified. Finally, the last chapter presents the conclusion and highlights the main findings of the dissertation.

11 Synopsis of Empirical Results

A large and growing body of literature has investigated the influence of empathic agents on human-agent interactions. Although it provides valuable insights into implementing empathy in virtual agents, it has several shortcomings, some of which are tackled in this dissertation. For example, several studies have focused on users' perceptions and opinions of the agent (e.g., Brave et al., 2005; Nass et al., 2001). However, no studies have investigated the role of user variables, such as age and culture, on the perception of emotional nonverbal responses of an empathic agent.

Accordingly, the main purpose of this dissertation was to investigate whether user variables can influence the perception of an empathic virtual agent based on its displayed emotional nonverbal behaviors. Furthermore, the aim was to scrutinize which emotional nonverbal behaviors elicit a higher perception of empathy in the users, with respect to the user factors that can affect their perception. Having insights into how empathic agents are perceived by different users and investigating which specific emotional nonverbal behaviors can contribute to a higher perception of empathy is important in enhancing the quality of users' experience. To this end, three empirical studies were conducted that examined the effect of users' age (Study 1), cultural background (Study 2), and potential bias toward an outgroup agent (Study 3) on the perception of emotional nonverbal behaviors of an agent. In the following chapters, the results of the studies are discussed in detail and with respect to the theoretical background: theory of mind (regarding its effect on understanding intentions and emotions of the virtual agent) and media equation theory (with regard to applying social rules of human-human interactions to human-agent interactions).

User Variables

To date, many researchers have investigated the role of individual variables in the perception of virtual entities (see e.g., Downing, Moore, & Brown, 2005; Gefen & Straub, 1997; Krämer et al., 2010; Zajicek & Hall, 2000). However, in research on embodied conversational agents, far too little attention has been paid to the effect of individual differences on the perception of emotional (empathic) nonverbal behaviors of an agent.

Accordingly, how different user variables and attitudes (age, culture, and potential ingroup/outgroup stereotypes toward the agent) can impact the user's perception of an empathic agent was aimed to be assessed in this dissertation. To this end, different emotional nonverbal behaviors, such as Smile, Head Nod plus Smile, Head Nod, Sad Face, Head Down, and Dropping the Arms plus Sad Face were created. It was investigated whether different user groups exhibit different perceptions in rating these behaviors as empathic.

What was the effect of users' age on perceiving the virtual agent?

The first user variable that was investigated in this dissertation was age (Study 1). The effect of the aging process in the function of ToM in late adulthood and its consequences related to people's social skills (Bailey et al., 2008) makes it important to investigate the perception of older users in an interaction with an empathic agent. This is specifically crucial in seniors' interactions with a virtual assistant (which is supposed to assist them with their daily activities, meaning that it will be engaged in their everyday lives). Moreover, when the agent strives to show empathic responses via nonverbal behaviors, it raises the question of whether older users can notice these behaviors, and precisely interpret the emotions. This is specifically important considering that the general cognitive decline of older people might influence the function of ToM, and hence their ability to accurately understand the emotional responses of the agent.

The results showed that the users' age indeed influenced their perception of an empathic virtual agent. Comparing the evaluation of older and younger users of a virtual assistant in Study 1, it was found that seniors perceived the agent as more empathic than their younger peers did. This might suggest that seniors' ability of ToM was not under the impression of aging, and therefore, they did not show difficulties in recognizing the intention of the agent. This is in line with the claim of the sociocognitive approach (see Phillips et al., 2002). According to this

approach, older people are more capable of understanding the emotions of others because they have experienced different expressions of emotions across their life span. Therefore, it may in fact be the case that the extensive experiences of social interactions that seniors had made them expect to see empathic responses in emotional situations and, therefore, they paid more attention to nonverbal behaviors. This, ultimately, could have contributed to understanding the intention of the agent more precisely and the emotional responses it was displaying.

Furthermore, comparing the participants' evaluations of the most empathically perceived nonverbal behaviors in the first study indicated that older people perceived Dropping the Arms plus Sad Face, Head Down, Sad Face, Head Nod plus Smile, and Smile as more empathic. The most empathic nonverbal behaviors according to younger adults were Dropping the Arms plus Sad Face, and Head Down. It is important to note that seniors were able to accurately perceive the emotional facial expression of the agent (Sad Face and Smile) as empathic. This contradicts the findings of prior studies that claimed the lower capability of seniors in recognizing emotional facial expressions (Beer et al., 2015). Moreover, the fact that older people perceived almost all nonverbal behaviors as empathic, except for Head Nod (which had less intensity in showing emotions) provides evidence that they were indeed capable of differentiating between different emotional nonverbal behaviors as well. This provides further support for the sociocognitive approach.

It is important to note that the factor of age did not significantly affect users' perceptions of the agent regarding all of the items presented to the participants. Besides the item empathy, older people showed a higher rating of the item trustworthy only. The nonsignificant effect of age on rating the items friendly, intelligent, and helpful can be interpreted as an additional evidence that older people did not simply evaluate the nonverbal responses as something positive, but they could associate them with the agent's purpose of empathizing. Therefore, they exhibited a higher rating of the item empathy instead. The higher rating of the trustworthiness of the agent could be interpreted as the side effect of seniors' higher perception of empathy since a correlation between these two items was also observed. Prior studies have shown that empathic agents lead to a higher perception of trustworthiness in users (Brave et al., 2005). However, in Study 1, the perception of empathy engendered the perception of trustworthiness only for the user group of older people, and younger users preferred to receive a neutral response in this respect. It is plausible that seniors, who are assumed to have less experience of technology, could feel more insecure when

encountering the agent. Therefore, it is possible that when they observed the agent showing emotional (empathic) responses, they felt less intimidated and thus could trust the agent more. However, it is important to note that users' prior experience with technology and their confidence with regard to confronting new technologies were not controlled for in this study. Therefore, this interpretation needs to be regarded with caution.

What was the effect of users' culture on perceiving the virtual agent?

The second individual variable that was investigated was the culture of the users (Study 2). Users from Iran and Germany as two culturally different societies were recruited to participate in the study. The results demonstrated that individual differences regarding the cultural background influenced users' perceptions of an empathic agent. In this respect, Iranian users showed a higher perception of empathy in the presence of the agent's emotional nonverbal behaviors compared with German users. It seems plausible that this result is due to the tendency toward collectivistic values and norms in Iranian participants and the effect these values have on understanding the emotions of others. Interpersonal harmony is a value that is learned from early childhood among Iranians, and in the process of upbringing, people learn to be sensitive to the emotions of others (Degens et al., 2012). Due to such sensitivity, it is possible that Iranians have a higher expectation of receiving emotional responses in emotional contexts. A lack of such responses can be interpreted as impoliteness and does not satisfy their expectations of harmonious communication. Furthermore, since the skill of ToM goes hand in hand with understanding other people's empathy, it can be assumed the collectivistic culture of Iranian users potentially influenced their ability to understand the empathic agent.

By contrast, it is plausible that German users with their relatively individualistic values rated the agent as less empathic since they might not have felt the need to observe emotional responses. Independence of individualists can make them less other-oriented and, therefore, understanding the emotions of others might not be as important as it is for collectivists. As was mentioned earlier in Chapter 3.3, environmental factors such as parenting style can influence the development and function of ToM (Vinden, 2001), and can result in people's ability to empathize and comprehend the empathic responses of others. Moreover, studies have reported that people from different cultures have different preferences with respect to emotional experience,

expression, and recognition (Mesquita & Leu, 2007). Taking it into account, it is plausible that German users might have had fewer expectations of receiving empathic responses because such expectations have not been nurtured during their upbringing process. However, it is important that in further research, the expectation to see emotional responses from the virtual agent is specifically tested.

People in collectivistic societies are more integrated into groups and they emphasize emotional dependence (Darwish & Huber, 2003; Hofstede, 2011). Therefore, caring about others and protecting the groups are highly valued in those cultures. Moreover, as the level of collectivism values increases in a society, it is reported that people show higher perspective-taking skills and empathic concerns for others (Heinke & Louis, 2009). By contrast, people in individualistic cultures mostly focus on promoting the individual's and their immediate family's interest, privacy, self-realization, and personal autonomy (Darwish & Huber, 2003; Hofstede, 2011). Therefore, they place less emphasis on ties between people and looking after each other. Accordingly, it can be stated that caring about others is not expected in individualistic cultures, as it is in collectivistic ones. Considering empathizing behaviors as "caring" behaviors (Brave et al., 2005), it is plausible that Iranian participants – who owing to collectivistic tendencies are assumed to care about the needs of others – might have expected the agent to care about them. Therefore, because of such expectations, the emotional nonverbal behaviors of the agent attracted their attention more compared with their German peers, hence, led them to evaluate the agent as more empathic.

As mentioned earlier, the different ratings of the intention of the agent could be due to different conceptualizations of ToM, which is assumed to stem from the dimension of collectivism/individualism (see e.g., Bradford et al., 2018). However, it is important that future studies additionally test the skill of ToM for each target group as well in order to draw a firmer conclusion.

The explanation can further be supported by the significant differences between Iranian users' ratings of the trustworthiness and helpfulness of the agent. The results of Study 2 indicated that Iranian users perceived the agent as more trustworthy and helpful compared with German users. The higher perception of these two items was correlated with the higher perception of Iranian users of the agents' empathic behavior. It is important to note that the presence of the nonverbal behavior significantly influenced users' evaluation of the agent, suggesting that

displaying nonverbal responses indeed mattered in users' perceptions. The higher perception that the agent was trustworthy and helpful can be interpreted as the effect of the agent being empathic: Iranian users comprehended the agent as being "caring" since it showed empathic concerns and in line with this as trustworthy and helpful. This, again, is consistent with the findings of Brave and colleagues (2005), who suggested that agents that can show empathic emotions are seen as more trustworthy by the users. However, the finding must be interpreted with caution because there was no interaction effect between the users' culture and the agent's nonverbal behaviors for rating the item trustworthy.

The data further demonstrated that compared with the neutral behavior, all emotional nonverbal behaviors significantly affected Iranian users' evaluation of empathy. The evaluations of the German participants, however, indicated that they perceived the agent as more empathic only when it showed the nonverbal behavior Head Nod. Explaining these differences in terms of cultural preferences, it can be said that all the nonverbal behaviors, with their diverse intensities in showing emotions, led the Iranian participants to perceive the agent as responsive, and therefore, empathic. It is plausible that different ways of caregiving and moral education in the two cultures led to different rating of the Iranian users. Accordingly, they possibly needed to observe any sort of emotional responses, regardless of how potent that response was, in order to feel that the agent was empathizing. Although the mean values show that Iranian users perceived Head Nod like other nonverbal behaviors as an empathic response, compared with other responses, they gave this nonverbal behavior lower scores. One possible explanation might be that even though at first glance it seems that merely the response itself was more important than its intensity for Iranian users, the intensity did in fact play a role in their evaluation of how empathic the agent was.

By contrast, the fact that German users evaluated only one nonverbal behavior, Head Nod, as an empathic response might stem from their cultural norms that value rationality. Accordingly, the only nonverbal behavior that German users rated as more empathic was the one that was not so intense and overemphasized and could be described as the most rational response. It is important to note that German users also cared about the emotional response of the agent. This is because they evaluated the agent positively (as empathic) when the agent showed a response – albeit the plainest one – compared with no response at all (neutral). This is in line with other findings, suggesting that empathic agents influence the positive perception of the users

(e.g., Brave et al., 2005; Klein et al., 2002; Prendinger et al., 2005b) even though there are cultural differences with regard to this perception. These findings can make an important contribution to designing a culture-specific agent. In this regard, if the aim of designing an agent is mostly centered on providing support for the users, it might be more effective to implement more intensified emotional nonverbal behaviors in an empathic agent for users with collectivistic tendencies, and plainer nonverbal behaviors for individualistic users.

What was the effect of users' cultural stereotypes on perceiving the virtual agent?

The third user factor that was investigated in this dissertation was users' attitude (potential bias) toward an agent that was designed and produced in a different country (Study 3). As media equation theory suggests, people can apply the same social rules that they use in their everyday interactions with other humans to their interactions with virtual entities (Reeves & Nass, 1996). In this respect, Study 3 attempted to test whether people apply ingroup/outgroup stereotypes toward the agent. This was not based on visual cues suggesting the cultural background of the agent, but solely by having the knowledge that the agent was designed and produced in a different country. For this purpose, young German participants were primed about the country of the agent in two different cover stories, which stated that the agent was designed and produced either in Germany or in Malaysia.

The results indicated that users' attributes toward the agent did not impact their perception of how empathic the agent was. One possible explanation can be that when people do not observe any cues that show the ethnicity of the agent, their attention does not get directed to notice the possible cultural differences between them and the agent. Since the virtual agent, Billie does not represent any specific culture, mere information about its country could not influence the two primed groups' understanding of how empathic it was. However, further analyses were conducted to test how all participants, independent of the groups they had been assigned to, exhibited preferences for specific nonverbal behaviors.

With regard to rating the responses as empathic, the participants perceived the following nonverbal behaviors as the most empathic ones: Dropping the Arms plus Sad Face, Head Down, Head Nod plus Smile, Sad Face, and Smile. It can be seen that when the agent showed body

movements in combination with the facial expressions, the users rated it as more empathic. This suggests that the intensity of the responses mattered for the participants in Study 3

The lack of difference between the two groups' perceptions is not limited to ratings of the item empathy. The other evaluations, for instance, regarding the agent's unfriendliness, reactivity, and rating the attributes of warmhearted/cold, incompetent/competent, and inactive/active, indicated that people's perceptions did not vary significantly. Moreover, to determine whether users' bias toward the agent could have a direct influence on their opinion that the agent was behaving appropriately or not, the participants were presented with a scale to rate the appropriateness of the agent's behavior. Similar to other evaluations, no differences were observed between the users' evaluations in this respect. Additionally, the two groups did not differ in their responses regarding further usage of the agent, suggesting, again, that they were not biased by the fact that the agent was from a different country.

The absence of any differences in all evaluations might be that the agent's display of emotional responses could have influenced participants' perception of the agent. The ratings of the nonverbal behaviors revealed that compared with the neutral behaviors almost all of the emotional responses were rated as empathic by all participants. Having an understanding that the agent is trying to be empathic, therefore, might have reduced the negative stereotypes toward the agent, leading to no differences between the evaluations of the two groups. The assumption that the existence of something positive between the users and the agent might change users' negative attitude was first made in the study by Vang and Fox (2014). In their study, they suggested that something positive such as sharing a task between White users and Black agents could have been the reason why people had less negativity toward the Black agent. Considering that the positive effect of emotional agents has long been proclaimed in different studies (Brave et al., 2005; Mascarenhas et al., 2013; Paiva et al., 2017), it can be assumed that it decreased users' stereotypes toward the agent.

Moreover, a follow-up study (see Chapter 10.5) was conducted to test whether users' evaluation of the agent was based on their knowledge of its design country only, or other factors, such as the appearance of the agent played a role. Accordingly, in 2 out of 4 conditions, the participants were asked to envision working with the virtual agent Billie, without seeing any picture of it. The other two conditions, however, included a picture of Billie, and the participants were asked to imagine working with it. In all four conditions, the participants were informed that

the agent is either from their own country (Germany) or Malaysia. The findings of the follow-up study also provided support for the findings of the third study.

Accordingly, it was shown that that merely having knowledge of the country of the agent does not make users biased toward the agent and, hence, does not change their rating of its empathic behavior. Comparing the participants' ratings demonstrated no differences between people's opinions that the agent would be capable of being an empathic agent. This can support the assumption that as long as people do not see a cue from the agent that explicitly shows cultural differences, such as differences in the look of the agent, the type of clothing, or accent, they cannot apply ingroup/outgroup stereotypes toward the agent.

The qualitative data also revealed that people did not specifically mention any possible cultural differences, both in Study 3 and in the follow-up study. They -independent of the group- mostly commented on the agent's way of dressing and its look, that is, it looked childish and, therefore, they could not take it seriously.

What was the moderating effect of users' experience with technology, mood, and level of empathy on perceiving the virtual agent?

As mentioned earlier, users' age and cultural background were found to have an influence on their rating of an empathic agent. However, in the first two studies, the moderating effect of users' experience with technology, mood, and level of empathy was not controlled for. Investigating the potential impact of these factors is important since in prior studies the influence of users' experience with technology was reported on users' nervousness with encountering the technology, the acceptance, and the probability of further usage of the specific technology (Agarwal & Prasad, 1999; Krämer et al., 2010; Wilkowska & Ziefle, 2009). Furthermore, people's mood was also demonstrated to have an effect on the inferences of the mental states and emotions of others (Converse et al., 2008; Juckel et al., 2018). And finally, how empathic people are can affect their judgment of the emotions of others (Baron-Cohen et al., 2001; Dimberg et al., 2011; Svetieva, & Frank, 2016).

Therefore, in the third study (also in the follow-up study) the moderating effect of these factors was examined. The results, however, indicated no significant moderating effect of

experience with technology and level of empathy on users' understanding of the empathic agent. Therefore, the findings of Study 3 (and the follow-up study) do not provide support to the findings of previous research that emphasized the role of users' computer literacy and experience with technology in perceiving the agent. However, the positive mood of the participants was found to have an effect on their evaluations (in Study 3 only). With regard to the possible effect of people's mood when encountering virtual agents, it is recommended that future studies test this factor in an interactive setting and with verbal responses as well. The effect of a conversation with the agent and the emotional responses that people can receive verbally from it might impact the users differently. Thus, it is possible that the individual differences in the aforementioned factors change the evaluation of the empathic agent in an interactive setting.

What was the effect of specific nonverbal behaviors?

One of the main purposes of this dissertation was to investigate whether user variables affect users' preferences of specific nonverbal behaviors of an agent in order to evaluate it as empathic. The results of the first two studies indicated that user groups who have diverse ages and cultural backgrounds perceive different nonverbal behaviors as empathic responses. However, priming the users about the design country of the agent did not bias them to evaluate different nonverbal behaviors as empathic responses.

It is important to note that in all three studies the nonverbal behaviors of the agent influenced the participants' perception of the agent. The significant effect of nonverbal behaviors on seniors' ratings suggests that they were indeed capable of noticing the differences in the nonverbal responses. Therefore, they rated the agent as empathic and trustworthy. The effective role of nonverbal behaviors for seniors is described in Chapter 5. It has been reported that seniors can specifically benefit from utilizing nonverbal behaviors which enables them to initiate, enhance, and maintain spoken conversations and makes them be more engaged in social interactions (Hubbard et al., 2002). Using nonverbal behaviors as a tool to maintain social relations could have been the reason that older users paid more attention to the displayed nonverbal behavior of the agent in Study 1.

Comparing the responses in Study 3 with the most empathically perceived responses of German users in Study 2, however, shows inconsistencies between the findings. Young German

participants in the second study rated only one nonverbal behavior, Head Nod, as the most empathic one. This inconsistency may be due to the different methods followed in the studies. In Study 2, people participated in the study online, and although they were instructed to perform the study on a laptop/Personal Computer, it cannot be definitely stated that they followed the instruction. Assuming that the participants used their cellphone increases the probability that they did not notice the differences between nonverbal behaviors on a small screen. It is, therefore, recommended that future research conduct a cultural study in a more controlled setting in order to avoid varying usage of the technical devices to perform the study.

Furthermore, it is important to note that how experienced users were in encounters with technologies was not controlled for in Study 2. This factor could have influenced users' expectations and preferences of nonverbal behaviors. Gamers, for instance, are reported to expect to see more sophisticated behaviors of a virtual agent (Krämer et al., 2010). It is recommended that future cultural studies measure this factor as well. Additionally, using a more developed method to create nonverbal behaviors, such as motion capturing with post-processing of bones, facial expressions, gestures, etc., can contribute to reflecting on details of the nonverbal behaviors. This can lead to a more sophisticated look of the agent, especially for experienced users.

A detailed look at the perceived nonverbal behaviors across all three studies indicates that the very intense nonverbal behaviors attracted more attention. It seems that these behaviors can be specifically utilized for people who might need and expect to see emotional responses more (such as seniors and people with collectivistic values). According to the findings, a simpler and shorter response such as Head Nod did not attract the attention of seniors and Iranian users. Therefore, it provides evidence that the perception of empathy is higher when the agent shows a strong nonverbal response, with more focus on body movement and facial expressions.

Additionally, these intense emotional expressions, such as Head Nod plus Smile, were found to be effective in giving the impression that the agent is more friendly, reactive, warm, and competent. Since Smile and Head Nod plus Smile both received higher scores regarding the appropriateness of the behavior (in Study 3), it is plausible that smiling behavior plays an important role in users' positive evaluations. Therefore, it can be suggested that when the agent shows responses with body movement, and also responses that include a smiling facial

expression, it attracts the attention of the users more, leading to higher positive ratings of the agent.

Interestingly, even though Head Nod did not increase the positive perception of the users (except for the rating that the agent is active) significantly in Study 3, it still had higher mean values than the neutral behavior. This suggests that people were able to differentiate between the agent being responsive and not responsive at all. However, this interpretation needs to be made with caution owing to the inconsistent findings of Study 2 and Study 3. Further research will need to investigate this specific nonverbal response for young German users (and generally, users with individualistic tendencies) in a more controlled setting.

12 Theoretical Implications

The present dissertation provides empirical evidence for theoretical assumptions that users' perception of an emotional virtual entity can be explained by their skill of ToM and by predictions of media equation theory. Having the ability of ToM can explain people's attempts to understand the intentions behind the actions of others and predict their behaviors. In order to have fruitful social interactions, it is crucial that people understand the mental states and intentions of others. They also need to know that these mental states and expectations can have an impact on the behaviors of others. As social relationships can be complicated by misunderstandings, it can be said that successful social interactions are dependent on having the skill of ToM.

Human-agent interactions, owing to many shortcomings, are more prone to misunderstandings. Therefore, to reduce the complexity of the interaction and possible misunderstandings, implementation of ToM in embodied agents has been examined in many studies (e.g., Belkaid & Sabouret, 2014; Hoogendoorn & Soumokil, 2010; Marsella & Pynadath, 2005; Peters, 2006). Notably, studies that investigate this ability from the perspective of the users are sparse. When encountering an emotional agent that intends to be empathic, it is important to have an understanding of how the users perceive that agent.

As mentioned in Chapter 3.2, there is a mixed pattern in the results of studies that investigated whether ToM and understanding the emotions and intentions of others are influenced by the aging process. For example, many scholars provided support for the sociocognitive

approach – that suggests older people are more capable of understanding emotions thanks to their extensive experiences across the life span – by showing that ToM remains intact during aging (see, e.g., Happé et al., 1998; Keightley et al., 2006). By contrast, the findings of many other researchers contradicted this claim and showed consistent results with the claim of the neuropsychological approach. These studies demonstrated that the general cognitive decline of older people decreases their ability to accurately recognize the emotions of others (see, e.g., Maylor et al., 2002; Sullivan & Ruffman, 2004).

The findings of Study 1 showed that age can influence users' ratings of the empathic agent. Consistent with the claims of the sociocognitive approach, the older users showed a higher perception of the empathic behaviors of the virtual agent. This can suggest that the ability to understand the intentions of others may not totally be under the influence of aging and people in their late adulthood are able to correctly recognize the empathic responses of an agent. The findings can support the claim that people's experiences of emotional situations can enhance their ability to understand the emotional responses of others. This is further supported by the fact that older users could differentiate between various exhibited nonverbal behaviors, and it affected their understanding of the intention of empathizing behavior by the agent.

Additionally, with regard to the cultural effect of ToM, the present findings demonstrated that when people have collectivistic tendencies – even though the tendencies are not assumed to be strong – they show different perceptions of the empathic agent. The results of Study 2 support the idea that ToM can be guided by cultural norms (see Mc Grath, 2009), suggesting that it can then influence comprehending the attempts of empathizing by others. Since the Iranian participants exhibited a higher perception of empathy in comparison with their German peers, the assumption that collectivistic inclinations impact users' perception of and expectations from an emotional agent can be supported. This is consistent with the collectivistic culture's emphasis on interpersonal harmony (Triandis, 1993) and on being sensitive and responsive to the communicative purpose of others (Harb & Smith, 2008; Shahaiein et al., 2014).

The fact that Iranian and German users rated different nonverbal behaviors as empathic is in line with studies suggesting that individualism/collectivism can influence verbal and nonverbal ways of declaring emotions (Cheon et al., 2010). Therefore, it is assumed that due to these cultural differences in expressing emotions, collectivistic users expected and preferred to see – any – emotional responses, while individualistic users instead saw the most rational response in

their opinion. Hence, it can be said that depending on whether a culture is collectivistic or individualistic, people's way of expressing, and therefore understanding the emotions of others might vary.

This finding has important implications for designers and developers. It can be argued that ToM as a basic requirement for human interactions can play a significant role in human-agent interaction as well. In this regard, besides the attempts that have been done to implement ToM in virtual entities, developers are recommended to consider the factors that can affect users' ToM as well. For example, according to the findings of this dissertation, accurate comprehension of the agent's emotional responses can be influenced by users' age and cultural tendencies. Designers can work in the direction of optimizing agents, that take the perspective of their users, and act in line with their expectations. Therefore, cross-cultural studies and studies that consider different age cohorts can contribute to providing insights into understanding how ToM functions in different cultures and different age stages, and how it all can be applied to interactions with virtual entities.

Additionally, the results of the first two studies appear to support the assumptions of media equation theory as well. The general assumption is that older people and younger adults (due to cognitive abilities or experiences in life), and collectivists and individualists (due to cultural differences) perceive emotional others differently. The findings of this dissertation indicated that people of varying ages and from a diverse cultural background showed social reactions to the virtual agent (attributing empathy to the agent) as they were expected to show the same social reactions to another human. Accordingly, people exhibited different perceptions of empathic behaviors (of an agent) due to their age and cultural background.

However, the result of Study 3 did not prove that people apply cultural stereotypes toward an empathic agent that is from another country. Therefore, it needs to be carefully assumed that media equation assumptions also apply for users' stereotypes toward the agent based on solely the information people have of the agent's country of origin.

The assumption that people show stereotypical attitudes were made on the basis of the findings of prior studies indicating the existence of stereotypes toward an agent with a different ethnicity (e.g., Dotsch & Wigboldus, 2008; Eastwick & Gardner, 2009; Lee & Nass, 1998). It is important to note that in these studies, the ethnic differences between the users and the agents

were represented in the appearance of the agent. In Study 3, the aim was to test whether users' ingroup/outgroup stereotypes are applicable when they know about the design and production country of the agent, without any apparent features indicating any cultural differences. The findings did not support the idea of applying cultural stereotypes to the agent based on this knowledge only. The data reported here, however, appear to support the assumption that when there is something positive between the agents and the users, such as sharing a task, it reduces users' stereotypical attitudes (Vang & Fox, 2014). It can be assumed that the agent's attempt to be empathic was perceived accurately by all participants, and thus, reduced any possible negative stereotypes toward the agent.

Moreover, the lack of differences between the two primed groups can be interpreted in line with the suggestion of prior research. Accordingly, the reason people apply social rules to virtual entities is due to the existence of anthropomorphic cues in those entities (Kim et al., 2012). It appears that the reason the participants did not show any bias toward the agent could stem from the absence of specific cues in the appearance of the agent. However, it should be considered that the manipulation check failed in this study because the participants were not asked at the end of the survey whether they could remember in which country the agent was designed. Therefore, the results need to be interpreted with caution.

13 Practical Implications

Practical implications for the design of an empathic agent can be derived from the empirical studies of this dissertation. It is recommended that designers and developers consider their target groups when designing an empathic virtual agent. The target groups are especially important when the agent has a specific purpose to fulfill, such as assisting the users. Accordingly, based on the findings of Study 1, it is recommended that the age of participants be specifically considered when designing empathic nonverbal behaviors for an agent. Older users seem to be more sensitive to the empathic responses of the agent and show trusting behavior in encounters with an empathic agent. This makes the implementation of empathy for these people crucial, since the trustworthiness of an agent, especially a virtual assistant, can predict the probability of future re-usage of the technology. These findings are important for designers of virtual assistants aimed at making users feel supported in their home environment.

When designing nonverbal behaviors to display empathy, older users are more sensitive to the intensity of the nonverbal behaviors as well. It is important to note that the older participants in Study 1 were able to accurately perceive the agent as empathic when it displayed facial expressions (Smile and Sad Face). However, they rated the agent as more empathic when the agent showed these expressions in combination with body movement. This suggests that although older users did not show difficulties in relating the facial expressions to the empathizing responses of the agent, it is better to create more intense emotional behaviors including body movements, to make the nonverbal responses more effective for them.

Furthermore, designers of virtual agents should consider that cultural preferences and expectations of users can influence their perception of the agent. Therefore, depending on the purpose of the design (e.g., whether the agent is aimed to be effective in assisting the users) it is crucial to pay attention to the cultural tendencies of the target group. The findings of Study 2 indicated that, depending on different emotional nonverbal responses, people with collectivistic and individualistic tendencies show different perceptions of an empathic agent. When the purpose of designing an agent is to be assistive, it is important to consider that collectivists value “caring” more, thus might need and expect to see empathic responses in emotional situations. Since collectivistic users perceived the agent as helpful and trustworthy as well, it can be stated that implementing empathic responses can fulfill the agent’s aim of being helpful for this target group specifically.

Moreover, collectivists showed different preferences regarding specific nonverbal behaviors compared with people with individualistic inclinations. Based on the findings it can be stated that the perception of empathy can be increased for collectivists when body movement is implemented in the nonverbal response as well. More specifically, intense emotional nonverbal behaviors that include both facial expression and body movement appear to be influential for these groups of users. This suggests that for users from a cultural background that values expressing emotions, highly intensive emotional responses can be comprehended as more effective. One speculation is that when the response is stronger, they are more certain of the agent’s intention, which is empathizing. Therefore, developers may consider a culture-specific design of nonverbal behaviors for Middle-Eastern users (or users who have collectivistic tendencies) to enhance the quality of their interactions with the agent.

It is also important that designers consider that cultures that do not place importance on the need to receive emotional responses might affect their people's preferences of seeing specific emotional nonverbal behaviors. Based on the findings of Study 2, it can be assumed that more plain nonverbal responses are more effective for users with individualistic tendencies.

However, participants in Study 3, who were also young German users, did not rate the plainest nonverbal behavior as the most empathic one. It is important to note, however, that people in Study 2 performed the study online, which lowers the control over the setting, especially the screens they used while watching the videos. Therefore, in order to derive a conclusion about individualists' preferences for specific nonverbal behaviors, it is recommended that future research replicate the effectiveness of the behaviors for this target group in a laboratory experiment.

What is remarkable in the findings of Study 3 is that knowledge of the design country of the agent solely does not influence users' attitudes toward the agent. This may help us to understand that when the agent does not represent any cues related to specific cultures, it has the possibility to diminish users' bias toward it. This finding has important implications for developing neutral-looking agents and provides the basis for future research in the domain of agent appearance.

In sum, this dissertation provides insights into the importance of focusing on user variables in design guidelines for nonverbal behaviors of empathic agents. Designers and developers are thus recommended to consider different versions of nonverbal behaviors for different target groups (e.g., specifically designed for older users and culture-specific behaviors for users worldwide). It is more effective that virtual agents show strong and intense emotional nonverbal behaviors for older people and people with collectivistic inclinations, in order to show its intention of empathizing in a more obvious way.

14 Limitations and Future Directions

Finally, a number of important limitations need to be considered. First, the settings of the studies in this dissertation were not interactive and the users were not able to respond to the agent. Since it can be assumed that in an interaction a higher level of attention might be needed

(especially for the target group of older people), people may show more difficulties in correctly understanding the responses of the agent. Furthermore, technical failures that might occur in an interactive scenario may lead to a misunderstanding between the interlocutors and, thereby, to user frustration. It is thus suggested that the effectiveness of different emotional nonverbal behaviors be tested in this kind of setting.

Additionally, the important role of ToM in understanding the emotions and intentions of others has been emphasized in prior studies and in this dissertation. However, in order to have a more precise knowledge of users' abilities of ToM (in different ages and different cultures), it is recommended that related tasks that evaluate this ability in adulthood be presented to the users before the study. Also, assessing the collectivistic/individualistic tendencies of the users, their willingness to care for others, and their expectations to receive emotional responses in emotional situations can contribute to derive precise conclusions regarding the cultural variables of the users.

Moreover, as the open-ended questions were not presented to the participants in the first two studies, no in-depth knowledge of older adults' and collectivistic users' preferences could be obtained. Future studies that consider these types of questions, or interviewing, could contribute to more precise findings. Also, because Study 2 was not laboratory-based, and the younger users in Study 1 were recruited online, it is recommended that future studies that consider these target groups investigate their perceptions in a more controlled setting. This way, it can be assured that all users watch the agent's exhibition of nonverbal behaviors on a similar scale and screen.

Study 1 and 2 are also limited by the lack of information on users' prior experience with technology. For example, considering the possibility that the usage of (or even confrontation with) virtual agents might vary in Germany and Iran, it is plausible that users have different expectancies of the capabilities of agents. Further work needs to be done to establish whether the experience with these technologies has a moderating effect on users' perception of an empathic agent.

Since the users in the last study mentioned their negative evaluation regarding the look of the agent (childish and not serious), it is recommended that future studies assess the effectiveness of emotional nonverbal behaviors with a different virtual agent. And finally, it is important to note that although the focus of this dissertation was on nonverbal behaviors, in order to express

the feeling of empathy, both verbal and nonverbal responses are likely to be more effective. Investigating the effect of individual variables on the evaluation of the combinations of these types of responses could then be taken into account.

15 Conclusion

The effect of user variables on the perception of an empathic agent based on its displayed emotional nonverbal behaviors was investigated in this dissertation. To this end, users' age, culture, and potential cultural stereotypes toward the agent were assessed. As the usage of virtual entities is increasing, it is important to have knowledge about users' perceptions and the factors that might influence their understanding of an agent. The results presented in this dissertation demonstrate that there are differences in perceptions, at least with regard to users' age and culture, and that considering them in the design of agents can contribute to qualified human-agent interactions.

The results of Study 1 indicated that age is an important factor in evaluating the emotional nonverbal behaviors of the agent. The older adults perceived the agent as more empathic, and in line with that, as more trustworthy compared with young people. This perception was due to the exhibition of emotional nonverbal behaviors after the agent uttered emotional sentences. Therefore, empathic agents can be specifically beneficial for older users who might have less experience with technology, as the empathic responses increase their trust in agents.

Investigating which specific nonverbal behaviors triggered the higher perceptions of empathy in older users indicated the effectiveness of the following responses: Dropping the Arms plus Sad Face, Head Down, Sad Face, Head Nod plus Smile, and Smile. Notably, the older users showed higher preferences for intense nonverbal behaviors (such as Dropping the Arms plus Sad Face and Head Down). When these preferences are considered in the design and production of the agent, a more positive outcome regarding their qualified interaction with the agent can be expected.

Furthermore, this dissertation provided insight into understanding how the cultural background of the users can influence their perception of an empathic agent based on its displayed nonverbal behaviors. In Study 2, it was observed that Iranian users as Middle-Eastern

samples with collectivistic tendencies showed higher perceptions of the agent's empathy, trustworthiness, and helpfulness compared with German users with individualistic inclinations. Similar to Study 1, the nonverbal behaviors of the agent were found to have an effect on users' evaluations in Study 2. Interestingly, Iranian users perceived all displayed emotional responses as empathic, suggesting that for users' who have collectivistic inclinations any emotional responses might be needed. However, they showed high preferences for more strong responses (such as Dropping the Arms plus Sad Face and Head Down) by which they could readily understand the intention of the agent. The data revealed that German users, by contrast, assessed only Head Nod as the most empathic response. Taken together, the findings can provide insights into designing a culture-specific agent, when the target groups have different cultural tendencies.

And finally, this dissertation could not provide support for the assumption that when the users are aware of the design and production country of the agent, they are biased toward it and, hence, perceive its empathic behaviors differently. Moreover, the assessment of the perceived nonverbal behaviors in Study 3 indicated inconsistent results regarding young German users' evaluations of the agent (as compared with the results of Study 2). Accordingly, it is recommended that for a better understanding of how individualists perceive an agent's emotional nonverbal behaviors, studies with a more controlled setting be conducted.

In summary, user variables are important and should be considered when evaluating an empathic agent, especially with respect to its emotional nonverbal behaviors. These sorts of investigations can provide us with a deeper knowledge of human-agent interactions from the perspective of the users.

References

- Agarwal, R., & Prasad, J. (1999). Are individual differences germane to the acceptance of new information technologies?. *Decision sciences*, *30*(2), 361-391.
- Alharbi, O., & Arif, A. S. (April 22, 2018). The Perception of Humanoid Robots for Domestic Use in Saudi Arabia. Paper. In CHI 2018 Workshop on Exploring Participatory Design Methods to Engage with Arab Communities. Montréal, QC, Canada: CHI Retrieved from *arXiv preprint arXiv:1806.09115*.
- Allen, V. L., & Wilder, D. A. (1979). Group categorization and attribution of belief similarity. *Small Group Behavior*, *10*(1), 73-80.
<https://doi.org/10.1177%2F0090552679101006>
- Ambady, N., LaPlante, D., Nguyen, T., Rosenthal, R., Chaumeton, N., & Levinson, W. (2002). Surgeons' tone of voice: a clue to malpractice history. *Surgery*, *132*(1), 5-9.
<https://doi.org/10.1067/msy.2002.124733>
- Apperly, I. A., Samson, D., Chiavarino, C., & Humphreys, G. W. (2004). Frontal and temporo-parietal lobe contributions to theory of mind: neuropsychological evidence from a false-belief task with reduced language and executive demands. *Journal of Cognitive Neuroscience*, *16*(10), 1773-1784. <https://doi.org/10.1111%2Fj.1467-9280.2006.01791.x>
- Arnett, J. J. (2008). The neglected 95%: why American psychology needs to become less American. *American Psychologist*, *63*(7), 602-614. DOI: 10.1037/0003-066X.63.7.602
- Avis, J., & Harris, P. L. (1991). Belief-desire reasoning among Baka children: Evidence for a universal conception of mind. *Child development*, *62*(3), 460-467.
<https://doi.org/10.1111/j.1467-8624.1991.tb01544.x>
- Aylett, R., Vannini, N., Andre, E., Paiva, A., Enz, S., & Hall, L. (2009). But that was in another country: agents and intercultural empathy. In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems - Volume 1* (pp. 329-336). (AAMAS '09). Richland, SC: International Foundation for Autonomous Agents and Multiagent Systems.
- Bailey, P. E., & Henry, J. D. (2008). Growing less empathic with age: Disinhibition of the self-perspective. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *63*(4), 219-226. DOI: 10.1093/geronb/63.4.P219
- Bailey, P. E., Henry, J. D., & Von Hippel, W. (2008). Empathy and social functioning in late adulthood. *Aging and Mental Health*, *12*(4), 499-503. DOI: 10.1080/13607860802224243
- Baron-Cohen, S. (1997). *Mindblindness: An essay on autism and theory of mind*. MIT press.
- Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a “theory of mind”? *Cognition*, *21*(1), 37-46. DOI: 10.1016/0010-0277(85)90022-8
- Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The “Reading the Mind in the Eyes” Test revised version: a study with normal adults, and adults with

- Asperger syndrome or high-functioning autism. *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 42(2), 241-251.
- Batson, C. D. (1990). How social an animal? The human capacity for caring. *American Psychologist*, 45(3), 336–346. <https://doi.org/10.1037/0003-066x.45.3.336>
- Batson, C. D., Fultz, J., & Schoenrade, P. A. (1987). Distress and Empathy: Two Qualitatively Distinct Vicarious Emotions with Different Motivational Consequences. *Journal of Personality*, 55(1), 19–39. <https://doi.org/10.1111/j.1467-6494.1987.tb00426.x>
- Baylor, A. L. (2005, January). The impact of pedagogical agent image on affective outcomes. In *Proceedings of Workshop “Affective Interactions: The Computer in the Affective Loop” at the 10th International Conference on Intelligent User Interfaces* (pp. 1-6). San Diego, CA: ACM Press.
- Baylor, A., & Kim, Y. (2003). The role of gender and ethnicity in pedagogical agent perception. In *E-Learn: World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education* (pp. 1503-1506). Association for the Advancement of Computing in Education (AACE).
- Baylor, A. L., & Kim, Y. (2004, August). Pedagogical agent design: The impact of agent realism, gender, ethnicity, and instructional role. In J.C. Lester, R.M. Vicari, F. Paraguaçu (Eds.), *International conference on intelligent tutoring systems* (pp. 592-603). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-540-30139-4_56
- Baylor, A. L., Kim, S., Son, C., & Lee, M. (2005, May). Designing effective nonverbal communication for pedagogical agents. In *Proceedings of the 2005 conference on Artificial Intelligence in Education: Supporting Learning through Intelligent and Socially Informed Technology* (pp. 744-746). IOS Press.
- Becker, C., Kopp, S., & Wachsmuth, I. (2007). Why emotions should be integrated into conversational agents. *Conversational Informatics*, 49–67. <https://doi.org/10.1002/9780470512470>
- Becker, B., & Mark, G. (2002). Social Conventions in Computermediated Communication: A Comparison of Three Online Shared Virtual Environments, (December), 19–39. https://doi.org/10.1007/978-1-4471-0277-9_2
- Beer, J. M., Smarr, C. A., Fisk, A. D., & Rogers, W. A. (2015). Younger and older users' recognition of virtual agent facial expressions. *International journal of human-computer studies*, 75, 1-20. <http://dx.doi.org/10.1016/j.ijhcs.2014.11.005>
- Beier, G. (1999). Locus of control when interacting with technology (Kontrollüberzeugungen im Umgang mit Technik). *Report Psychologie*, 24(9), 684-693.
- Belkaid, M., & Sabouret, N. (2014). A logical model of theory of mind for virtual agents in the context of job interview simulation. *arXiv preprint arXiv:1402.5043*.
- Bickmore T., Cassell J. (2005). Social dialogue with embodied conversational agents. In: van Kuppevelt J.C.J., Dybkjær L., Bernsen N.O. (Eds.), *Advances in Natural Multimodal Dialogue Systems. Text, Speech and Language Technology* (pp. 23-54). Springer,

Dordrecht. https://doi.org/10.1007/1-4020-3933-6_2

- Blair, R. J. R. (2005). Responding to the emotions of others: Dissociating forms of empathy through the study of typical and psychiatric populations. *Consciousness and Cognition*, *14*(4), 698–718. <https://doi.org/10.1016/j.concog.2005.06.004>
- Blanchard-Fields, F. (2007). Everyday problem solving and emotion: An adult developmental perspective. *Current Directions in Psychological Science*, *16*(1), 26-31.
- Bradford, E. E., Jentsch, I., Gomez, J. C., Chen, Y., Zhang, D., & Su, Y. (2018). Cross-cultural differences in adult theory of mind abilities: a comparison of native-English speakers and native-Chinese speakers on the self/other differentiation task. *Quarterly Journal of Experimental Psychology*, *71*(12), 2665-2676.
- Brave, S., Nass, C., & Hutchinson, K. (2005). Computers that care: Investigating the effects of orientation of emotion exhibited by an embodied computer agent. *International Journal of Human Computer Studies*, *62*(2), 161–178. <https://doi.org/10.1016/j.ijhcs.2004.11.002>
- Brothers, L. (1989). A Biological Perspective on Empathy. *The American Journal of Psychiatry*, *146*, 10–19.
- Callaghan, T., Rochat, P., Lillard, A., Claux, M. L., Odden, H., Itakura, S., ... & Singh, S. (2005). Synchrony in the onset of mental-state reasoning: Evidence from five cultures. *Psychological Science*, *16*(5), 378-384. DOI: 10.1111/j.0956-7976.2005.01544.x
- Caravita, S. C., Di Blasio, P., & Salmivalli, C. (2010). Early adolescents' participation in bullying: Is ToM involved? *The Journal of Early Adolescence*, *30*(1), 138-170. <https://doi.org/10.1177/0272431609342983>
- Carney, D. R., Cuddy, A. J., & Yap, A. J. (2010). Power posing: Brief nonverbal displays affect neuroendocrine levels and risk tolerance. *Psychological science*, *21*(10), 1363-1368. doi: 10.1177/0956797610383437
- Carruthers, P., & Smith, P. K. (Eds.). (1996). *Theories of theories of mind*. Cambridge, UK: Cambridge University Press.
- Castelli, I., Baglio, F., Blasi, V., Alberoni, M., Falini, A., Liverta-Sempio, O., ... & Marchetti, A. (2010). Effects of aging on mindreading ability through the eyes: an fMRI study. *Neuropsychologia*, *48*(9), 2586-2594. <https://doi.org/10.1016/j.neuropsychologia.2010.05.005>
- Chasiotis, A., Kiessling, F., Hofer, J., & Campos, D. (2006). Theory of mind and inhibitory control in three cultures: Conflict inhibition predicts false belief understanding in Germany, Costa Rica and Cameroon. *International Journal of Behavioral Development*, *30*(3), 249-260. DOI: 10.1177/0165025406066759
- Cheon, B. K., & Mathur, V. A., & Chiao, J.Y. (2010). Empathy as Cultural Process: Insights from the Cultural Neuroscience of Empathy. *World Cultural Psychiatry Research Review*, *5*, 32-42.
- Chiao, J., & Ambady, N. (2007). Cultural neuroscience: Parsing universality and diversity across

- levels of analysis. In S. Kitayama & D. Cohen (Eds.), *The Handbook of Cultural Psychology* (pp. 237-254). New York, NY: Guilford Press.
- Comfort, A. (1984). *Reality and empathy: Physics, mind, and science in the 21st century*. Albany, NY: SUNY Press.
- Converse, B. A., Lin, S., Keysar, B., & Epley, N. (2008). In the mood to get over yourself: mood affects theory-of-mind use. *Emotion*, 8(5), 725.
- Chopik, W. J., O'Brien, E., & Konrath, S. H. (2017). Differences in empathic concern and perspective taking across 63 countries. *Journal of Cross-Cultural Psychology*, 48(1), 23-38. <https://doi.org/10.1177/0022022116673910>
- Clark III, R. D., & Maass, A. (1988). The role of social categorization and perceived source credibility in minority influence. *European Journal of Social Psychology*, 18(5), 381-394. <https://doi.org/10.1002/ejsp.2420180405>
- Crandall, C. S., & Eshleman, A. (2003). A justification-suppression model of the expression and experience of prejudice. *Psychological bulletin*, 129(3), 414. DOI:10.1037/0033-2909.129.3.414
- Danilava, S., Busemann, S., & Schommer, C. (2012). Artificial conversational companions a requirements analysis. In *Proceedings" 4th International Conference on Agents and Artificial Intelligence"* (pp. 282-289). SciTePress 2012.
- Darwish, A. F. E., & Huber, G. L. (2003). Individualism vs collectivism in different cultures: a cross-cultural study. *Intercultural Education*, 14(1), 47-56.
- Davis, M. H. (2018). *Empathy: A social psychological approach*. Routledge.
- Decety, J., & Jackson, P. L. (2004). *The functional architecture of human empathy*. *Behavioral and cognitive neuroscience reviews*, 3(2), 71-100. <https://doi.org/10.1177/1534582304267187>
- Degens, N., Hofstede, G. J., McBreen, J., Beulens, A., Mascarenhas, S., Ferreira, N., ... & Dignum, F. (2012, June). When agents meet: empathy, moral circle, ritual, and culture. In *Workshop on Emotional and Empathic Agents (EEA) at 11th International Conference on Autonomous Agents and Multiagent Systems*.
- Dennett, D. C. (1988). Précis of the intentional stance. *Behavioral and brain sciences*, 11(3), 495-505.
- DeSteno, D., Breazeal, C., Frank, R. H., Pizarro, D., Baumann, J., Dickens, L., & Lee, J. J. (2012). Detecting the trustworthiness of novel partners in economic exchange. *Psychological science*, 23(12), 1549-1556. <https://doi.org/10.1177%2F0956797612448793>
- Di Blasi, Z., Harkness, E., Ernst, E., Georgiou, A., & Kleijnen, J. (2001). Influence of context effects on health outcomes: a systematic review. *The Lancet*, 357(9258), 757-762. DOI:10.1016/s0140-6736(00)04169-6

- Dimberg, U., Andréasson, P., & Thunberg, M. (2011). Emotional empathy and facial reactions to facial expressions. *Journal of Psychophysiology*.
- Dotsch, R., & Wigboldus, D. H. (2008). Virtual prejudice. *Journal of experimental social psychology*, 44(4), 1194-1198. DOI: 10.1016/j.jesp.2008.03.003
- Downing, R. E., Moore, J. L., & Brown, S. W. (2005). The effects and interaction of spatial visualization and domain expertise on information seeking. *Computers in Human Behavior*, 21(2), 195-209.
- Dunn, J., Brown, J., & Beardsall, L. (1991). Family talk about feeling states and children's later understanding of others' emotions. *Developmental Psychology*, 27(3), 448-455.
- Eastwick, P. W., & Gardner, W. L. (2009). Is it a game? Evidence for social influence in the virtual world. *Social Influence*, 4(1), 18-32.
- Eid, M., & Diener, E. (2009). Norms for experiencing emotions in different cultures: Inter-and intranational differences. In: Diener E. (Eds.), *Culture and Well-Being. Social Indicators Research Series*, (pp. 169-202). Springer, Dordrecht.
- Eimler, S. C., Krämer, N. C., & Von Der Pütten, A. M. (2011). Empirical results on determinants of acceptance and emotion attribution in confrontation with a robot rabbit. *Applied Artificial Intelligence*, 25(6), 503–529. <https://doi.org/10.1080/08839514.2011.587154>
- Eisenberg, N., Spinrad, T. L., & Sadovsky, A. (2006). Empathy-related responding in children. In M. Killen, & J. Smetana (Eds.), *Handbook of moral development* (pp. 517–549). Mahwah, NJ: Erlbaum.
- Ekman, P. (2004). Emotions revealed. *BMJ*, 328(Suppl S5), 0405184. New York: Henry Holt & Co.
- Elfenbein, H. A. (2015). In-group advantage and other-group bias in facial emotion recognition. In *Understanding facial expressions in communication* (pp. 57-71). Springer, New Delhi.
- Elfenbein, H. A., & Ambady, N. (2002). Is there an in-group advantage in emotion recognition? *Psychological Bulletin*, 128(2), 243-249. <https://doi.org/10.1037/0033-2909.128.2.243>
- Elfenbein, H. A., & Ambady, N. (2003). Cultural similarity's consequences: A distance perspective on cross-cultural differences in emotion recognition. *Journal of Cross-Cultural Psychology*, 34(1), 92-110. <https://doi.org/10.1177/0022022102239157>
- Endrass, B., André, E., Rehm, M., Lipi, A. A., & Nakano, Y. (2011, May). Culture-related differences in aspects of behavior for virtual characters across Germany and Japan. In *The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 2* (pp. 441-448). International Foundation for Autonomous Agents and Multiagent Systems, Taipeh, Taiwan.
- Eysenck, S. B. G. (1981). Impulsiveness and antisocial behaviour in children. *Current Psychological Research*, 1(1), 31–37. <https://doi.org/10.1007/BF02684423>
- Ezer, N., Fisk, A. D., & Rogers, W. A. (2009, October). More than a servant: Self-reported

- willingness of younger and older adults to having a robot perform interactive and critical tasks in the home. In *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 53, No. 2, pp. 136-140). Sage CA: Los Angeles, CA: SAGE Publications.
- Finset, A., & Ørnes, K. (2017). Empathy in the Clinician–Patient Relationship. *Journal of Patient Experience, 4*(2), 64–68. <https://doi.org/10.1177/2374373517699271>
- Frith, U., & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 358*(1431), 459-473.
- Gefen, D., & Straub, D. W. (1997). Gender differences in the perception and use of e-mail: An extension to the technology acceptance model. *MIS quarterly, 389*-400.
- German, T. P., & Hehman, J. A. (2006). Representational and executive selection resources in ‘theory of mind’: Evidence from compromised belief-desire reasoning in old age. *Cognition, 101*(1), 129-152. <https://doi.org/10.1016/j.cognition.2005.05.007>
- Germain, M. H. (1998). *Worldly teachers: Cultural learning and pedagogy*. Westport, CT: Bergin & Garvey
- Ghorbani, N., Bing, M. N., Watson, P. J., Davison, H. K., & LeBreton, D. L. (2003). Individualist and collectivist values: Evidence of compatibility in Iran and the United States. *Personality and Individual Differences, 35*(2), 431-447. [https://doi.org/10.1016/S0191-8869\(02\)00205-2](https://doi.org/10.1016/S0191-8869(02)00205-2)
- Gini, G., Albiero, P., Benelli, B., & Altoè, G. (2007). Does empathy predict adolescents' bullying and defending behavior? *Aggressive Behavior: Official Journal of the International Society for Research on Aggression, 33*(5), 467-476. DOI: 10.1111/j.1532-7795.2012.00809.x
- Goleman, D. (1998). *Working with emotional intelligence*. New York, NY: Bantam Books.
- Greenfield, P. M., Keller, H., Fuligni, A., & Maynard, A. (2003). Cultural pathways through universal development. *Annual review of psychology, 54*(1), 461-490. <https://doi.org/10.1146/annurev.psych.54.101601.145221>
- Greenwood, P. M. (2000). The frontal aging hypothesis evaluated. *Journal of the International Neuropsychological Society, 6*(6), 705-726. <https://doi.org/10.1017/S1355617700666092>
- Grice, H. P. (1957). Meaning. *The philosophical review, 66*(3), 377-388.
- Halpern, J. (2003). What is clinical empathy? *Journal of General Internal Medicine, 18*(8), 670–674. <https://doi.org/10.1046/j.1525-1497.2003.21017.x>
- Happé, F. G. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of autism and Developmental disorders, 24*(2), 129-154. <https://doi.org/10.1007/BF02172093>
- Happé, F. G., Winner, E., & Brownell, H. (1998). The getting of wisdom: theory of mind in old age. *Developmental psychology, 34*(2), 358-362. <https://doi.org/10.1037/0012-1649.34.2.358>

- Harb, C., & Smith, P. B. (2008). Self-construals across cultures: Beyond independence—interdependence. *Journal of Cross-Cultural Psychology*, 39(2), 178-197. <https://doi.org/10.1177%2F0022022107313861>
- Heal, J. (1986). Replication and Functionalism. In J. Butterfield (Ed.): *Language, Mind, and Logic* (pp.135-150).Cambridge, UK: Cambridge University Press.
- Hedden, T., & Gabrieli, J. D. (2004). Insights into the ageing mind: a view from cognitive neuroscience. *Nature reviews neuroscience*, 5(2), 87. doi:10.1038/nrn1323
- Heinke, M. S., & Louis, W. R. (2009). Cultural background and individualistic–collectivistic values in relation to similarity, perspective taking, and empathy. *Journal of Applied Social Psychology*, 39(11), 2570-2590. <https://doi.org/10.1111/j.1559-1816.2009.00538.x>
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world?. *Behavioral and brain sciences*, 33(2-3), 61-83. <https://doi.org/10.1017/S0140525X0999152X>
- Heylen, D., Van Es, I., Nijholt, A., & van Dijk, B. (2002, June). Experimenting with the gaze of a conversational agent. In *Proceedings international CLASS workshop on natural, intelligent and effective interaction in multimodal dialogue systems* (pp. 93-100). Copenhagen: NIS Lab.
- Hoffman, M. L. (1994). The contribution of empathy to justice and moral judgment. *Reaching out: Caring, altruism, and prosocial behavior*, 7, 161-194.
- Hoffman, M. L. (2001). *Empathy and moral development: Implications for caring and justice*. Cambridge, UK: Cambridge University Press.
- Hofstede, G. (2001). *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations* (2nd ed). Thousand Oaks, CA: Sage.
- Hofstede, G. (2011). Dimensionalizing cultures: The Hofstede model in context. *Online readings in psychology and culture*, 2(1), 8. DOI:10.9707/2307-0919.1014
- Hofstede Insights *Country Comparison* (2019a). Retrieved from <https://www.hofstede-insights.com/country-comparison/china,iran/>.
- Hofstede Insights *Country Comparison* (2019b). Retrieved from <https://www.hofstede-insights.com/country-comparison/germany/>.
- Hofstede Insights *Country Comparison* (2019c). Retrieved from <https://www.hofstede-insights.com/country-comparison/germany,the-usa/>.
- Hofstede Insights *Country Comparison* (2019d). Retrieved from <https://www.hofstede-insights.com/country-comparison/iran/>.
- Hone, K. (2006). Empathic agents to reduce user frustration: The effects of varying agent characteristics. *Interacting with computers*, 18(2), 227-245. <https://doi.org/10.1016/j.intcom.2005.05.003>

- Hoogendoorn, M., & Soumokil, J. (2010, May). Evaluation of virtual agents utilizing theory of mind in a real time action game. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1-Volume 1* (pp. 59-66). International Foundation for Autonomous Agents and Multiagent Systems.
- Hubbard, G., Cook, A., Tester, S., & Downs, M. (2002). Beyond words: Older people with dementia using and interpreting nonverbal behaviour. *Journal of Aging Studies, 16*(2), 155-167. doi:10.1016/S0890-4065(02)00041-5
- Hughes, C., & Leekam, S. (2004). What are the links between theory of mind and social relations? Review, reflections and new directions for studies of typical and atypical development. *Social development, 13*(4), 590-619. <https://doi.org/10.1111/j.1467-9507.2004.00285.x>
- Hynes, C. A., Baird, A. A., & Grafton, S. T. (2006). Differential role of the orbital frontal lobe in emotional versus cognitive perspective-taking. *Neuropsychologia, 44*(3), 374-383. <https://doi.org/10.1016/j.neuropsychologia.2005.06.011>
- Ickes, W. (1993). Empathic Accuracy. *Journal of Personality, 61*(4), 587–610. <https://doi.org/10.1111/j.1467-6494.1993.tb00783.x>
- Ickes, W., Stinson, L., Bissonnette, V., & Garcia, S. (1990). Naturalistic social cognition: Empathic accuracy in mixed-sex dyads. *Journal of personality and social psychology, 59*(4), 730-742. DOI: 10.1037/0022-3514.59.4.730
- Ioannidou, F., & Konstantikaki, V. (2008). Empathy and emotional intelligence: What is it really about? *International Journal of caring sciences, 1*(3), 118-123.
- Isbister, K., & Nass, C. (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>
- Jan, D., Herrera, D., Martinovski, B., Novick, D., & Traum, D. (2007, September). A computational model of culture-specific conversational behavior. In *International Workshop on Intelligent Virtual Agents* (pp. 45-56). Berlin, Heidelberg: Springer.
- Jaques, P. A., & Viccari, R. M. (2005). Considering students' emotions in computer-mediated learning environments. *Web-Based Intelligent E-Learning Systems: Technologies and Applications, 122–138*. <https://doi.org/10.4018/978-1-59140-729-4.ch006>
- Juckel, G., Heinisch, C., Welpinghus, A., & Brüne, M. (2018). Understanding another person's emotions-An interdisciplinary research approach. *Frontiers in psychiatry, 9*, 414.
- Keightley, M. L., Winocur, G., Burianova, H., Hongwanishkul, D., & Grady, C. L. (2006). Age effects on social cognition: Faces tell a different story. *Psychology and Aging, 21*(3), 558–572. <https://doi.org/10.1037/0882-7974.21.3.558>
- Khan, R., & De Angeli, A. (2007, September). Mapping the demographics of virtual humans. In *Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI... but not as we know it-Volume 2* (pp. 149-152). BCS Learning & Development Ltd.

- Kim, Y., Baylor, A. L., & Shen, E. (2007). Pedagogical agents as learning companions: The impact of agent emotion and gender. *Journal of Computer Assisted Learning*, 23(3), 220–234. <https://doi.org/10.1111/j.1365-2729.2006.00210.x>
- Kim, Youjeong, & Sundar, S. S. (2012). Anthropomorphism of computers: Is it mindful or mindless? *Computers in Human Behavior*, 28(1), 241–250. <https://doi.org/10.1016/j.chb.2011.09.006>
- Klein, J., Moon, Y., & Picard, R. W. (2002). This computer responds to user frustration—Theory, design, and results. *Interacting with Computers*, 14(2), 119–140. [https://doi.org/10.1016/S0953-5438\(01\)00053-4](https://doi.org/10.1016/S0953-5438(01)00053-4)
- Koda, T. (2007, January). Cross-cultural study of avatars' facial expressions and design considerations within Asian countries. In *International Workshop on Intercultural Collaboration* (pp. 207-220). Springer, Berlin, Heidelberg.
- Koda, T., & Ishida, T. (2006, January). Cross-cultural study of avatar expression interpretations. In *International Symposium on Applications and the Internet (SAINT'06)* (pp. 7-pp). IEEE.
- Koda, T., Rehm, M., & André, E. (2008, September). Cross-cultural evaluations of avatar facial expressions designed by western designers. In *International Workshop on Intelligent Virtual Agents* (pp. 245-252). Springer, Berlin, Heidelberg.
- Krämer, N. C. (2001). *Bewegende Bewegung: sozio-emotionale Wirkungen nonverbaler Verhaltens und deren experimentelle Untersuchung mittels Computeranimation* [Moving movement. Socio-emotional effects of nonverbal behavior and their experimental investigation through computer animation]. Pabst.
- Krämer, N. C. (2005, September). Social communicative effects of a virtual program guide. In *International Workshop on Intelligent Virtual Agents* (pp. 442-453). Springer, Berlin, Heidelberg.
- Krämer, N. C. (2008). Theory of mind as a theoretical prerequisite to model communication with virtual humans. In *Modeling communication with robots and virtual humans* (pp. 222-240). Springer, Berlin, Heidelberg.
- Krämer, N. C., Eimler, S., von der Pütten, A., & Payr, S. (2011). Theory of companions: what can theoretical models contribute to applications and understanding of human-robot interaction?. *Applied Artificial Intelligence*, 25(6), 474-502.
- Krämer, N. C., Hoffmann, L., & Kopp, S. (2010, September). Know Your Users! Empirical Results for Tailoring an Agent's Nonverbal Behavior to Different User Groups. In *International Conference on Intelligent Virtual Agents* (pp. 468-474). Springer, Berlin, Heidelberg.
- Krämer, N. C., Simons, N., & Kopp, S. (2007, September). The effects of an embodied conversational agent's nonverbal behavior on user's evaluation and behavioral mimicry. In *International Workshop on Intelligent Virtual Agents* (pp. 238-251). Springer, Berlin, Heidelberg.
- Kraml, H., & Michlmayr, M. (2002). Simulation versus theory theory. theories concerning the

- ability to read minds (Masters thesis). *Innsbruck, Austria, Leopold-Franzens University*.
- Langenhove, L. V., & Harré, R. (1994). Cultural stereotypes and positioning theory. *Journal for the Theory of Social Behaviour*, 24(4), 359-372. <https://doi.org/10.1111/j.1468-5914.1994.tb00260.x>
- Lamm, C., Batson, C. D., & Decety, J. (2007). The neural substrate of human empathy: Effects of perspective-taking and cognitive appraisal. *Journal of Cognitive Neuroscience*, 19(1), 42–58. <https://doi.org/10.1162/jocn.2007.19.1.42>
- Lee, Y. T. (1993). Ingroup preference and homogeneity among African American and Chinese American students. *The Journal of social psychology*, 133(2), 225-235. <https://doi.org/10.1080/00224545.1993.9712140>
- Lee, E. J. (2003). Effects of “gender” of the computer on informational social influence: the moderating role of task type. *International Journal of Human-Computer Studies*, 58(4), 347-362.
- Lee, E. J., & Nass, C. (1998, October). Does the ethnicity of a computer agent matter? An experimental comparison of human-computer interaction and computer-mediated communication. In *Proceedings of the 1998 Workshop on Embodied Conversational Characters*.
- Leite, I., Castellano, G., Pereira, A., Martinho, C., & Paiva, A. (2014). Empathic Robots for Long-term Interaction: Evaluating Social Presence, Engagement and Perceived Support in Children. *International Journal of Social Robotics*, 6(3), 329–341. <https://doi.org/10.1007/s12369-014-0227-1>
- Leite, I., Pereira, A., Mascarenhas, S., Martinho, C., Prada, R., & Paiva, A. (2013). The influence of empathy in human–robot relations. *International journal of human-computer studies*, 71(3), 250-260. <https://doi.org/10.1016/j.ijhcs.2012.09.005>
- Litvack-Miller, W., McDougall, D., & Romney, D. M. (1997). The structure of empathy during middle childhood and its relationship to prosocial behavior. *Genetic, Social, and General Psychology Monographs*, 123(3), 303-325.
- MacPherson, S. E., Phillips, L. H., & Della Sala, S. (2002). Age, executive function and social decision making: a dorsolateral prefrontal theory of cognitive aging. *Psychology and aging*, 17(4), 598-609. DOI: 10.1037//0882-7974.17.4.598
- Madea, A. R. (1987). Empathy and nursing care outcomes. *Research and Theory for Nursing Practice*, 1(3), 197-213.
- Maldonado, H., Lee, J. E. R., Brave, S., Nass, C., Nakajima, H., Yamada, R., ... & Morishima, Y. (2005, May). We learn better together: enhancing elearning with emotional characters. In *Proceedings of th 2005 conference on Computer support for collaborative learning: learning 2005: the next 10 years!* (pp. 408-417). International Society of the Learning Sciences.
- Marsella, S. C., & Pynadath, D. V. (2005). Modeling influence and theory of mind. *Virtual Social Agents*, 199.

- Mascarenhas, S., Martinho, C., Prada, R., Paiva, A., & Leite, I. (2013). The influence of empathy in human – robot relations, *71*, 250–260. <https://doi.org/10.1016/j.ijhcs.2012.09.005>
- Matsumoto, D., Yoo, S. H., Fontaine, J., Anguas-Wong, A. M., Arriola, M., Ataca, B., ... Grossi, E. (2008). Mapping expressive differences around the world: The relationship between emotional display rules and individualism versus collectivism. *Journal of Cross-Cultural Psychology*, *39*(1), 55–74. <https://doi.org/10.1177/0022022107311854>
- Maylor, E. A., Moulson, J. M., Muncer, A. M., & Taylor, L. A. (2002). Does performance on theory of mind tasks decline in old age? *British Journal of Psychology*, *93*(4), 465–485. <https://doi.org/10.1348/000712602761381358>
- McCrae, R. R., & Terracciano, A. (2005). Personality profiles of cultures: aggregate personality traits. *Journal of personality and social psychology*, *89*(3), 407–425. <https://doi.org/10.1037/0022-3514.89.3.407>
- Mc Grath, M. (2009). *Assessing the effect of culture on theory of mind performance in South African students*. Unpublished honours thesis. University of Cape Town, Rondebosch, South Africa.
- McInnis, M. A. (2014). *The relation between theory of mind and empathy in preschool: The case of fantasy orientation*. Doctoral dissertation. University of Alabama Libraries.
- McKinnon, M. C., & Moscovitch, M. (2007). Domain-general contributions to social reasoning: Theory of mind and deontic reasoning re-explored. *Cognition*, *102*(2), 179–218. <https://doi.org/10.1016/j.cognition.2005.12.011>
- Mesquita, B., & Leu, J. (2007). The cultural psychology of emotion. *Handbook of cultural psychology*, 734–759.
- Miller, P. A., & Eisenberg, N. (1988). The Relation of Empathy to Aggressive and Externalizing/Antisocial Behavior. *Psychological Bulletin*, *103*(3), 324–344. <https://doi.org/10.1037/0033-2909.103.3.324>
- Moon, Y. (2000). Intimate Exchanges: Using Computers to Elicit Self-Disclosure From Consumers. *Journal of Consumer Research*, *26*(4), 323–339. <https://doi.org/10.1086/209566>
- Moran, J. M. (2013). Lifespan development: The effects of typical aging on theory of mind. *Behavioural brain research*, *237*, 32–40. <https://doi.org/10.1016/j.bbr.2012.09.020>
- Nass, C., Foehr, U., Brave, S., & Somoza, M. (2001). The effects of emotion of voice in synthesized and recorded speech. In *Proceedings of the AAAI symposium emotional and intelligent II: The tangled knot of social cognition*. North Falmouth, MA: AAAI.
- Nass, C., Isbister, K., & Lee, E. J. (2000). Truth is beauty: Researching embodied conversational agents. *Embodied conversational agents*, 374–402. Cambridge, MA: MIT Press.
- Nass, C., & 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., & 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., 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., 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>
- Newhagen, J. E., & Reeves, B. (1992). The Evening's Bad News: Effects of Compelling Negative Television News Images on Memory. *Journal of Communication*, 42(2), 25–41. <https://doi.org/10.1111/j.1460-2466.1992.tb00776.x>
- Nisbett, R. E. (2007). Eastern and Western ways of perceiving the world. In Y. Shoda, D. Cervone, & G. Downey (Eds.), *Persons in context: Building a science of the individual*, 62–83. New York, NY: Guilford Press.
- Nisbett, R. E., & Masuda, T. (2003). Culture and point of view. *Proceedings of the National Academy of Sciences*, 100(19), 11163–11170. <https://doi.org/10.1073/pnas.1934527100>
- Oliner, P. M., & Oliner, S. P. (1995). *Toward a caring society: Ideas into action*. Westport, CT: Praeger.
- Paiva, A., Leite, I., Boukricha, H., & Wachsmuth, I. (2017). Empathy in virtual agents and robots: A survey. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 7(3), 1–40. <https://doi.acm.org/10.1145/2912150>
- Perner, J. (1996). Simulation as explicitation of predication-implicit knowledge about the mind: arguments for a simulation-theory mix. In P. Carruthers & P.K. Smith (Eds.), *Theories of theories of mind* (pp. 90–104). Cambridge, UK: Cambridge University Press.
- Peters, C. (2006, September). Designing synthetic memory systems for supporting autonomous embodied agent behaviour. In *ROMAN 2006-The 15th IEEE International Symposium on Robot and Human Interactive Communication* (pp. 14–19). IEEE.
- Peterson, R. & Leonhardt, J. M. (2015). The complementary effects of empathy and nonverbal communication training on persuasion capabilities. *Administrative Issues Journal*, 5(1), 77–88. <https://dx.doi.org/10.2139/ssrn.2726150>
- Phillips, L. H., Bull, R., Allen, R., Inch, P., Burr, K., & Ogg, W. (2011). Lifespan aging and belief reasoning: Influences of executive function and social cue decoding. *Cognition*, 120(2), 236–247. <https://doi.org/10.1016/j.cognition.2011.05.003>
- Phillips, L. H., MacLean, R. D., & Allen, R. (2002). Age and the understanding of emotions: Neuropsychological and sociocognitive perspectives. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 57(6), 526–530. <https://doi.org/10.1093/geronb/57.6.P526>

- Plutchik, R. (1987). Evolutionary bases of empathy. *Empathy and its development*, 1, 38-46.
- Pratt, J. A., Hauser, K., Ugray, Z., & Patterson, O. (2007). Looking at human–computer interface design: Effects of ethnicity in computer agents. *Interacting with Computers*, 19(4), 512-523. <https://doi.org/10.1016/j.intcom.2007.02.003>
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and brain sciences*, 1(4), 515-526. <https://doi.org/10.1017/S0140525X00076512>
- Prendinger, H., & Ishizuka, M. (2005). The empathic companion: A character-based interface that addresses users' affective states. *Applied Artificial Intelligence*, 19(3–4), 267–285. <https://doi.org/10.1080/08839510590910174>
- Prendinger, H., Ma, C., Mori, J., & Ishizuka, M. (2005a, May). Does non-verbal behavior of an embodied agent matter? In *Proceedings of the 2005 International Conference on Active Media Technology, 2005.(AMT 2005)*. (pp. 535-540). IEEE.
- Prendinger, H., Mori, J., & Ishizuka, M. (2005b). Using human physiology to evaluate subtle expressivity of a virtual quizmaster in a mathematical game. *International Journal of Human Computer Studies*, 62(2), 231–245. <https://doi.org/10.1016/j.ijhcs.2004.11.009>
- Reeves, B., & Nass, C. I. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. Cambridge, UK: Cambridge University Press.
- Remland, M. S., Jones, T. S., & Brinkman, H. (1995). Interpersonal distance, body orientation, and touch: Effects of culture, gender, and age. *The Journal of social psychology*, 135(3), 281-297. <https://doi.org/10.1080/00224545.1995.9713958>
- Reniers, R. L., Corcoran, R., Drake, R., Shryane, N. M., & Völlm, B. A. (2011). The QCAE: A questionnaire of cognitive and affective empathy. *Journal of personality assessment*, 93(1), 84-95.
- Rickenberg, R., & Reeves, B. (2000). The effects of animated characters on anxiety, task performance, and evaluations of user interfaces. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '00*, 2(1), 49–56. <https://doi.org/10.1145/332040.332406>
- Riek, L. D., Mavridis, N., Antali, S., Darmaki, N., Ahmed, Z., Al-Neyadi, M., & Alketheri, A. (2010, April). Ibn sina steps out: Exploring arabic attitudes toward humanoid robots. In *Proceedings of the 2nd international symposium on new frontiers in human–robot interaction, AISB, Leicester* (Vol. 1).
- Rieser, V., & Lemon, O. (2011). Learning and evaluation of dialogue strategies for new applications: Empirical methods for optimization from small data sets. *Computational Linguistics*, 37(1), 153–196. https://doi.org/10.1162/coli_a_00038
- Riess, H., & Kraft-Todd, G. (2014). EMPATHY: a tool to enhance nonverbal communication between clinicians and their patients. *Academic Medicine*, 89(8), 1108-1112. DOI: 10.1097/ACM.0000000000000287
- Rossen, B., Johnsen, K., Deladisma, A., Lind, S., & Lok, B. (2008, September). Virtual humans

- elicit skin-tone bias consistent with real-world skin-tone biases. In *International Workshop on Intelligent Virtual Agents* (pp. 237-244). Berlin, Heidelberg: Springer.
- Rousseau, P. C., & Blackburn, G. (2008). The touch of empathy. *Journal of Palliative Medicine*, *11*(10), 1299–1300. <https://doi.org/10.1089/jpm.2008.0174>
- Ruffman, T., Henry, J. D., Livingstone, V., & Phillips, L. H. (2008). A meta-analytic review of emotion recognition and aging: Implications for neuropsychological models of aging. *Neuroscience & Biobehavioral Reviews*, *32*(4), 863-881. <https://doi.org/10.1016/j.neubiorev.2008.01.001>
- Ruffman, T., Slade, L., Devitt, K., & Crowe, E. (2006). What mothers say and what they do: The relation between parenting, theory of mind, language and conflict/cooperation. *British Journal of Developmental Psychology*, *24*(1), 105-124. <https://doi.org/10.1348/026151005X82848>
- Ryder, A. G., Yang, J., Zhu, X., Yao, S., Yi, J., Heine, S. J., & Bagby, R. M. (2008). The cultural shaping of depression: somatic symptoms in China, psychological symptoms in North America? *Journal of abnormal psychology*, *117*(2), 300-313. <https://doi.org/10.1037/0021-843X.117.2.300>
- Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L. J., & Lee, K. (2006). The development of executive functioning and theory of mind: A comparison of Chinese and US preschoolers. *Psychological science*, *17*(1), 74-81. <https://doi.org/10.1111%2Fj.1467-9280.2005.01667.x>
- Saltzman, J., Strauss, E., Hunter, M., & Archibald, S. (2000). Theory of mind and executive functions in normal human aging and Parkinson's disease. *Journal of the International Neuropsychological Society*, *6*(7), 781-788. <https://doi.org/10.1017/S1355617700677056>
- Senju, A., Vermetti, A., Kikuchi, Y., Akechi, H., Hasegawa, T., & Johnson, M. H. (2013). Cultural background modulates how we look at other persons' gaze. *International journal of behavioral development*, *37*(2), 131-136. <https://doi.org/10.1177%2F0165025412465360>
- Shahaeian, A., Nielsen, M., Peterson, C. C., & Slaughter, V. (2014). Cultural and family influences on children's theory of mind development: A comparison of Australian and Iranian school-age children. *Journal of Cross-Cultural Psychology*, *45*(4), 555-568. <https://doi.org/10.1177%2F0022022113513921>
- Shahaeian, A., Peterson, C. C., Slaughter, V., & Wellman, H. M. (2011). Culture and the sequence of steps in theory of mind development. *Developmental psychology*, *47*(5), 1239-1247. <https://doi.org/10.1037/a0023899>
- Shin-Young, P., & Harada, A. (2006). A study of Non-verbal Expressions in a computer-Mediated communication context (CMC).
- Shneiderman, B., Plaisant, C., Cohen, M., Jacobs, S., Elmqvist, N., & Diakopoulos, N. (2016). *Designing the user interface: strategies for effective human-computer interaction*. Harlow, UK: Pearson Education Ltd.
- Slessor, G., Phillips, L. H., & Bull, R. (2007). Exploring the specificity of age-related differences

in theory of mind tasks. *Psychology and aging*, 22(3), 639-643.
<https://doi.org/10.1037/0882-7974.22.3.639>

- Straßmann, C., von der Pütten, A. R., Yaghoubzadeh, R., Kaminski, R., & Krämer, N. (2016, September). The effect of an intelligent virtual agent's nonverbal behavior with regard to dominance and cooperativity. In *International Conference on Intelligent Virtual Agents* (pp. 15-28). Springer, Cham.
- Sullivan, S., & Ruffman, T. (2004). Social understanding: How does it fare with advancing years? *British Journal of Psychology*, 95(1), 1-18.
<https://doi.org/10.1348/000712604322779424>
- Svetieva, E., & Frank, M. G. (2016). Empathy, emotion dysregulation, and enhanced microexpression recognition ability. *Motivation and Emotion*, 40(2), 309-320.
- Takaku, S., Weiner, B., & Ohbuchi, K. I. (2001). A cross-cultural examination of the effects of apology and perspective taking on forgiveness. *Journal of Language and Social Psychology*, 20(1-2), 144-166.
- Tepper, D. T., & Haase, R. F. (1978). Verbal and nonverbal communication of facilitative conditions. *Journal of counseling psychology*, 25(1), 35.
- Tooby, J., & Cosmides, L. (1995). 7 8 Mapping the Evolved Functional Organization of Mind and Brain.
- Triandis, H. C. (1993). Collectivism and individualism as cultural syndromes. *Cross-cultural research*, 27(3-4), 155-180. <https://doi.org/10.1177%2F106939719302700301>
- Vang, M. H., & Fox, J. (2014). Race in virtual environments: Competitive versus cooperative games with black or white avatars. *Cyberpsychology, Behavior, and Social Networking*, 17(4), 235-240. <https://doi.org/10.1089/cyber.2013.0289>
- Vannini, N., Enz, S., Sapouna, M., Wolke, D., Watson, S., Woods, S., ... & Aylett, R. (2011). "FearNot!": a computer-based anti-bullying-programme designed to foster peer intervention. *European journal of psychology of education*, 26(1), 21-44.
<https://doi.org/10.1007/s10212-010-0035-4>
- Vinden, P. G. (2001). Parenting attitudes and children's understanding of mind: a comparison of Korean American and Anglo-American families. *Cognitive Development*, 16(3), 793-809.
[https://doi.org/10.1016/S0885-2014\(01\)00059-4](https://doi.org/10.1016/S0885-2014(01)00059-4)
- Völlm, B. A., Taylor, A. N., Richardson, P., Corcoran, R., Stirling, J., McKie, S., ... & Elliott, R. (2006). Neuronal correlates of theory of mind and empathy: a functional magnetic resonance imaging study in a nonverbal task. *Neuroimage*, 29(1), 90-98.
<https://doi.org/10.1016/j.neuroimage.2005.07.022>
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology*, 54(6), 1063.
- Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development:

- The truth about false belief. *Child development*, 72(3), 655-684.
<https://doi.org/10.1111/1467-8624.00304>
- Westerhof, G. J., & Barrett, A. E. (2005). Age identity and subjective well-being: A comparison of the United States and Germany. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 60(3), 129-136. <https://doi.org/10.1093/geronb/60.3.S129>
- Wilkowska, W., & Ziefle, M. (2009, November). Which factors form older adults' acceptance of mobile information and communication technologies?. In *Symposium of the Austrian HCI and Usability Engineering Group* (pp. 81-101). Springer, Berlin, Heidelberg.
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1), 103-128.
[https://doi.org/10.1016/0010-0277\(83\)90004-5](https://doi.org/10.1016/0010-0277(83)90004-5)
- Wooldridge, M. J., & Jennings, N. R. (1995). Intelligent Agents: Theory and Practice. *Journal of Knowledge Engineering Review*, 10(2), 115-152.
<https://doi.org/10.1017/S0269888900008122>
- Yaghoubzadeh, R., Kramer, M., Pitsch, K., & Kopp, S. (2013). Virtual Agents as Daily Assistants for Elderly or Cognitively Impaired People. *Proceedings of the 13th International Conference on Intelligent Virtual Agents*, 8108(Iva), 91. https://doi.org/10.1007/978-3-642-40415-3_7
- Yalçın, Ö. N. (2018, October). Modeling Empathy in Embodied Conversational Agents. In *Proceedings of the 2018 on International Conference on Multimodal Interaction* (pp. 546-550). ACM.
- Zajicek, M., & Hall, S. (2000). Solutions for elderly visually impaired people using the Internet. In *People and Computers XIV—Usability or Else!* (pp. 299-307). Springer, London.
- Zanbaka, C., Goolkasian, P., & Hodges, L. (2006, April). Can a virtual cat persuade you?: the role of gender and realism in speaker persuasiveness. In *Proceedings of the SIGCHI conference on Human Factors in computing systems* (pp. 1153-1162). ACM
- Zhou, J., Gotch, C., Zhou, Y., & Liu, Z. (2008). Perceiving an object in its context—is the context cultural or perceptual?. *Journal of Vision*, 8(12), 2-2.

Appendix

Please contact the author if you are interested to obtain the stimulus materials and questionnaires that were used in Study 1– Study 3.