Using Avatars to Study Social Cognition in Cross-cultural Psychology and High-functioning Autism

Inauguraldissertation

zur Erlangung des Doktorgrades

der Humanwissenschaftlichen Fakultät

der Universität zu Köln

nach der Promotionsordnung vom 10.05.2010

vorgelegt von

Thomas Dratsch

aus

Wuppertal

November 2013

Diese Dissertation wurde von der Humanwissenschaftlichen Fakultät der Universität zu Köln im Dezember 2013 angenommen.

1. Gutachter: Prof. Dr. Gary Bente
Universität zu Köln
Department Psychologie
Richard-Strauss-Straße 2
50931, Köln

2. Gutachter: Prof. Dr. Dr. Kai Vogeley
Universitätsklinik Köln
Zentrum für Neurologie und Psychiatrie
Kerpener Straße 62
50924, Köln

Acknowledgements

First of all, I would like to thank my supervisors Prof. Gary Bente and Prof. Kai Vogeley for their great support and encouragement at all stages of this thesis. I would like to thank Prof. Bente for hiring me right after graduation, for giving me the opportunity to pursue my PhD, and for introducing me to science. I would especially like to thank Prof. Vogeley for trusting me to conduct two studies with his patients at the Autism Ambulance of the University Clinic of Cologne.

I am very grateful for all the amazing colleagues/friends that I have had the pleasure to meet through my work at our lab. Each one of you is special and awesome, so I list everyone in alphabetical order: Daniel—the Baguette—Brunsch, thank you for always taking time out of your busy schedule for hour-long conversations on the hallway; Freya—History Genius—Elvert, thank you for dropping little nuggets of wisdom during lunch breaks; Lena—Master of Corrections—Frischlich, thank you for catching the most embarrassing errors in this thesis and for staying calm when everybody panics; Björn—MoCap—Günter, thank you for wearing pyjamas to work with me and pretending they were MoCap-suits—nobody noticed; Julia—Grunge—Kneer, thank you for awesome conversations at after-parties; Diana—the أستاذ—Rieger, thank you for also being afraid of flying and for being the best dissertation buddy I could wish for; Daniel—the DJ—Roth, thank you for playing old-school HipHop in our office when we both need it the most; Roman—Dances with Wolves—Rybak, thank you for launching our piano and voice duo; Sutude—the German—Shishechian, thank you for speaking the clearest Deutsch at the institute and for always being kind; Oana—Born to Run—Teodosiu, thank you for running the Unilauf with us; Monika—Soul of the Institute—Wegener, thank you for moving all bureaucratic barriers out of the way; Tim—Mario Kart Grandmaster—Wulf, thank you for being a great HiWi and for bringing three avocados on lunch breaks.

In addition, I would like to thank all members of the Autism Ambulance of the University Clinic of Cologne for giving me access to their patients and for helping me with the data collection. The studies in this thesis would not have been possible without their help and support. I would like to thank Prof. Kai Vogeley, Katharina Krämer, Alexandra Georgescu, Tabea von der Lühe, Nenia Schulze, Astrid Gawronski, Kathleen Volpert, Uli Pfeiffer, Bojana Kuzmanovic, Ralf Tepest, and Leonhard Schilbach.

I would also like to thank Kliment Yanev and Iva Barišić for programming the experimental paradigms for Studies 3 and 4.

I wish to thank my parents for supporting me spiritually during the time I was working on this thesis. Thanks to Hans, Gabi, Monika, Robert.

Last, but not least, a cordial, anonymous thank you goes out to all the men and women who were brave enough to participate in my studies. I would mention you all by name here, but I would probably get in trouble for that.

Table of Contents

List of Tables

List of Figures

Introduction

Over the last 20 years, virtual avatars have become a popular research tool in psychology and neuroscience for studying social cognition. As opposed to photographs or movie recordings of actual human beings, avatars allow for the precise control over all aspects of the stimulus, ranging from the avatar's gaze and movement behavior to its physical appearance, such as age, gender, or ethnicity (Vogeley & Bente, 2010). Additionally, avatars have made it possible to create interactive paradigms that enable the study of social interactions in real-time (Wilms et al., 2010). Benefitting from this recent development, the present thesis set out to use avatars to study social cognition in two areas: cross-cultural psychology and high-functioning autism. The unifying connection between the four studies combined in this thesis is that they all rely on specific advantages offered by avatars and that they could not have been conducted without them.

The basic structure of the thesis is as follows: **Study 1** lays the methodological ground for the other three studies (**Study 2**, **Study 3,** and **Study 4**) by testing the validity of avatars in a trust game. Once the validity of avatars has been established, they are used to tackle particular demands in the study of social cognition in cross-cultural psychology and high-functioning autism.

The goal of **Study 1** was to test whether virtual avatars can produce similar behavioral responses as photographs of actual human beings. Even though there are many studies that show that avatars can lead to similar impressions as photographs or videos (e.g., Moser et al., 2007), only few studies so far have tested if avatars also lead to similar behavioral responses. As two studies have shown, the differences between avatars and real human beings can be subtle. In a study on the perception of nonverbal behavior, Bente et al. (2001) investigated whether observers form similar impressions based on animated avatars as opposed to video recordings of nonverbal behavior. The results showed that impression ratings did not differ between avatars and video recordings. However, eye-tracking results revealed that participants focused less attention on the faces of the avatars as opposed to the faces of the humans in the videos. In a study on interpersonal trust, Riedl et al. (2011) could show that participants trusted virtual avatars in a similar way as other humans. However, fMRI results revealed a reduced activity in the mentalizing network when participants interacted with an avatar as opposed to a human.

Based on these two findings, **Study 1** set out to devise a more rigorous test for the validity of avatars by assessing participants' behavioral trust in avatars. First, by letting participants play a trust game for actual money, it was tested whether impressions of avatars translate into behavioral trust when actual money is at stake. Second, by combining avatars with another objective information source, it was tested whether avatars still have an effect on decision making if other more objective information is available. The results of **Study 1** show that humans trust avatars in a similar way as they trust other human beings, supporting the validity of avatars in the study of psychological phenomena. The results, however, also hint at subtle differences between avatars and humans. When combined with another source of information, avatars have a slightly weaker effect than photographs of actual humans.

As the validity of avatars has been established in **Study 1,** they are used to tackle particular demands in the study of social cognition in cross-cultural psychology and high-functioning autism (**Study 2, Study 3, and Study 4).** Taken together, all three studies (**Study 2** to **Study 4**) make use of the unique features that avatars offer researchers.

Study 2 investigated the influence of two sources of information—factual information and tacit, social cues—on decision making in an intercultural trust game with participants form an individualistic (Germany) and a collectivistic culture (United Arab Emirates). For **Study 2**, a database of Arab and German avatars was developed that controlled for all facial features except the avatar's ethnicity. Therefore, the results of the final study cannot be attributed to confounding factors, such as facial expressions.

In **Study 3**, two features of avatars were used: precise control over the avatar's movement behavior and interactive control of the avatar through the participants. To study the ability to detect direct gaze in high-functioning autism, a gaze sequence consisting of direct gaze and slightly averted gaze was programmed and displayed by a virtual avatar, taking into consideration even small degrees of deviation. Such a stimulus sequence would not have been possible to create using conventional techniques, such as video recordings or photographs. As the results show, the methodology was successful at uncovering subtle impairments in the detection of direct gaze by individuals with autism that would have otherwise been left undetected. To further explore whether the impaired ability to perceive direct gaze in high-functioning autism could be explained by an aversive reaction towards uncontrollable gaze, a virtual avatar was created whose gaze could be controlled by the participants. Correspondingly, such an interactive paradigm could not have been realized using conventional means. As the results show, the methodology was successful at showing that individuals with high-functioning autism can overcome part of their impairment in detecting direct gaze when given active control over the social gaze of the avatar.

For **Study 4**, a database of attractive and unattractive avatars was developed. Because previous studies with heterogeneous facial stimuli had shown conflicting results with regard to the ability to detect attractiveness by individuals with autism (Da Fonseca, Santos, Rosset, & Deruelle, 2011; White, Hill, Winston, & Frith, 2006), it was crucial to develop a homogenous stimulus set free of possible confounding factors. In addition, because the study also used eye-tracking to record participants visual attention to different facial areas, facial stimuli had to be free of artifacts that could distract participants gaze. The final stimulus set consisted of faces that only differed with regard to attractiveness and kept all other influence factors constant—facial expressions of all avatars were neutral and the gaze direct. The results show that individuals with high-functioning autism are impaired in their ability to detect attractiveness.

Goal of the Present Thesis

As a whole, the aim of the present thesis is to advance knowledge in two major fields of social cognition—cross-cultural psychology and high-functioning autism—by using virtual avatars. **Study 1** assessed the validity of social avatars as a research tool. **Study 2** focused on cross-cultural differences in trust. **Study 3** and **Study 4** investigated whether two core abilities of social cognition—detection of direct gaze and perception of attractiveness—are impaired in individuals with high-functioning autism.

Study 1: Do You Trust My Avatar?

Introduction

Despite their widespread use in science and on the web, only few studies so far have compared the effects of virtual avatars to humans (e.g., Moser et al., 2007). In a study focusing on the perception of nonverbal behavior, Bente et al. (2001) could show that virtual avatars displaying nonverbal behavior produced the same impressions as the original videos that the movements of the avatars had been transcribed from. However, results of eye-tracking revealed that participants focused less attention on the faces of the avatars and more on the body as compared to the original videos. In a different study, Riedl et al. (2011) could show that participants trusted virtual avatars in a similar way as other humans. However, fMRI results revealed a reduced activity in the mentalizing network when participants interacted with an avatar as opposed to a human. Taken together, both studies show that virtual avatars can have similar effects as real humans. Nevertheless, they also hint at subtle differences in the response to avatars that should be further investigated. The goal of the present study was to further explore whether virtual avatars are trusted in a similar way as humans.

A recent study by Bente et al. (2012) provides a framework to study trust in online transactions. Using a standard trust-game (Bolton, Katok, & Ockenfels, 2004), Bente et al. (2012) could show that both reputation scores and seller photographs—two information sources often found on e-commerce websites—had an independent influence on buyers' trust and purchase decisions. The aim of the present study was to investigate whether the results of the study by Bente et al. (2012) could be replicated with virtual avatars. Additionally, the present study also investigated whether participants' involvement in the trust game would influence the effect of seller avatars on purchase decisions.

Experiment 1

Background and Hypotheses

Reputation. One key problem that every website offering an open market place for its users must solve is how to establish trust between sellers and buyers. Because both parties usually do not meet in person, everyday heuristics used to establish trust cannot be applied. To solve this problem, almost all of the major websites (e.g., eBay, Amazon) have adopted reputation systems—usually in the form of a five-star-index—that give users the ability to quickly assess the trustworthiness of their potential business partners. As a large body of studies has shown, reputation scores can effectively increase trust among sellers and buyers and thus boost the number of successful transactions (e.g., Diekmann, Jann, & Wyder, 2007; Gregg & Scott, 2006). In their study, Bente et al. (2012) did find a significant effect of reputation scores on purchase decisions. Positive reputation lead to higher purchase rates than both negative reputation and missing reputation. Additionally, negative reputation led to significantly higher purchase rates than missing information, indicating that the uncertainty of missing information may be more aversive than the information conveyed by the negative reputation score (Bente et al., 2012). In line with the results by Bente et al. (2012), the following hypotheses are formulated:

Hypothesis 1: Reputation scores influence purchase decisions in online transactions. Positive reputation leads to higher purchase rates than both negative and missing reputation (H1a.) Negative reputation leads to higher purchase rates than missing reputation (H1b).

Seller avatars. In addition to reputation scores, seller representations, such as photographs, are another source of information that can help to increase trust. For instance, Jones et al. (1998) could show that photos of sellers could positively influence buyers' trust and purchase decisions. However, seller photographs can also activate stereotypes and have a negative effect on trust building if they depict members of minorities (Doleac & Stein, 2010). For instance, Ayres et al. (2011) found that eBay products accompanied by photographs of African Americans sold for 20% less than products accompanied by Caucasians. Taken together, these results show that the effects of photographs in e-commerce—both positive and negative—are similar to real-life interactions. Bente et al. (2012) found that trustworthy photos led to higher purchase rates than untrustworthy photos. Contrary to the results on the effect of missing reputation, Bente et al. (2012) did not find a difference in purchase rates between untrustworthy photos and missing photos. Regarding the effect of avatars on purchase decisions, the following hypotheses are stated based on the results by Bente et al. (2012):

> *Hypothesis 2:* Seller avatars influence purchase behavior in online transactions. Trustworthy seller avatars lead to higher purchase rates than both untrustworthy and missing avatars (H2a). Untrustworthy and missing avatars do not differ in purchase rates. (H2b).

Involvement. According to the Elaboration Likelihood Model (Petty & Cacioppo, 1986), decisions can be influenced by two different pathways or routes of information processing: the peripheral and the central route. When humans use the peripheral route, they are easily influenced by superficial cues, such as the attractiveness or supposed expertise of the information source. When humans use the central route, they process the available information in great detail and try to make rational decisions. Humans generally use the central route, when they are highly involved in the decision to be made and cognitive resources are available. When the involvement in the decision is low and cognitive resources are scarce, participants use the peripheral route. From a rational perspective, the best strategy in the trust game would be to solely rely on the reputation scores and to ignore the seller photographs because the reputation scores as opposed to the seller photographs represent an objective measure of the sellers' cooperativeness. In their original study, Bente et al. (2012) found equal effect sizes for reputation ($\eta^2_p = .27$) and seller photographs ($\eta^2_p = .20$), indicating that participants were equally influenced in their decisions by tacit, social cues (seller photos) and factual information (reputation scores). One possible explanation for this may be that participants' involvement in the trust game was low, which caused them to use a shallow mode of processing, thus giving both cues equal weight in their decisions. One goal of the present study was to investigate whether participants' could be influenced to use the central route of information processing and thus decrease the influence of seller avatars on purchase decisions. In order to vary the involvement of the participants in this study, the present study varied the amount of virtual units at stake in the trust game. In the high involvement condition (high amount at stake), it was expected that participants would use the central route of processing based on factual information (reputation scores). In the low involvement condition (low amount at stake), it was expected that participants would use the peripheral route based on tacit cues (seller avatars). Therefore, it was hypothesized:

Hypothesis 3: Involvement and seller avatar show an interaction effect. High involvement leads to a weaker influence of seller avatars on purchase decisions (H3).

Methods

Participants. A total of 126 students (66 male, 60 female; M_{age} = 24.15, SD_{age} = 5.66) participated in the online trust game. Participants were invited via email to take part in an online trust game in which they could earn between 5 and 7.25 Euros. Emails were distributed through mailing lists of the University of Cologne.

Design. The influence of avatars, reputation scores, and involvement on purchase decisions was analyzed in a 3 × 3 × 2 design, with avatar (trustworthy vs. untrustworthy vs. no avatar) and reputation (positive vs. negative vs. no reputation) as within-subjects factors and involvement (high vs. low) as a between-subjects factor.

Stimulus materials.

***Trust game*.** The standard trust game by Bolton (2004) provided the framework for the present study. In this trust game, a buyer (trustor) has to decide whether he will trust a seller (trustee) and buy a product from him/her. However, instead of actual products buyer and seller only exchange virtual units that are converted into real money at the end of the experiment. All possible decisions and outcomes of the trust game are shown in the pay-off matrix (see Figure 1). If the buyer decides not to buy (Case 1), both buyer and seller keep their 35 units. If the buyer decides to buy and the seller ships the product (Case 2a), both buyer and seller receive 50 units for the successful trade. If the buyer decides to buy and the seller does not ship the product (Case 2b), the buyer loses his/her 35 units to the seller, who receives 70 units.

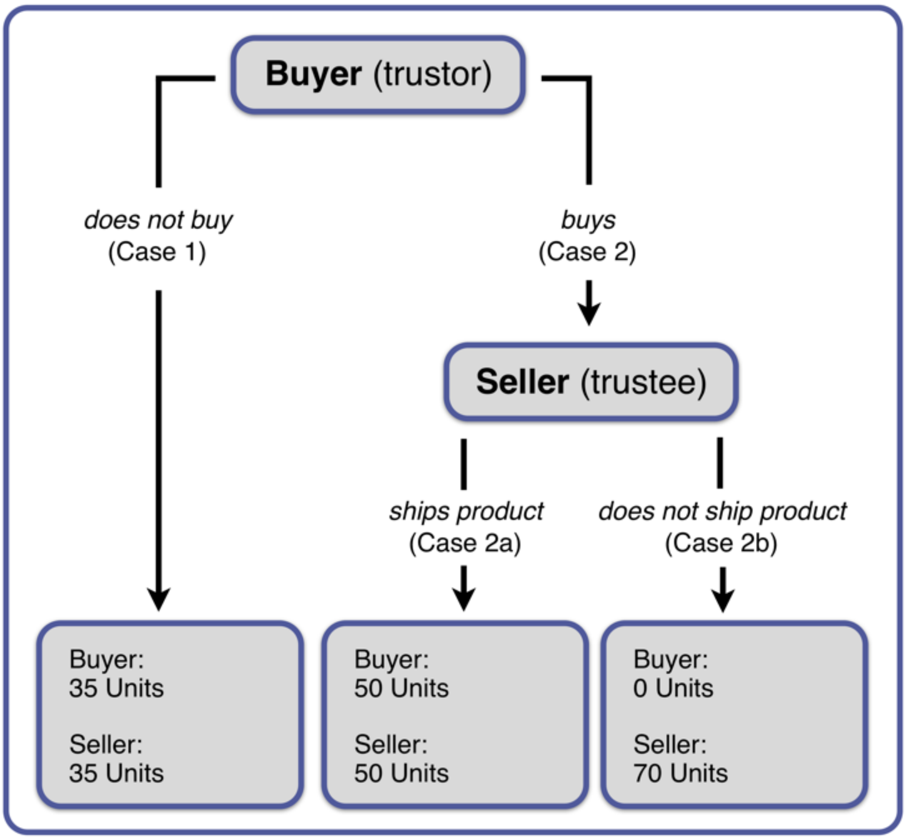

Figure 1. Pay-off matrix for the trust game showing all possible decisions and outcomes.

Seller avatars. To select appropriate avatars for the study, 30 virtual avatars (15 male, 15 female) were created using the software FaceGen (Singular Inversions, 2011). In a prestudy, 35 participants (8 male, 27 female; $M_{age} = 27.00$, $SD_{age} = 6.76$) rated the trustworthiness of the avatars on a 7-point scale ranging from 1 (*very untrustworthy*) to 7 (*very trustworthy*). Based on the trustratings in the prestudy, six trustworthy and six untrustworthy avatars were selected for the study. On average, the six trustworthy avatars ($M = 4.59$, $SD = 0.82$) were rated to be significantly more trustworthy than the six untrustworthy avatars ($M = 2.71$, $SD = 0.86$), $t(34) = 8.69$, $p < .001$, $d = 1.48$. The remaining avatars were used in filler trials.

Reputation scores. The reputation of the sellers was symbolized using a reputation index with five stars—with the number of stars representing the percentage of previous trades in which the seller had shipped the product after receiving the payment (one star: 0–20%; two stars: 21–40%; three stars: 41–60%; four stars: 61–80%; five stars: 81–100%). Appropriate reputation scores were selected based on the prestudy by Bente et al. (2012): 30 participants (16 male, 14 female) rated the trustworthiness of the different levels of the reputation score on a 7-point scale ranging from 1 (*very untrustworthy*) to 7 (*very trustworthy*). Based on the results of the prestudy, two of the star-indexes were selected for the study: Three stars (the seller shipped the product 41–60% of the time) led to a significantly lower trust rating (M = 2.90, SD = 1.30) than four stars (the seller shipped the product 61–80% the time), $t(29) = 8.95$, $p < .001$, $d = 1.70$.

Involvement. Participants' involvement in the trust game was modified through the amount of units at stake in the game. In the low involvement condition, both buyer and seller received a basic amount of 35 units. In the high involvement condition, the amount was raised to 3500 units. Participants were randomly assigned to the high or low involvement condition.

Procedure. All participants completed the study online using their own computers. At the beginning of the study, participants were informed that they would now play an online trust game with buyers and sellers, and that they their own role in the trust game would be determined randomly. On the next page, participants were informed that they had been randomly assigned the role of the buyer—all participants were assigned the role of the buyer—and that they would now play the trust game with several other players who had been assigned the role of the seller. Participants were then informed that the sellers would be represented through virtual avatars generated out of real photographs. Two photographs and their avatars were provided as examples. Participants were further informed that in some cases the avatar would not be available. After that, participants were introduced to the five-star reputation score. Again, participants were told that sometimes the reputation score would not be available. Participants were then informed that they would not receive immediate feedback on the decisions of the sellers, but that their final earnings would be calculated based on the decisions of the sellers. After that, participants proceeded to the actual profiles of the sellers that included the sellers' avatar and reputation, and decided for each profile whether they would buy from that particular seller (see Figure 2 for a sample profile).

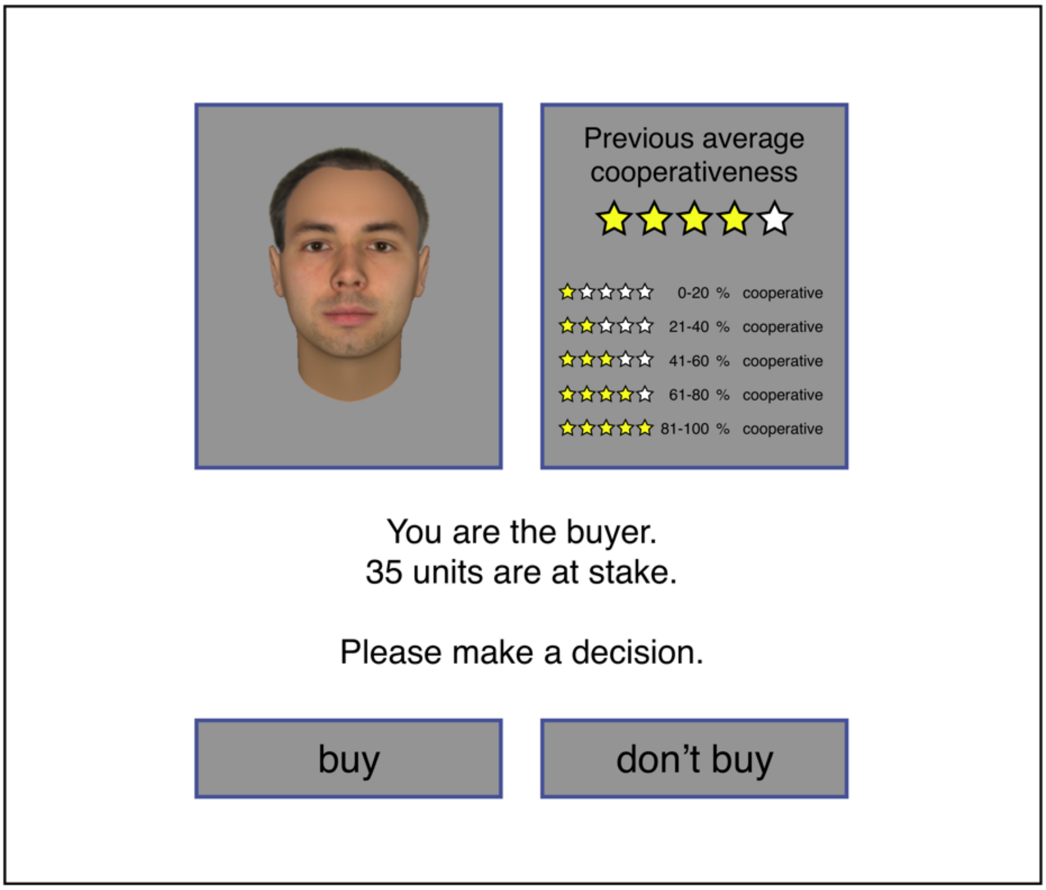

Figure 2. Screenshot of the profile of one seller in the trust game, showing a trustworthy avatar with a high reputation score (four stars) in the low involvement condition (35 units at stake).

Participants then completed 15 experimental trials and 14 filler trials (four filler trials in the beginning, eight filler trials randomly mixed with the experimental trials, and two filler trials at the end). Filler trials always used reputation scores other than three or four and were included to distract participants from the repeated presentation of reputation scores three and four. After participants had completed all trials, they were thanked for their participation and issued an anonymous code that allowed them to collect their earnings. Payment was calculated based on participants' decisions in the trust game. If participants had decided to buy the product, payment was always calculated as if the seller had shipped the product. Final payment to participants ranged from 5.07 to 6.95 Euros (M = 6.05, SD = 0.32).

Results and Discussion

Effect of involvement. Buying decisions were analyzed in a 3 × 3 × 2 mixed ANOVA (avatar × reputation × involvement). There was neither a significant main effect of involvement nor any significant interactions between involvement and the other variables, $F < 1$. Thus, hypothesis H3 could not be supported. The degree of involvement did not influence participants' buying decisions.

Effect of avatar and reputation. Because there was no significant main effect of involvement, data of both conditions were combined into one single group. Then, buying decisions were analyzed in a 3 × 3 repeated-measures ANOVA (avatar × reputation). There was a significant main effect of avatar, $F(2, 250) = 39.57$, $p < .001$, $\eta^2_p = .240$, a significant main effect of reputation, $F(2, 250) = 124.58$, $p < .001$, $\eta^2_p = .499$, and a significant interaction effect between avatar and reputation, $F(4, 500) = 3.66$, $p = .006$, $\eta^2_p = .028$. The results are shown in Figure 3.

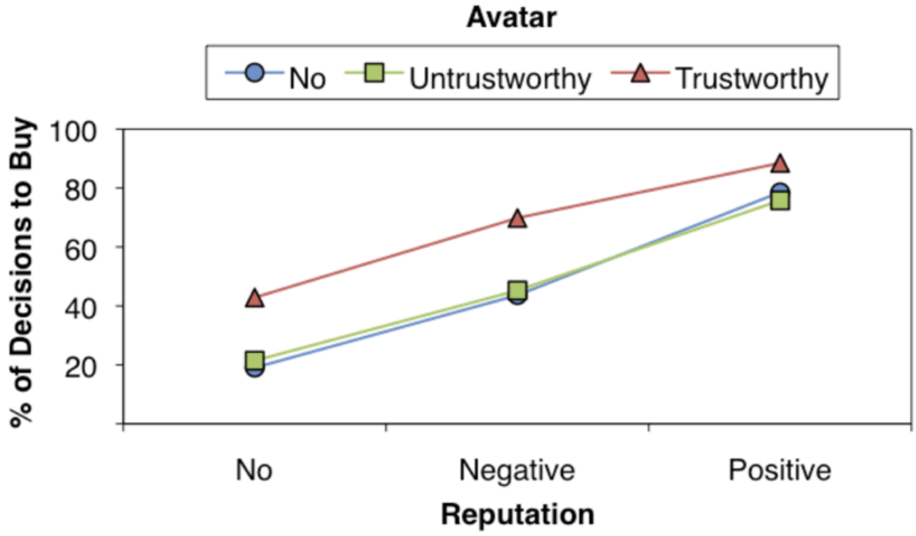

Figure 3. Effects of reputation and avatar on purchase decisions.

To break down these effects, pairwise comparisons were performed.

Reputation scores did significantly affect purchase behavior in online transactions. Positive reputation lead to higher purchase rates than negative and missing reputation, supporting hypothesis H1a. In addition, negative reputation lead to higher purchase rates than missing reputation, supporting hypothesis H1b (see Table 1).

Table 1

Results of Post-hoc Comparisons for Reputation

Avatar	Reputation	$t(125)$	p	Cohen's d
No	Negative vs. No	5.35	< .001	0.48
	Positive vs. No	12.39	< .001	1.10
	Positive vs. Negative	7.45	< .001	0.67
Untrustworthy	Negative vs. No	5.52	< .001	0.50
	Positive vs. No	13.51	< .001	1.21
	Positive vs. Negative	7.19	< .001	0.64
Trustworthy	Negative vs. No	6.15	< .001	0.55
	Positive vs. No	10.63	< .001	0.98
	Positive vs. Negative	5.25	< .001	0.49

Seller avatars also had a significant effect on purchase behavior in online transactions. Trustworthy seller avatars lead to higher purchase rates than untrustworthy or missing avatars, supporting hypothesis H2a. There was no difference in purchase rates between avatars with an untrustworthy appearing seller or missing avatars, supporting H2b (see Table 2).

Table 2

Results of Post-hoc Comparisons for Avatar

Reputation	Avatar	$t(125)$	p	Cohen's d
No	Untrustworthy vs. No	0.75	.455	0.07
	Trustworthy vs. No	5.71	< .001	0.51
	Trustworthy vs. Untrustworthy	5.77	< .001	0.52
Negative	Untrustworthy vs. No	0.37	.714	0.03
	Trustworthy vs. No	6.07	< .001	0.55
	Trustworthy vs. Untrustworthy	6.32	< .001	0.56
Positive	Untrustworthy vs. No	0.82	.415	0.07
	Trustworthy vs. No	2.75	.007	0.26
	Trustworthy vs. Untrustworthy	4.43	< .001	0.41

As for the combined effect of reputation and seller avatars, there was a significant interaction between reputation and seller avatars. In their original study, Bente et al. (2012) did not find a significant interaction between reputation and seller photographs, which they interpreted as evidence for the independent influence of both cues on purchase decisions. The results of the present study, however, hint at an important difference in the effect of seller photographs as opposed to seller avatars: In the study by Bente et al. (2012) the effect of the seller photographs was independent of the effect of reputation. In the present study, the effect of the seller avatars differed with the levels of reputation. In the negative reputation condition, where there was high uncertainty whether the seller would ship the product (only 41–60% of the time), participants were more susceptible to the influence of the seller avatars. In the positive reputation condition, where there was high certainty that the seller would ship the product (61–80% of the time), the influence of the seller avatars was weaker. This suggests that even though participants do use seller avatars in their decision process, they give less weight to this cue as opposed to reputation scores. This is also supported by the larger effect size for reputation (η^2_p = .499) as opposed to seller avatars (η^2_p = .240). However, if the reputation is negative, participants rely more strongly on seller avatars in their decisions. Because this interaction effect was not predicted, Experiment 2 was conducted to replicate the interaction between reputation and seller avatars.

Experiment 2

Background and Hypotheses

As the results of Experiment 1 have shown, there was a significant interaction effect between reputation and trustworthiness of the seller avatar. More precisely, the effect of avatar was stronger if there was high uncertainty that the seller would ship the product. If the uncertainty that the seller would ship the product was low, the effect of the trustworthiness of the avatar was weaker. However, Experiment 1 suffered from one major limitation: Reputation score number 3 was both more negative and represented a higher uncertainty than reputation score number 4 —thus confounding negativity and uncertainty. To rule out this possibility, Experiment 1 was replicated using all five reputation scores. Especially reputation scores 1 (seller shipped 0–20% of the time) and 2 (seller shipped 21–40% of the time) are important in this context because they are more negative than reputation score number 3 but represent less uncertainty—it is certain that the seller will not ship the product. Therefore, to replicate the effect and to rule out other explanations, Experiment 2 was performed. Based on the results of Experiment 1, I hypothesize:

> *Hypothesis 4:* The effect of seller avatars is strongest in the condition with the highest uncertainty that the seller will ship the product compared to all other conditions.

Methods

Participants. Overall, 147 students (21 male, 126 female; M_{age} = 23.88, SD_{age} = 4.56) participated in the online trust game. Participants were recruited through email invitations sent via mailing lists of several universities in Germany. Participants were invited to play an online trust game in which they could earn between 7 and 10 Euros.

Design. To analyze the interaction between reputation scores and trustworthiness of the seller avatars, I used a 2 (avatar: trustworthy vs. untrustworthy) × 5 (reputation: 1 vs. 2 vs. 3 vs. 4 vs. 5 stars) design, with both avatar and reputation as within-subjects factors.

Stimulus materials. Based on the pretest of Experiment 1, ten trustworthy and ten untrustworthy avatars were selected. On average, the ten trustworthy avatars (M = 4.35, SD = 0.62) were rated to be significantly more trustworthy than the ten untrustworthy avatars (M = 3.01, SD = 0.77), $t(34) = 8.45$, $p < .001$, $d = 1.43$. The remaining avatars were used in filler trials.

To vary the reputation of the sellers, the same five-star-index as in Experiment 1 was used.

Procedure. The procedure was the same as in Experiment 1. Payment ranged from 7.00 to 10.00 Euros (M = 8.52, SD = 0.39).

Results and Discussion

Effect of avatar and reputation. To analyze the effect of reputation and avatar on purchase decisions, I performed a 2 × 5 repeated-measures ANOVA (avatar × reputation). There was a significant main effect of avatar on purchase decisions, $F(1, 146) = 28.5$, $p < .001$, $\eta^2_p = .163$. There was a significant main effect of reputation on purchase decisions, $F(4, 584) = 517.6$, $p < .001$, $\eta^2_p = .780$. There was a significant interaction effect between avatar and reputation, $F(4, 584) = 3.7$, $p = .006$, $\eta^2_p = .024$.

To break down this significant interaction effect, I subtracted the mean purchase decisions for the untrustworthy avatars from the mean purchase decisions for the trustworthy avatars for each of the five levels of reputation (see Figure 4).

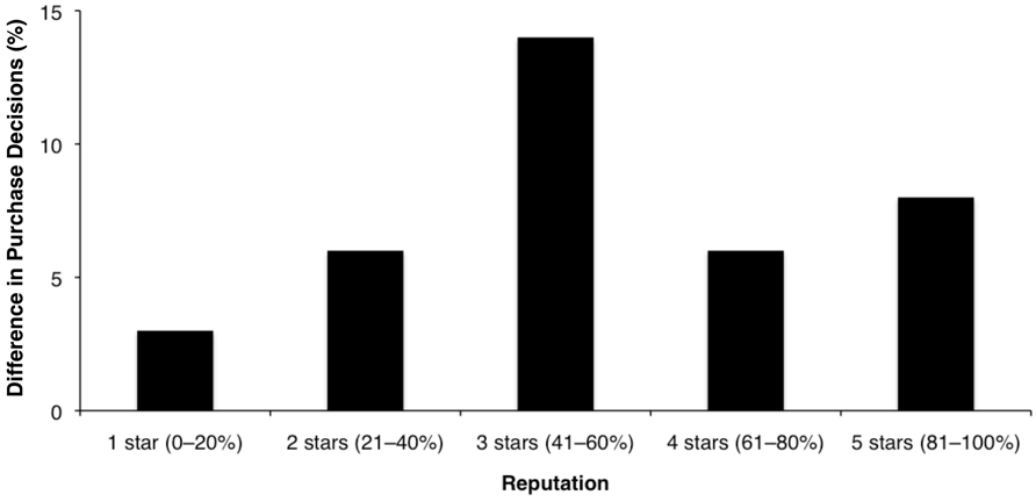

Figure 4. Difference in purchase decisions between trustworthy and untrustworthy avatars for all five levels of reputation.

To test the specific hypothesis that the difference in purchase decisions between trustworthy and untrustworthy avatars should be largest for the reputation score with the highest uncertainty that the seller would ship the product, I used a Helmert contrast. The difference in purchase decisions between trustworthy and untrustworthy avatars for the reputation score with the highest uncertainty (3 stars: 41–60 %) was compared to the mean difference in purchase decisions between trustworthy and untrustworthy avatars of all other reputation scores (1, 2, 4, and 5 stars). As predicted, this contrast was significant, $F(1, 146) = 6.26$, $p = .013$, $\eta^2_p = .041$, supporting hypothesis H4, which states that the effect of the seller avatar should be strongest in the condition with highest uncertainty that the seller would ship the product.

General Discussion

The aim of the present study was to investigate whether people trust avatars in a similar way as they trust photographs in an e-commerce setting. Using the paradigm by Bente et al. (2012), the present study replicated the main effects found by Bente et al. (2012) for seller reputation and seller avatar: Both positive reputation and trustworthy avatars led to higher purchase rates than negative reputation and untrustworthy avatars. With regard to missing information, the results of the present study also mirror the results by Bente et al. (2012). Negative reputation still led to significantly higher purchase rates than missing information, indicating that the uncertainty of missing information may be more aversive than the information conveyed by the negative reputation score (Bente et al., 2012). In line with the results by Bente et al. (2012), untrustworthy avatars did no differ significantly from missing avatars.

Based on the Elaboration Likelihood Model (Petty & Cacioppo, 1986), the present study also investigated whether participants' involvement—varied by the amount of units at stake in each trial—would influence the effect of seller avatars on purchase decisions. More precisely, it was hypothesized that when the stakes are high participants would use a central route of information processing relying more on reputation scores as opposed to seller avatars. However, the amount of units at stake in each trial (35 units vs. 3500 units) did not influence purchase decisions. This may be in part due to the fact that the conversion rate for each condition was adjusted (1:2 in the low involvement condition and 1:200 in the high involvement condition) so that the actual amount of money at stake in each trial (17.5 cents) was the same across the two different involvement conditions. Therefore, future studies investigating the influence of participants' involvement on purchase decisions should also vary the actual amount of money at stake.

Nevertheless, the present study did find some subtle difference in the effect of avatars as compared to photographs. In contrast to the results by Bente et al. (2012), the present study did find an interaction effect between seller avatar and reputation score: The effect of the avatar was stronger if there was high uncertainty that the seller would ship the product—as was the case in the negative reputation condition. If the uncertainty that the seller would ship the product was low—as was the case in the positive reputation condition—, the effect of the trustworthiness of the avatar was weaker. Supporting this notion, the effect size for reputation was greater ($\eta^2_p = .499$) than the effect size for avatar ($\eta^2_p = .240$), indicating that the effect of seller avatars was weaker than the effect of reputation. Hinting at an important difference in the effect of seller avatars as opposed to seller photographs, this result is in contrast to the findings by Bente et al. (2012) who found similar effect sizes for reputation ($\eta^2 = .27$) and seller photographs ($\eta^2 = .20$). However, as shown by the significant interaction between reputation and seller avatars in both Experiment 1 and 2, participants' responses to seller avatars were influenced by the uncertainty inherent in the reputation scores. When the uncertainty that the seller would ship the product was high, participants were more susceptible to the influence of the seller avatars. As the results show, there are subtle differences in the effects of virtual avatars as opposed to photographs. The present study therefore shows for the first time that avatars can produce different behavioral trust than real photographs. In a previous study, Riedl et al. (2011) could show a reduced activity in the mentalizing network when participants interacted with an avatar as opposed to a human. However, contrary to the findings of the present study, Riedl et al. (2011) did not find any differences between avatars and humans with regard to behavioral trust. The results of the present study show that the behavioral differences between avatars and humans can be subtle and may only be detectable when avatars are combined with other information sources.

In sum, the present study shows that in online transactions we are influenced by both reputation scores and seller avatars in our decisions. However, the effect of seller avatars can be moderated by reputation scores. The findings of the present study support the notion that seller avatars can serve as cues to build trust in online transactions, but that reputation scores have a greater impact on purchase decisions than avatars.

Study 2: Cultures of Trust

Introduction

Generalized trust is one of the corner stones of society (Knight, 2001). On a macro-level, high levels of trust are related to a society's productivity and economic success (Beugelsdijk, De Groot, & Van Schaik, 2004); on an individual level, trust has a positive effect on subjective well-being (Hudson, 2006). Generally, people can make their decision on whom to trust based on two sources of information: factual information or tacit cues. In the past, trust has been mostly studied from an economic viewpoint with a strong focus on humans as rational actors that use objective information from past interactions to decide when and whom to trust (e.g., Berg, Dickhaut, & McCabe, 1995). However, contrary to traditional economic theories stating that humans as rational actors should only be influenced by factual information, a large body of psychological studies has shown that humans are also strongly influenced by tacit, social cues. These social cues can lead to fast impressions of trustworthiness that have a strong impact on influential decisions. For instance, Willis and Todorov (2006) could show that exposure to a face for only 100 ms was sufficient for participants to form judgments of trustworthiness. Underlining the influence of such snap judgments, Ballew and Todorov (2007) showed that the trustworthiness of political candidates strongly predicted their success in elections—candidates who looked more trustworthy received more votes than candidates who looked less trustworthy.

In a recent study combining research from both the economic and psychological tradition on trust research, Bente et al. (2012) could show that German participants in an online trust game did use both sources of information—factual information and tacit, social cues—equally when deciding which sellers to trust, hinting at a general tendency to give equal weight to both sources of information.

However, as cross-cultural research of the past 30 years has shown, culture does have a strong influence on behavior and cognition (Markus & Kitayama, 1991), and even basic processes of cognition and perception are influenced by culture (e.g., Nisbett, Peng, Choi, & Norenzayan, 2001; Chua, Boland, & Nisbett, 2005). As recent studies using the prisoners' dilemma paradigm have shown (e.g., Gächter, Herrmann, & Thöni, 2010; Herrmann, Thöni, & Gächter, 2008), culture does have a great influence on the decision to trust others and to cooperate. It therefore remains an open question whether different cultures are equally influenced by both types of information when making judgments of trustworthiness.

Thus, the present study investigated whether individualistic and collectivistic cultures differ in their use of two different sources of information—factual information and tacit cues—in building trust by having participants from Germany, a more individualistic culture, and the United Arab Emirates, a more collectivistic culture, play an online trust game in which they made buying decisions based on sellers' reputation (factual information) and sellers' avatars (tacit cues).

Influence of Culture on Trust

Cross-cultural research on trust has been largely based on the cultural dimensions introduced by Hofstede (e.g., Hofstede, Hofstede, & Minkov, 2010). Especially the individualism/collectivism dimension has been used to explain and predict cross-cultural differences in trust and trust-related behavior. According to Hofstede (2010), members of collectivistic cultures define themselves over their membership in groups—the self and the group are strongly connected. On the other hand, members of individualistic cultures define themselves as separate, unique individuals—the self can be part of a group but remains autonomous (Markus & Kitayama, 1991).

In a large-scale cross-cultural study involving 31 different countries, Gheorghiu et al. (2009) could show that individualism/collectivism was strongly related to generalized social trust. Countries with strong collectivistic norms had lower levels of generalized social trust than countries with strong individualistic values. However, collectivists are not generally less trusting than individualists. As Huff and Kelley (2003) showed, the distinction between in- and out-group members is particularly important for collectivists and can explain lower levels of trust in collectivists. Even though both individualists and collectivists show a bias to trust the in-group, this bias is stronger in collectivists. Therefore, general lower levels of trust in collectivistic cultures can be explained by less trust to out-group members.

Several studies have investigated how different sources of information affect trust building in individualistic and collectivistic cultures: With regard to the effect of reputation, Branzei et al. (2007) showed that both individualists and collectivists were negatively influenced in their decision to trust by a low reputation. In addition, they found that members of individualistic cultures were more influenced by dispositional factors of trustees whereas members of collectivistic cultures paid more attention to situational factors. In a different study, Vishwanath (2004) demonstrated that the generalized trust level in a culture influenced the effect of reputation on buying decisions. When the level of trust within a culture was high, a low reputation did not negatively affect decisions to buy. However, when the level of trust within a culture was low, a low reputation did lead to less trust.

With regard to the effect of tacit, social cues, pictures had a strong influence on buying decisions in online auctions on buyers form a collectivistic culture as compared to buyers from an individualistic culture (Vishwanath, 2003). In addition, trust based on personal relationships is more important in collectivistic cultures, whereas trust based on formal rules is more important in individualistic cultures (Bohnet, Herrmann, & Zeckhauser, 2010). For instance, Yamagishi and Yamagishi (1994) found that members of collectivistic cultures rely more on social information and members of individualistic cultures more on reputation. Additionally, trust and cooperation are more strongly influenced on the basis of affect in collectivistic cultures, whereas individualistic cultures put more emphasis on cognitive elaboration (Dakhli, 2009).

To summarize the results on cross-cultural differences in trust, members from collectivistic cultures are less likely to trust out-group members (Huff & Kelley, 2003), put more emphasis on personal relationships (Bohnet et al., 2010), and pay more attention to situational factors (Branzei et al., 2007). On the other hand, members of individualistic cultures are more likely to trust out-group members (Huff & Kelley, 2003), prioritize formal rules over personal relationships (Bohnet et al., 2010), and rely more on dispositional information (Branzei et al., 2007). Members from both cultures are negatively influenced in their decision to trust by a low reputation (Branzei et al., 2007; Vishwanath, 2004). Also, social, affective information does seem to have a stronger influence on members of collectivistic cultures (Yamagishi & Yamagishi, 1994; Dakhli, 2009).

Nevertheless, there is so far no study that has investigated the influence of both factual information and tacit, social cues on trust building in individualistic and collectivistic cultures. Based on the paradigm by Bente et al. (2012), the present study therefore investigated two different cultures—Germany and the United Arab Emirates—that have been described as distinct on the individualism/collectivism dimension (Hofstede et al., 2010), with Germany being a more individualistic culture and the United Arab Emirates being a more collectivistic culture.

In line with the results by Bente et al. (2012), it was expected that despite of cross-cultural differences both cultures would be generally influenced by the two types of information—reputation scores and seller avatars—in the trust game. Thus, the following general hypothesis is stated:

> Hypothesis 1: Both reputation scores and seller avatars have a positive effect on purchase decisions. Positive reputation leads to more purchases than low reputation (H1a). Trustworthy seller avatars lead to more purchases than untrustworthy avatars (H1b).

Based on research showing that members of collectivistic cultures rely more on social information and members of individualistic cultures rely more on reputation (1994), it was hypothesized that both cultures differ in the weight they give to both sources of information:

> Hypothesis 2: The relative influence of reputation scores and seller avatars on purchase decisions does differ between Germans and Arabs. Purchase decisions of Germans are more influenced by reputation scores (H2a). Purchase decisions of Arabs are more influenced by seller avatars (H2b).

In sum, the present study investigated whether German and Arab participants differ in their purchase decisions in an online trust game when presented with both reputation scores (factual information) and seller avatars (tacit, social cues).

Methods

Participants. Overall, 88 Arab and German participants were invited via email to take part in an online trust game in which they could earn between 6.30 and 9 Dollar. Invitations were distributed through mailing lists of the University of Cologne (Germany) and the American University of Sharjah (United Arab Emirates).

Four participants were excluded from this original sample for the following reasons: One participant did not report his/her own culture and could therefore not be categorized as Arab or German. Two German participants incorrectly selected the English version of the study and were therefore presented with the Arab avatars instead of the German Avatars. One German participant was excluded from the study because his/her purchasing behavior differed from all other participants. That participant never made a purchase in all 18 experimental trials, resulting in a z score of -3.59.

Therefore, the final sample consisted of 42 German (9 male, 33 female; M_{age} = 27.19, SD_{age} = 8.95) and 42 Arab participants (9 male and 33 female; M_{age} = 27.19, SD_{age} = 8.95).

Design. The influence of culture, seller avatars, and reputation scores on purchase decisions was analyzed in a 2× 3 × 3 design, with culture (Arab vs. German) as a between-subjects factor and seller avatars (untrustworthy vs. trustworthy avatar) and reputation (high vs. low) as within-subjects factors

The main dependent variable was the number of purchases made in each condition.

Stimulus materials.

Trust game. Again, the standard trust game by Bolton (2004) provided the framework for the present study. In this trust game, a buyer (trustor) has to decide whether he will trust a seller (trustee) and buy a product from him/her. However, instead of actual products buyer and seller only exchange virtual units that are converted into real money at the end of the experiment. All possible decisions and outcomes of the trust game are shown in the pay-off matrix (see Figure 5). If the buyer decides not to buy (Case 1), both buyer and seller keep their 35 units. If the buyer decides to buy and the seller ships the product (Case 2a), both buyer and seller receive 50 units for the successful trade. If the buyer decides to buy and the seller does not ship the product (Case 2b), the buyer loses his/her 35 units to the seller, who receives 70 units.

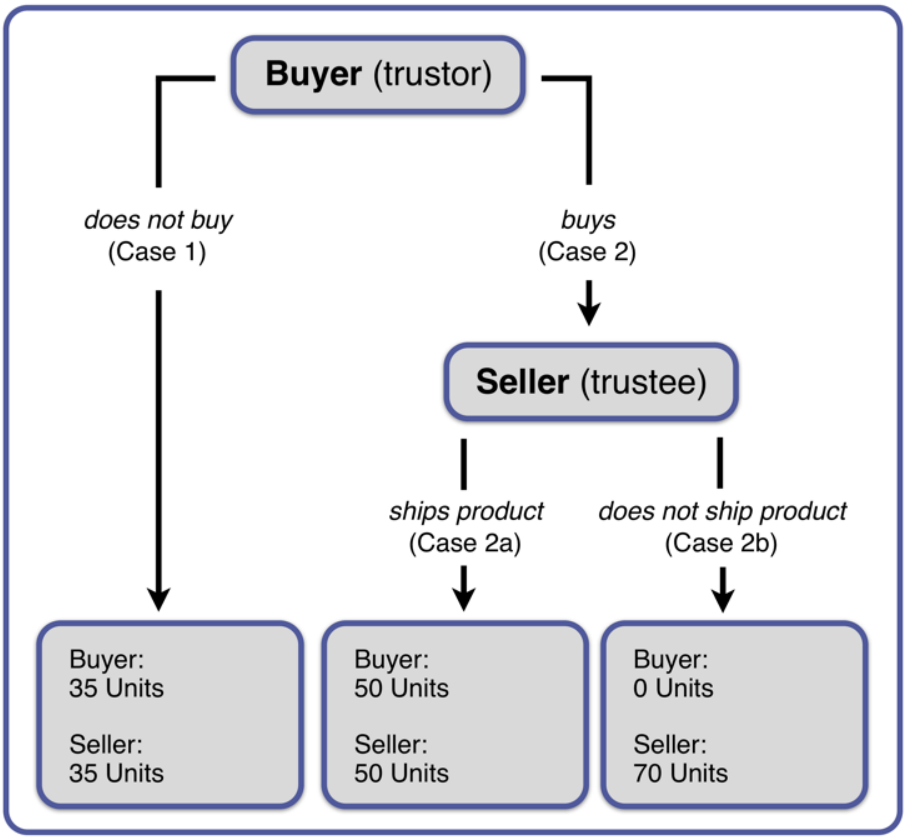

Figure 5. Pay-off matrix for the trust game showing all possible decisions and outcomes.

Seller avatars. A prestudy with Arab and German participants was conducted to select appropriate avatars for the study. To avoid the influence of stereotypes on decision making, participants in the final trust game were presented only with avatars from their own culture. German participants saw only German avatars, and Arab participants saw only Arab avatars. Therefore, the goal was to select two separate sets of trustworthy and untrustworthy avatars for each culture.

18 male German and 18 male Arab avatars matched for attractiveness and cultural typicality were selected from a database of Arab and German avatars (Bente et al., in preparation). German avatars were created with the software FaceGen (Singular Inversions, 2011), using portrait photographs of students at the University of Cologne. All participants in the database gave written consent for the use of their

photos. Arab avatars were created on the basis of the German avatars using the software's ethnicity transformation tool. Facial expression and gaze of all avatars was standardized.

Attractiveness ratings for all avatars in the database had been collected in a prestudy with 42 participants from Germany (4 male, 38 female; M_{age} = 23.4, SD_{age} = 4.4) and 31 students from the United Arab Emirates (7 male, 24 female; M_{age} = 26.4, SD_{age} = 7.9). Ratings for the cultural typicality and the gender of the avatars had been collected in another prestudy with 32 participants from Germany (4 male, 28 female; M_{age} = 23.7, SD_{age} = 3.9) and 19 participants from the United Arab Emirates (9 male, 10 female; M_{age} = 24.1, SD_{age} =5.7). Because several studies have shown that ratings of attractiveness and trustworthiness are strongly linked (Theodoridou, Rowe, Penton-Voak, & Rogers, 2009), attractiveness ratings from the prestudy were used as a proxy for trustworthiness. Therefore, avatars with medium high and low attractiveness ratings were selected.

To ensure that avatars would be clearly identified as Arab or German, only avatars with cultural typicality ratings higher than 80% were selected. In addition, only avatars that were unambiguously judged to be male were chosen. Thus, 18 male German and 18 male Arab avatars matched for attractiveness and cultural typicality were selected. To acquire ratings of trustworthiness for these 36 avatars, they were presented to 54 participants in another prestudy with 18 Arab (9 male, 9 female; M_{age} = 20.0, SD_{age} = 1.4) and 36 German participants (10 male, 26 female; M_{age} = 36.7, SD_{age} = 11.3) who rated the trustworthiness of the avatars on a 7-point scale ranging from 1 (very untrustworthy) to 7 (very trustworthy). Both German and Arab participants showed strong intercultural agreement with regard to the trustworthiness ratings of the avatars, indicated by a strong positive correlations between the trustworthiness from the two cultures, $r(40) = .73$, $p < .01$.

Avatars were ranked according to their trustworthiness ratings. The six avatars with the highest trustworthiness ratings in each culture were selected as the trustworthy avatars; the six avatars with the lowest trustworthiness ratings in each culture were selected as the untrustworthy avatars (see Figure 6 for two avatars from each culture).

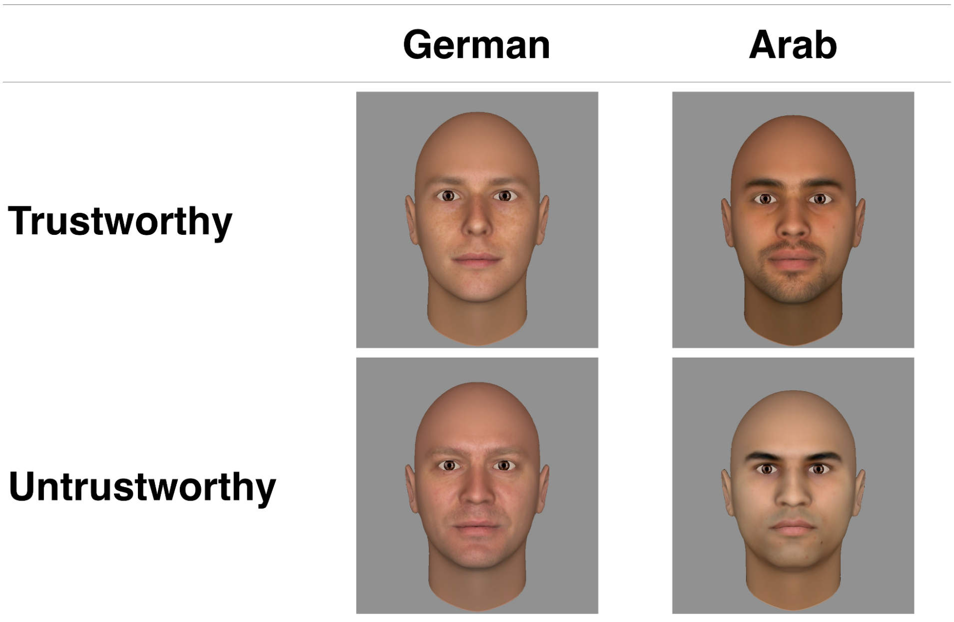

Figure 6. Trustworthy and untrustworthy avatars from both cultures (German and Arab).

In both cultures, trustworthy avatars received significantly higher trustworthiness ratings than untrustworthy avatars: Trustworthy Arab avatars (M = 4.32, SD = 1.03) received significantly higher trustworthiness ratings from Arab participants than untrustworthy Arab avatars, (M = 3.07, SD = 0.96), $t(17) = 5.14$, $p < .001$, d = 1.23. Trustworthy German (M = 3.96, SD = 0.90) avatars received significantly higher trustworthiness ratings from German participants than untrustworthy German avatars (M = 3.15, SD = 0.89), $t(35) = 7.83$, $p < .001$, d = 0.91.

The remaining six avatars from each culture were used as filler trails in the final experiment.

Reputation scores. The reputation of the sellers was symbolized using a reputation index with five stars—with the number of stars representing the percentage of previous trades in which the seller had shipped the product after receiving the payment (one star: 0–20%; two stars: 21–40%; three stars: 41–60%; four stars: 61–80%; five stars: 81–100%). Appropriate reputation scores were selected based on the prestudy by Bente et al. (2012): 30 participants (16 male, 14 female) rated the trustworthiness of the different levels of the reputation score on a 7-point scale ranging from 1 (*very untrustworthy*) to 7 (*very trustworthy*). Based on the results of the prestudy, two of the star-indexes were selected for the study: Three stars (the seller shipped the product 41–60% of the time) led to a significantly lower trust rating (M = 2.90, SD = 1.30) than four stars (the seller shipped the product 61–80% the time), $t(29) = 8.95$, $p < .001$, $d = 1.7$.

Procedure. All participants completed the study online using their own computers. At the beginning of the study, participants were informed that they would now play an online trust game with buyers and sellers, and that their own role in the trust game would be determined randomly. On the next page, participants were informed that they had been randomly assigned the role of the buyer—all participants were assigned the role of the buyer—and that they would now play the trust game with several other players who had been assigned the role of the seller. Participants were then informed that the sellers would be represented through virtual avatars generated out of real photographs. One photograph and one avatar were provided as an example. To avoid the influence of stereotypes on decision making, participants in the final trust game were presented only with avatars from their own culture. German participants saw only German avatars, and Arab participants saw only Arab avatars. After participants had been informed about the avatars, participants were introduced to the five-star reputation score. Participants were then informed that they would not receive immediate feedback on the decisions of the sellers, but that their final earnings would be calculated based on the decisions of the sellers. After that, participants proceeded to the actual profiles of the sellers that included the sellers' avatar and reputation, and decided for each profile whether they would buy from that particular seller (see Figure 7 for a sample profile).

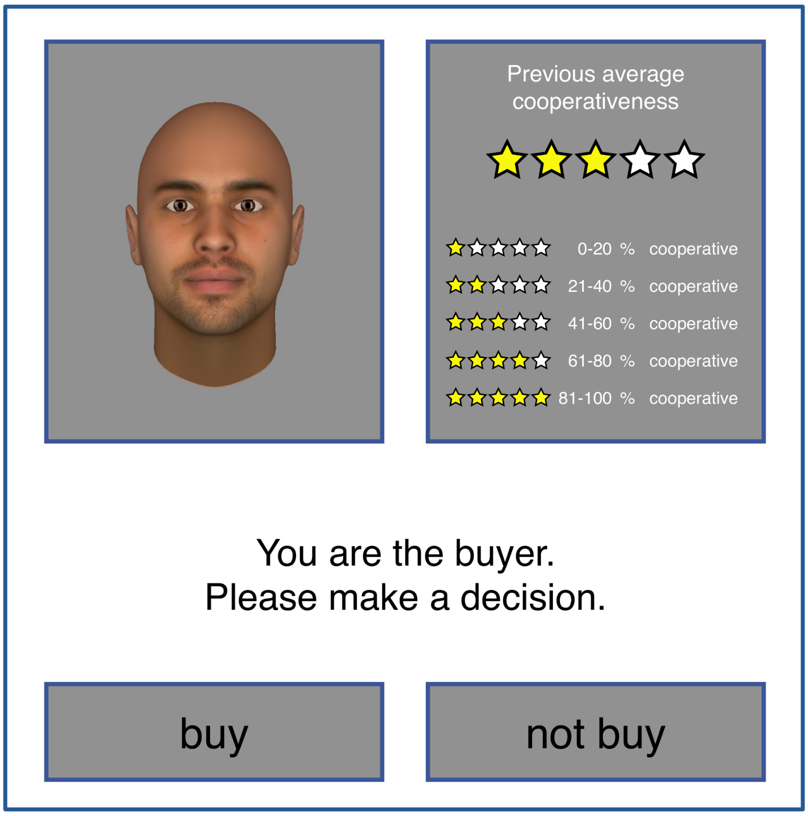

Figure 7. Screenshot of the profile of one seller in the trust game, showing a trustworthy Arab avatar with a low reputation score (three stars).

Participants then completed 12 experimental trials and 6 filler trials (two filler trials in the beginning, four filler trials randomly mixed with the experimental trials). Filler trials always used reputation scores other than three or four stars and were included to distract participants from the repeated presentation of reputation scores three and four. After participants had completed all trials, they were thanked for their participation and informed that they would receive a voucher for Amazon worth their payment. Payment was calculated based on participants' decisions in the trust game. If participants had decided to buy the product, payment was always calculated as if the seller had shipped the product. Final payment to participants ranged from 6.30 to 8.55 Dollar (M = 7.78, SD = 0.41).

Dependent variable. The number of purchases in each condition was used as the dependent variable.

Results

Because there was an unequal distribution of male and female participants in both cultures, $\chi^2(1, N = 84) = 15.75$, $p < .001$, it was first checked whether participant's gender had any influence on the results. A 2 × 2 × 2 × 2 (gender × culture × reputation × seller avatars) mixed ANOVA revealed no significant main effect for gender, $F(1, 80) = .31$, $p = .58$, $\eta^2_p = .004$, nor any significant interaction effects, $F <$ 2.9. Therefore, gender was dropped as a factor from all further analyses.

Buying decisions were analyzed in a 2 × 2 × 2 (culture × reputation × seller depictions) mixed ANOVA. There was a significant main effect for reputation, $F(1, 82) = 56.37$ $p < .001$, $\eta^2_p = .41$, indicating that a high reputation (M = 2.40, SD = 0.71) caused participants to buy significantly more often than a low reputation (M = 1.40, SD =.99). Thus, Hypothesis H1a, stating that reputation has a positive influence on buying decisions, was supported.

There was also a significant main effect for seller avatars, $F(1, 82) = 17.74$, $p < .001$, $\eta^2_p = .18$, indicating that a trustworthy avatar (M = 2.08, SD = 0.65) caused participants to buy significantly more often than an untrustworthy avatar (M = 1.71, *SD* = 0.81). Therefore, Hypothesis H1b, stating that seller avatars have a positive influence on buying decisions, was also supported.

There was no significant main effect for culture, $F(1, 82) = 1.56$, $p = .215$, $\eta^2_p = .019$, indicating that neither of the cultures bought generally more in the trust game than the other. There was also no significant three-way interaction between reputation, seller avatars, and culture, $F(1, 82) = 0.34$, $p = .854$, $\eta^2_p < .001$.

Thus, Hypothesis H2 in general, stating that reputation and seller avatars would have different effects in the two cultures, with Germans being more influenced by reputation and Arabs being more influenced by seller avatars, was not supported.

There was no significant two-way interaction between culture and avatar, $F(1, 82) = .30$, $p = .\ 59$, $\eta^2_p = .004$. Thus, hypothesis H2a could not be supported. There was no significant interaction between avatar and reputation, $F(1, 82) < .01$, $p = 1.00$, $\eta^2_p < .001$.

However, There was significant two-way interaction between reputation and culture, $F(1, 82) = 4.23$, $p = .043$, $\eta^2_p = .049$, showing that German and Arab participants responded differently to sellers with high and low reputation (see Figure 8). Thus, hypothesis H2b could be partially supported.

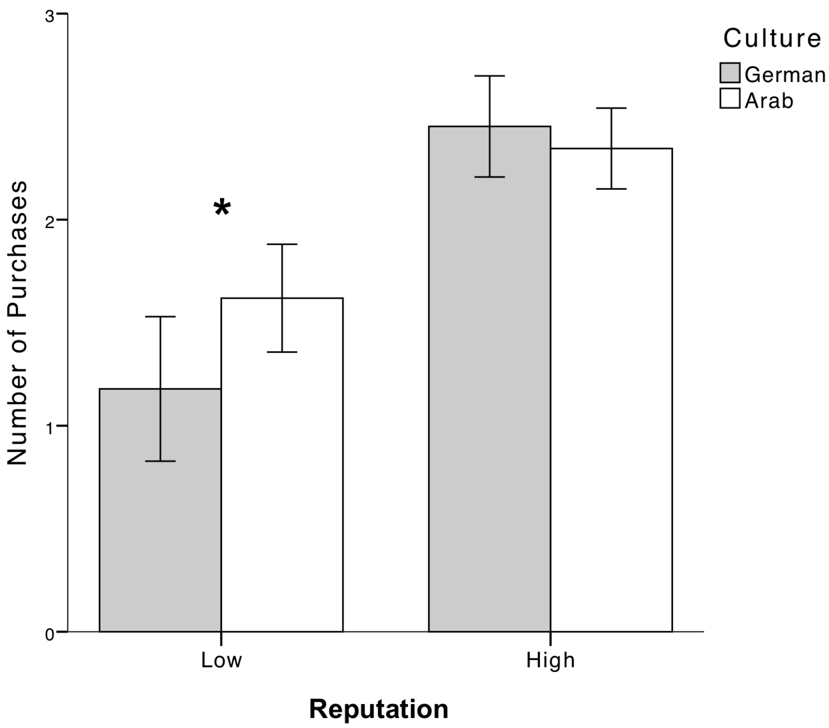

Figure 8. Two-way interaction between reputation and culture. * $p < .05$.

Pairwise comparisons showed that German participants ($M = 1.18$, $SD = 1.13$) bought significantly less often when the reputation was low than Arab participants ($M = 1.62$, $SD = 0.84$), $t(82) = -2.03$, $p = .045$, $d = -0.44$. There was no difference in the number of purchases between German ($M = 2.45$, $SD = 0.79$) and Arab participants ($M = 2.35$, $SD = 0.63$) when the reputation of the seller was high, $t(82) = 0.69$, $p = .493$, $d = 0.14$.

Discussion

The present study investigated whether individualistic and collectivistic cultures differ in their use of two different sources of information—factual information and tacit cues—in building trust. As the results of the trust game show, participants from Germany, a more individualistic culture, and the United Arab Emirates, a more collectivistic culture, were influenced by both reputation scores and seller avatars. These results are in line with previous studies that have found a positive effect of reputation and seller photographs on buying decisions (e.g., Bente et al., 2012; Branzei et al., 2007; Vishwanath, 2004). The results of the present study extend the findings of Bente et al. (2012)—found in Western participants—to members of a collectivistic culture: the United Arab Emirates.

In contrast to previous studies that have found that members of collectivistic cultures show lower levels of trust than members of individualistic cultures (2009), the present study did not find a general effect of culture on the willingness to trust. In the trust game, Arab participants were not less likely to trust potential sellers than German participants. As Huff and Kelley (2003) could show, the distinction between in- and out-group members may be particularly important for members of collectivistic cultures. Therefore, in the present study, the culture of the seller avatars and of the participants had been matched—Arab participants saw only Arab avatars, and German participants saw only German avatars. The results show that this matching was successful in preventing a potential in-group bias in the Arab participants from distorting the results. However, a future study should present sellers from both cultures to Arab and German participants to test if the reported in-group bias in collectivists can be replicated (Huff & Kelley, 2003).

Based on previous cross-cultural research on the influence of social information as opposed to reputation (Yamagishi & Yamagishi, 1994), it was hypothesized that Arab and German participants would differ in their susceptibility to reputation scores and seller avatars. However, this hypothesis was only partially supported by the results. As for the effect of seller avatars, the present study did not find a stronger influence of seller avatars on members of a collectivistic culture.

Nevertheless, both cultures did differ in the effect of reputation on purchase decisions. German participants bought significantly less often than Arab participants when the reputation of the seller was low. This result is in line with previous studies that have shown that members of individualistic cultures rely more on formal criteria and factual information (Bohnet et al., 2010; Yamagishi & Yamagishi, 1994). In the present study, reputation was operationalized as the percentage of previous trades in which the seller had successfully shipped the product. The reputation could therefore be conceived of as a disposition of the seller to ship the product. As a previous study by Branzei et al. (2007) has shown, members of individualistic cultures are more influenced by dispositional factors of trustees whereas members of collectivistic cultures pay more attention to situational factors. More generally, it has been shown that members of collectivistic cultures rely more on situational factors when predicting behavior whereas members of individualistic cultures rely more on dispositional factors (Lee, Hallahan, & Herzog, 1996). It is therefore possible that German participants conceived of the reputation of the seller as a disposition predictive of his/her behavior in the current transaction. Arab participants, on the other hand, may have considered other situational factors competing with the predictive value of the reputation. Thus, future studies should investigate why particularly negative information may have a strong predictive value for individualists as opposed to collectivists.

Another explanation might involve both cultures' generalized level of trust. As Vishwanath (2004) has shown, the generalized trust level in a culture influences the effect of reputation on buying decisions. When the level of trust within a culture is high, a low reputation does not negatively affect buying decisions. However, when the level of trust within a culture is low, a low reputation does lead to less trust.

Germany was among the cultures investigated by Vishwanath (2004), displaying medium levels of generalized trust, and, consequently, German participants were moderately affected by a low reputation in their buying decisions. It is therefore possible that the difference in purchase decisions under low reputation in the present study can be explained by different levels of generalized trust between the German and Arab participants. However, generalized trust in Arab countries tends to be low (Jamal, 2007), casting at least some doubt on this explanation. Nevertheless, future cross-cultural studies on the influence of reputation on buying decisions should measure participants' generalized trust level as a possible covariate.

With regard to the use of virtual avatars as stimuli in the present study, the results of the present study are similar to the outcome of Study I. Again, the effect sizes for virtual avatars were smaller ($\eta^2_p = .18$) than for reputation ($\eta^2_p = .41$). Taking into account the equal effect sizes for photographs and reputation found by Bente et al. (2012), the results further support the notion that virtual avatars produce slightly weaker responses than actual photographs. Nevertheless, virtual avatars were used by members of both cultures as a source of information that influenced behavioral trust, hinting at the possibility that the tendency to trust avatars may not be limited to Western cultures.

Taken together, the present study shows that members of individualistic and collectivistic cultures both rely on factual information and tacit, social cues in building trust. However, negative reputation seems to have a special salience for members of individualistic cultures, translating into a decrease in behavioral trust and a drop in purchases.

Study 3: Direct Gaze Detection in High-functioning Autism

Introduction

The perception of direct gaze—i.e., being gazed at directly by another person—is crucial to human social interaction: It can signal the attention of others and mediate the perception of displayed emotions (Adams & Kleck, 2005). In addition, it provides the basis for mutual gaze, the shared gaze between two persons. The ability to detect direct gaze is a fundamental function established early in our evolution (Myowa-Yamakoshi, Tomonaga, Tanaka, & Matsuzawa, 2003) and can be measured as early as in 2- to 5-day-old infants, who are able to distinguish between faces with direct and averted gaze, and who prefer to look at faces with direct gaze (Farroni, Csibra, Simion, & Johnson, 2002).

Recent studies have focused on the question of how accurate individuals with autism are at detecting direct gaze (see Nation & Penny, 2008; Senju & Johnson, 2009). Even though there are several studies that hint at deficits (Gepner, de Gelder, & de Schonen, 1996; Howard et al. 2000; Swettenham et al., 2001; Senju, Yaguchi, Tojo, & Hasegawa, 2003; Wallace, Coleman, Pascalis, & Bailey, 2006), there are also studies that show no deficits (Senju, Hasegawa, & Tojo, 2005; Senju, Kikuchi, Hasegawa, Tojo, & Osanai, 2008; Webster & Potter, 2011). In addition, the developmental trajectory of the putative impairment is unclear. Because accumulating neuropsychological evidence hints at a general difference between individuals with autism and neurotypical individuals in the way they process direct gaze (Grice et al., 2005; Kylliäinen & Hietanen, 2006; Joseph, Ehrman, McNally, & Keehn, 2008; Elsabbagh et al., 2009; Akechi et al., 2010; Pitskel et al., 2011), it might be possible that the ability to detect direct gaze is impaired even in adults with autism.

Only two studies so far have investigated the ability to detect direct gaze in adults with autism: Howard et al. (2000) found differences between individuals with autism and neurotypical individuals on a task in which participants had to select one image displaying direct gaze out of two. However, using the same paradigm, Webster and Potter (2011) were not able to replicate this finding. The fact that this impairment has not been consistently observed can in part be explained by several shortcomings of previous studies.

First, all studies so far have used tasks that do not measure the critical ability in question—namely, being able to differentiate between direct gaze and averted gaze. All studies did use either a comparison task (Gepner et al., 1996; Howard et al., 2000; Swettenham et al., 2001; Webster & Potter, 2011) or a visual search task (Senju et al., 2003; Senju et al., 2005; Senju et al., 2008). In the comparison task, participants had to select out of pairs of images the one displaying direct gaze. In the visual search task, participants had to find one face displaying direct gaze that was presented along with several faces showing averted gaze. Whereas these tasks might be useful to find crude differences in accuracy to detect direct gaze, different task settings modeling natural situations that require detection of direct gaze might provide a better framework to detect subtle impairments (Wilms et al., 2010).

Second, all studies used large variations of gaze direction. The visual search task described above included only two variations—direct gaze or averted gaze. Even though the degrees of deviation were not reported in the studies (e.g., Senju et al., 2003), compared to the stimuli used in the current experiment it is safe to conclude that the gaze was at least averted by 15°. In the comparison task (e.g., Webster & Potter, 2011), gaze was averted by either 20°, 15°, 10°, or 5°. Psychophysical studies investigating the accuracy to detect deviations from direct gaze, however, have found that humans are able to distinguish degrees between direct and averted gaze that are far more subtle (Gamer & Hecht, 2007). Thus, previous studies did not use stimuli in a range (between 0°–10°) that provides a sufficient level of ambiguity to detect group differences.

All these factors might explain why previous studies of direct gaze detection in autism have failed to find clear differences between individuals with autism and typically developing individuals. Thus, these shortcomings were addressed in the present study by using a realistic virtual character showing realistic gaze behavior that varied in small increments between direct and averted gaze.

Additionally, it seems important to consider what other factors of gaze behavior might contribute to a deficit in detecting direct gaze in autism. One important aspect of gaze behavior that previous studies have ignored is the social nature of gaze (Wilms et al., 2010). In naturalistic situations, gaze is embedded in a social context, signaling communicative meaning and creating affordances to act (Schilbach, Eickhoff, Cieslik, Kuzmanovic, & Vogeley, 2011). Most importantly, gaze is often ambiguous and unpredictable. It might be this unpredictability that could interfere with the ability to detect direct gaze in individuals with autism. Several theories have identified the predictability of social situations as an important factor influencing the behavior of individuals with autism (Gomot & Wicker, 2011; Qian & Lipkin, 2011; Baron-Cohen, 2009; Dawson, Meltzoff, Osterling, Rinaldi, & Brown, 1998). For instance, Baron-Cohen et al. (2009) proposed that autism is characterized by a strong drive to predict and control the environment. Consequently, it has been shown that children with autism favor predictable over unpredictable environments (Ferrara & Hill, 1980) and exhibit more problematic behaviors, such as self-hits and aggression, in unpredictable environments (Flannery & Horner, 1994). If the unpredictability of a situation is removed by imitating children with autism, communicative gaze increases (Sanefuji & Ohgami, 2011). This preference for predictable over unpredictable stimuli seems to extend to simple displays of motion. When presented with point-light-displays of either biological motion or non-social contingencies, two-year-olds with autism spend more time looking at non-social contingencies (Klin, Lin, Gorrindo, Ramsay, & Jones, 2009). Besides preferring predictable stimuli, children with autism seem to be impaired in their ability to predict the variability of common real-life events (Loth, Happé, & Gómez, 2010).

One of the key features of social gaze is its independence and goal-directedness that can only be partly predicted by relying on general knowledge about social situations or taking the perspective of another person. Inanimate objects and non-social systems, on the other hand, can be understood and predicted in a mechanistic way by observing and controlling the relations between the input and output of the system.

Social gaze, therefore, poses a double burden to individuals with autism: First, because of a general deficit in social interaction, individuals with autism are impaired in their ability to understand and interpret the gaze of another person and are therefore not able to predict it. Second, because the gaze of another person usually cannot be controlled, individuals with autism are not able to explore the underlying rules of the gaze behavior to make it more predictable in the future. In other words, individuals with autism might be impaired in the processing of social gaze because it appears unpredictable to them. A possible explanation for the development of this impairment has been provided by the social orienting hypothesis (Dawson et al., 1998), according to which children with autism are less drawn to social stimuli because they seem unpredictable to them. Because they spend less time attending to social stimuli than typically developing children, they miss the opportunity to learn about social stimuli, which would make them more predictable to them. Thus, a negative feedback loop develops—unpredictable social stimuli cause children with autism to spend less time attending to them, which makes social stimuli even more unpredictable and causes children with autism to spend even less time attending to them. One way to break this feedback loop would be to create predictable social environments that can be controlled by individuals with autism to help them overcome their deficits.

Based on these assumptions, the following prediction can be made: If given the opportunity to control the gaze of another person and thus increase the predictability of the situation, individuals with autism should be able to overcome their deficit. Because the gaze of another person usually cannot be controlled in real-life situations, individuals with autism have no opportunity to experience social gaze in a predictable environment. However, computer-generated stimuli allow to create social gaze stimuli that are ecologically valid and yet can be well controlled (Wilms et al., 2010). This approach was used in the present study to investigate the relation between the controllability of social stimuli and the detection of direct gaze by individuals with autism.

In sum, the aim of the present study was twofold: First, using a realistic virtual character displaying dynamic gaze behavior with fine-tuned variations, it was investigated whether individuals diagnosed with autism are impaired in their ability to detect direct gaze. Second, by giving participants control over the gaze of a social character, the effect of having control over a social stimulus on gaze detection abilities was examined.

Experiment 1

Methods

Participants. Nineteen adults with high-functioning autism (10 male, 9 female; M_{age} = 39.1, age range: 23–53 years) were recruited at the Adult Autism Outpatient Clinic of the Department of Psychiatry at the University Hospital of a large city in Western Germany. All participants with high-functioning autism were diagnosed by two independent physicians according to ICD-10 criteria. In line with other studies (e.g., Frith & de Vignemont, 2005), high-functioning autism is used as an umbrella term for both high-functioning autism and Asperger syndrome. Nineteen control participants (11 male, 8 female; M_{age} = 33.8, age range: 20–48 years) were recruited at the Department of Psychology at the University of a large city in Western Germany. Both groups were matched for age, gender, years of formal education, and intelligence (measured with the German version of the Wechsler Adult Intelligence Scale-Revised (WAIS-R; Tewes, 1994). Participants also completed the Beck Depression Inventory (BDI; Beck & Steer, 1987; Hautzinger, 1995), the Autism Spectrum Quotient (AQ; Baron-Cohen, Hoekstra, Knickmeyer, & Wheelwright, 2006), and the "Reading the Mind in the Eyes Test" (ToM-Eyes; Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001).

On average, there was no difference in age between participants diagnosed with autism (M = 39.1, SD = 9.4) and participants in the control group (M = 33.8, SD = 7.3), $t(36) = 1.95$, $p = .060$, $d = 0.63$. See Table 3 for an overview of demographic and psychopathological variables for both groups.

Table 3

Demographic, Psychopathological, and IQ Results for the Pilot Study

Variable	HFA (n = 19) *M*	HFA (n = 19) *SD*	Control (n = 19) *M*	Control (n = 19) *SD*	$t(36)$	*p*	Cohen's *d*
Age (years)	39.1	9.4	33.8	7.3	1.95	.060	0.63
Education (years)	18.3	3.0	20.1	4.5	-1.42	.163	-0.47
WAIS-R verbal IQ	124.9	11.5	126.6	7.4	-0.52	.605	-0.18
WAIS-R performance IQ	123.3	13.8	127.7	11.4	-1.08	.288	-0.35
WAIS-R IQ (total)	127.0	12.3	132.2	8.1	-1.53	.135	-0.50
BDI	11.3	7.1	4.9	3.1	3.60	.001	1.17
AQ	41.1	3.8	14.2	5.5	17.66	<.001	5.69
ToM-Eyes	16.2	4.2	18.6	3.1	-2.03	.050	-0.65

Note. WAIS-R = Wechsler Intelligence Scale for Adults, BDI = Beck Depression Inventory, AQ = Autism Spectrum Quotient, ToM-Eyes = Reading the Mind in the Eyes Test

There was also no difference between the HFA group (M = 18.3, SD = 3.0) and the control group (M = 20.1, SD = 4.5) with regard to years of formal education, $t(36) = -1.42$, $p = .163$, $d = -0.47$. Furthermore, there was no difference in intelligence between participants diagnosed with autism (M = 127.0, SD = 12.3) and participants in the control group (M = 132.2, SD = 8.1), $t(36) = -1.53$, $p = .135$, $d = -0.50$. Participants diagnosed with autism had higher scores on the Beck Depression Inventory (BDI; Beck & Steer, 1987; Hautzinger, 1995) (M = 11.3, SD = 7.1) than participants in the control group (M = 4.9, SD = 3.1), $t(36) = 3.60$, $p = .001$, $d = 1.17$. This was to be expected because high-functioning autism is known to have a higher prevalence of depression (Stewart, Barnard, Pearson, Hasan, & O'Brien, 2006; Lehnhardt et al., 2011). Confirming clinical diagnosis, participants diagnosed with autism scored higher on the Autism Spectrum Quotient (*M* = 41.1, *SD* = 3.8) than participants in the control group (*M* = 14.2, *SD* = 5.5), $t(36) = 17.66$, $p < .001$, d = 5.69. In addition, there was a marginally significant difference on the ToM-Eyes test between the HFA group (*M* = 16.2, *SD* = 4.2) and the control group (*M* = 18.6, *SD* = 3.1), $t(36) = -2.03$, $p = .050$, $d = -0.65$.

Stimulus materials. In the detection task, participants' accuracy to detect when a realistic virtual character was gazing directly at them was tested.

As the stimulus in the detection task, the present study used a realistic virtual character created in the commercially available software Poser 6 (see Figure 9a).

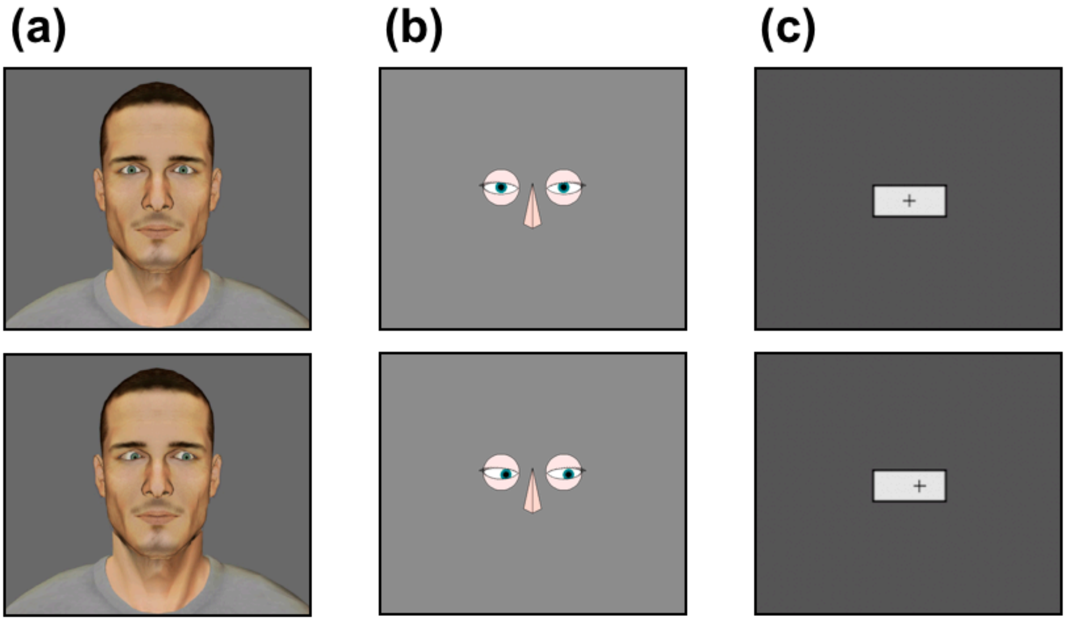

Figure 9. **(a)** Realistic virtual character, **(b)** reduced virtual character, and **(c)** non-social, geometric stimulus.

Because the criterion for judging gaze as direct can be biased by a smiling expression, with smiles increasing the chance of perceiving gaze as direct (Martin & Rovira, 1982), the virtual character showed a neutral expression. The virtual character was moving its eyes and fixated different positions. Fixations included direct gaze, slightly averted gaze, and clearly averted gaze. To describe the position of the eyes, a system with two rotations (y and z) was used. Y-rotation described the degree of deviation from the center of the eyes (see Figure 10a); z-rotation described the direction of the deviation from the center of the eyes (see Figure 10b).

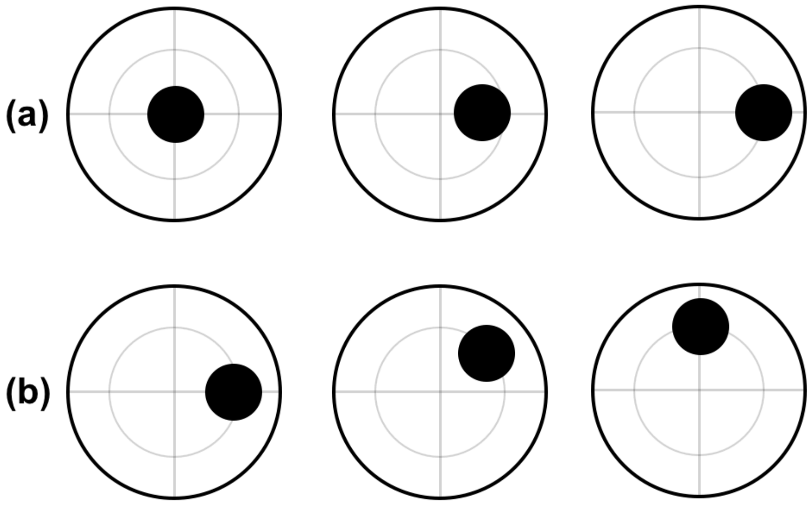

Figure 10. Gaze positions at different degrees of **(a)** y-rotation and **(b)** z-rotation.

To determine which degrees of y-rotation would be perceived as direct gaze, a pilot study was conducted. 20 participants (13 female, 7 male) saw pictures of the virtual character with the gaze angle varying on the y-axis from 1° to 10°. For every picture, the participants rated whether they felt that the virtual character's gaze was direct. Rotations of 1° and 2° were mistaken for direct gaze by 80% of the participants (3°: 70%, 4°: 60%, 5°: 35%, 6°: 10%). All participants correctly identified deviations of 7° and more as averted gaze. As a result, gaze angles ranging from 0° to 7° on the y-axis were used in the experiment, because they provided a sufficient level of ambiguity. Because the y-axis only determined the amount of deviation from the center of the eye, different degrees on the z-axis were used to vary the direction of the virtual character’s gaze. Combining both y- and z-angles a realistic gaze sequence with 10 hits and 20 distractors was created. The order in which the different y-angles appeared in the gaze sequence was completely randomized and a set of 30 trials was programmed. At the beginning of each trial, the virtual character showed clearly averted gaze (15 – 25° on the y-axis) for 2,000 ms. After that, it moved its eyes to a fixation point of either 0° (hit) or 1-7° (false alarm). The duration of the fixation was randomly chosen to be either 2,000 ms or 4,000 ms based on earlier studies on social gaze (Bente, Donaghy, & Suwelack, 1998; Bente, Eschenburg, & Krämer, 2007; Kuzmanovic et al., 2009) that have shown that a gaze duration of around 4,000 ms is perceived as pleasant. Therefore, each trial lasted either 4,000 ms or 6,000 ms. Participants used a mouse to signal when the virtual character was showing direct gaze, clicking the left mouse button each time they felt the virtual character was gazing directly at them. Dependent variables were reaction times, the number of correct responses (hits), and the number of false alarms. Because previous studies have shown that social stimuli, such as faces, are less pleasant for individuals with autism (Corden, Chilvers, & Skuse, 2008; Hutt & Ounsted, 1966; Richer & Coss, 1976), participants were asked to rate the pleasantness (1 = *very unpleasant* to 5 = *very pleasant*) of the task on a 5-point scale. If gaze detection is impaired in individuals with autism, a task demanding participants to detect direct gaze might be perceived as difficult. Therefore, participants were also asked to rate the difficulty (1 = *very easy* to 5 = *very difficult*) of the task on a 5-point scale.

In the setting task, participants' accuracy to actively establish direct gaze with a virtual character was assessed. As the virtual character, a reduced avatar was used (see Figure 9b) that had been employed in previous experiments of this research group (Bente et al., 2007). Instead of detecting when the virtual character was gazing at them, participants actively established direct gaze using the mouse. In each trial, the eyes of the virtual character appeared in a random position of averted gaze. In addition, the rotation of the virtual character´s head was also varied between 0° and 30° on the horizontal, vertical, and sagittal axis. Participants used the mouse to move the eyes of the virtual character into a position that matched direct gaze as closely as possible. When the participants had established that position, they pressed a button on the mouse and the eyes of the virtual character moved to a new random position. Participants completed a sequence of 16 trials. There was no time limit for each trial to allow the participants to work as accurately as possible. Dependent variables were reaction times and degrees of deviation from a perfect eye contact.

Procedure. All participants were tested at the Department of Psychiatry at the University Hospital of a large city in Western Germany. Prior to the experiment, all participants gave consent to participate. All participants had normal or corrected-to-normal visual acuity. The experimental stimuli were presented on a 17-in. monitor with a resolution of 1280 x 1024 pixels. Each participant completed both the setting and detection task in one single session. To eliminate carry-over-effects, all tasks were presented in random order. At the beginning of the experiment, the experimenter greeted the participants and told them that they were going to take part in an experiment on social perception. After that, they sat down in a chair approximately 50 cm from the computer. They positioned their head so that the eyes were looking at the centre of the screen, and the experimenter started the experiment. The participants then saw a grey screen with white instructions explaining the first of the two tasks. Participants could read the instructions at their own pace and then proceeded to the actual task. After they had completed the first task, participants proceeded to the instructions for the second task. Participants rated the pleasantness and the difficulty of the detection task on a paper questionnaire next to the computer. The experiment ended, after participants had completed all tasks. After the experiment, participants completed the BDI (Beck & Steer, 1987; Hautzinger, 1995), the AQ (Baron-Cohen et al., 2006), the WAIS-R (Tewes, 1994), and the ToM-Eyes test (Baron-Cohen et al., 2001).

Results and Discussion

Detection task. Accurate gaze detection includes both correctly detecting direct gaze and classifying gaze as averted that is not direct. A person classifying every gaze direction in the experiment as direct would correctly detect all instances of direct gaze, but the high rate of false alarms would indicate that this person is not truly able to distinguish between direct and averted gaze. To compare both groups' ability to detect direct gaze, a sensitivity index *d'* was calculated out of the number of correct responses and the number of false alarms. There was a significant difference in *d'* between the HFA group (M = 0.18, SD = 0.51) and the control group (M = 1.00, SD = 0.45), $t(36) = -5.19$, $p < .001$, $d = -1.71$ (see Figure 11a).

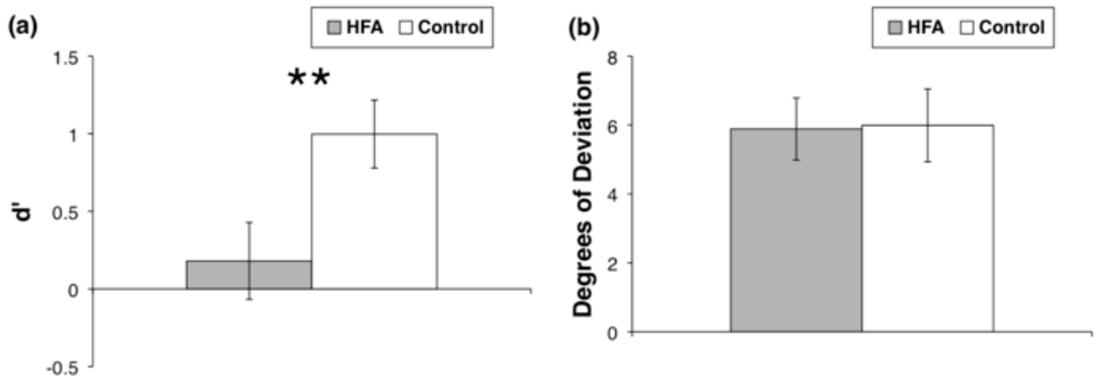

Figure 11. Overall results of the pilot study: **(a)** detection task and **(b)** setting task. Error bars represent 95%-CI of the mean. * $p < .05$, ** $p < .001$.

To analyze at which degrees of deviation both groups differed, the percentage of false alarms at each angle, ranging from 1°–7°, was compared in a 2 × 7 mixed ANOVA (group × angle). As expected, there was a significant effect of group on percent of false alarms, $F(1, 36) = 36.82$, $p < .001$, $\eta_p^2 = .506$. There was also a significant effect of angle, $F(6, 216) = 6.41$, $p < .001$, $\eta_p^2 = .151$. There was no significant interaction effect between group and angle, $F(6, 216) = 0.26$, $p = .957$, $\eta_p^2 = .007$. Independent sample t tests using a Bonferroni-corrected alpha ($\alpha = .05/7 = .007$) were used to compare both group's performance for the different degrees (1°–7°). As shown in Table 4 and Figure 12, participants in the HFA group produced significantly more false alarms than control participants at 3°, 4°, 5°, and 6°.

Table 4

Percent of False Alarms at Different Angles (1°–7°)

	HFA (n = 19)		Control (n = 19)				
Angle	M	SD	M	SD	t(36)	p	Cohen's d
1°	100.00	0.00	78.95	41.89	2.19	.042	0.71
2°	100.00	0.00	68.42	47.76	2.88	.010	0.94
3°	92.11	18.73	63.16	36.67	3.06	.005	0.99
4°	83.16	17.97	58.95	22.58	3.66	.001	1.19
5°	86.32	21.14	63.16	23.35	3.21	.003	1.04
6°	72.37	35.25	38.16	33.71	3.06	.004	0.99
7°	76.32	34.83	47.37	42.41	2.30	.028	0.75

Note. Using a Bonferroni-corrected alpha ($\alpha = .05/7 = .007$), only p values below .007 are considered to be significant.

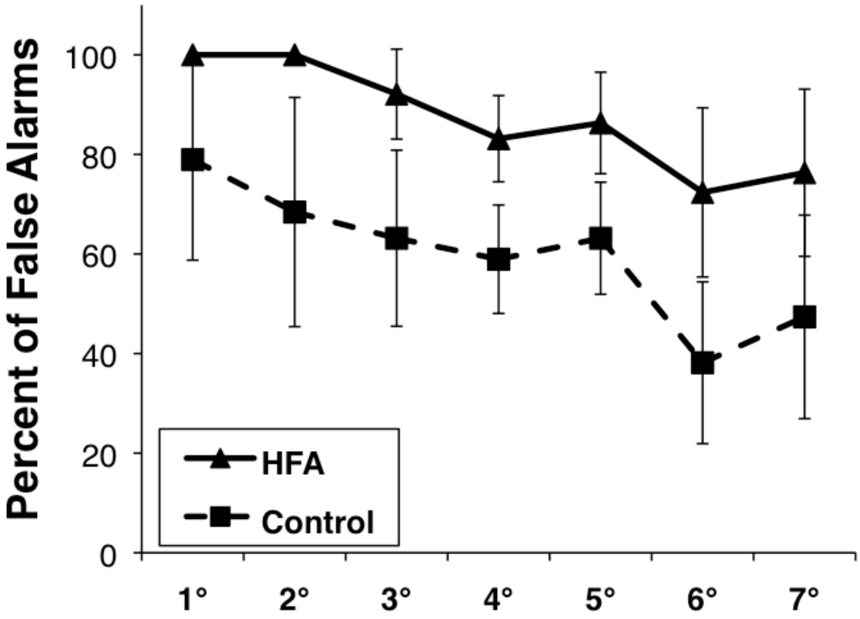

Figure 12. Percent of false alarms at different angles (1°–7°). Error bars represent 95%-CI of the mean.

With regard to reaction times, there was no significant difference between the HFA group (M = 1,180 ms, SD = 669 ms) and the control group (M = 1,340 ms, SD = 421 ms), $t(36) = -0.99$, $p = .329$, $d = -0.32$.

Perceived pleasantness of the task did differ significantly between the HFA group (M = 2.56, SD = 0.92) and the control group (M = 3.47, SD = 0.90), $t(35) = -3.06$, $p = .004$, $d = -1.00$. This indicates that participants in the control group perceived the task with the virtual character to be more pleasant than participants in the HFA group.

In addition, participants in the HFA group (M = 2.89, SD = 1.08) did perceive the task to be more difficult than participants in the control group (M = 1.79, SD = 0.85), $t(35) = 3.45$, $p = .001$, $d = 1.13$.

Setting task. There was no significant difference between the HFA group (*M* = 5.89, *SD* = 1.87) and the control group (*M* = 5.99, *SD* = 2.19) in the accuracy to establish direct gaze, $t(36) = -0.16$, $p = .876$, $d = -0.05$ (see Figure 11b).

However, there was a significant difference with regard to reaction times between the HFA group (*M* = 4,501 ms, *SD* = 1,703 ms) and the control group (*M* = 3,542 ms, *SD* = 1,148 ms), $t(36) = 2.04$, $p = .049$, $d = 0.66$, indicating that participants in the HFA group did take longer to complete the setting task.

Taken together, the results of the pilot study provide evidence that adults diagnosed with autism are impaired in the detection of direct gaze indexed by difficulties to distinguish subtle degrees of difference between averted and direct gaze. Especially gaze that is averted by 3°, 4°, 5°, and 6° was routinely mistaken to be direct by the participants in the autism group. However, there seems to be a threshold at around 7° at which both groups agree that gaze is averted. In line with theories about the aversiveness of social gaze for individuals with autism (e.g., Corden et al., 2008, participants in the autism group did perceive the detection task to be less pleasant and more difficult than the control group. As for the setting task, participants with autism were as accurate as participants in the control group at establishing direct gaze with the virtual character, indicating that the control over a stimulus might be an important mediating factor in the detection of direct gaze.

Experiment 2

Even though the pilot study did provide support for the hypotheses, it suffered from several limitations. First, there was no non-social control stimulus in either the detection or setting task to which the performance of both groups could have been compared. Thus, it is possible that the lower performance of the HFA group in the detection task is not limited to the social domain and does rather reflect a general processing deficit. To address this issue, a non-social control task was added to the design of the present study to assess the participant's performance in a non-social setting.

Second, whereas the detection task did use a realistic virtual character, the setting task did only use a reduced virtual character that consisted of the eye-region including the nose. Therefore, the fact that the present study did not find any difference in the performance on this task between the HFA group and the control group might be due to the reduced appearance of the virtual character. To exclude this alternative explanation, in the main experiment the same realistic virtual character was used for both the detection and the setting task.

Third, in the pilot study pleasantness and difficulty was only assessed for the detection task. In order to investigate whether control over the stimulus does affect the subjective experience of the task, pleasantness and difficulty were also measured for the setting task in the main experiment.

Methods

Participants. Thirty-seven adults with high-functioning autism (23 male, 14 female; M_{age} = 34.1, age range: 18–51 years) were recruited at the Adult Autism Outpatient Clinic of the Department of Psychiatry at the University Hospital of a large city in Western Germany. All participants with high-functioning autism were diagnosed by two independent physicians according to ICD-10 criteria. Thirty-nine control participants (23 male, 16 female; M_{age} = 31.1, age range: 24–45 years) were recruited at the Department of Psychology at the University of a large city in Western Germany. Both groups were matched for age, gender, and intelligence (measured with the German version of the Wechsler Adult Intelligence Scale-III (WIE – Wechsler Intelligenztest für Erwachsene; Von Aster, Neubauer, & Horn, 2006). Participants also completed the Beck Depression Inventory (BDI; Beck & Steer, 1987; Hautzinger, 1995), the Autism Spectrum Quotient (Baron-Cohen et al., 2006), the Empathy Quotient (Baron-Cohen & Wheelwright, 2004), and the Systemizing Quotient (Wheelwright et al., 2006).

On average, there was no difference in age between participants diagnosed with autism (M = 34.1, SD = 10.1) and participants in the control group (M = 31.1, SD = 4.3), $t(74) = 1.71$, $p = .092$, $d = 0.39$. See Table 5 for an overview of demographic and psychopathological variables for both groups.

Table 5

Demographic, Psychopathological, and IQ results for the Main Experiment

	HFA (n = 37)		Control (n = 39)					
Variable	*M*	*SD*	*M*	*SD*	*df*	*t*	*p*	Cohen's d
Age (years)	34.1	10.1	31.1	4.3	74	1.71	.092	0.39
WIE verbal IQ$_a$	112.4	17.0	113.8	12.9	72	-0.42	.679	-0.09
WIE performance IQ$_a$	103.5	20.4	107.2	13.2	72	-0.93	.357	-0.22
WIE IQ (total)$_a$	109.0	18.8	112.1	13.0	72	-0.82	.417	-0.19
BDI	11.7	10.7	6.0	5.1	74	2.98	.004	0.68
AQ	40.2	4.8	15.8	3.9	74	24.43	<.001	5.58
EQ$_b$	20.2	9.9	45.6	10.6	72	-10.63	<.001	-2.48
SQ$_b$	42.2	16.3	25.9	10.1	72	5.25	<.001	1.2

Note. WIE = Wechsler Intelligenztest für Erwachsene, BDI = Beck Depression Inventory, AQ = Autism Spectrum Quotient, EQ = Empathizing Quotient, SQ = Systemizing Quotient

[a]Two participants diagnosed with autism completed an older version of the intelligence test (WAIS-R, German version; Tewes, 1994) and were thus excluded from these analyses.

[b]Two participants diagnosed with autism did not complete the EQ and SQ and were thus excluded from the analyses.

Furthermore, there was no difference in intelligence between participants diagnosed with autism (M = 109.0, SD = 18.8) and participants in the control group (M = 112.1, SD = 13.0), $t(72) = -.82$, $p = .417$, $d = -0.19$.

Participants diagnosed with autism had higher BDI scores (Beck & Steer, 1987; Hautzinger, 1995) (M = 11.7, SD = 10.7) than participants in the control group (M = 6.0, SD = 5.1), $t(74) = 2.98$, $p = .004$, $d = 0.68$.

As expected, participants diagnosed with autism scored higher on the AQ (HFA: M = 40.2, SD = 4.8; Control: M = 15.8, SD = 3.9), $t(74) = 24.43$, $p < .001$, d = 5.58, lower on the EQ (HFA: M = 20.2, SD = 9.9; Control: M = 45.6, SD = 10.6), $t(72) = -10.63$, $p < .001$, $d = -2.48$, and higher on the SQ (HFA: M = 42.2, SD = 16.3; Control: M = 25.9, SD = 10.1), $t(72) = 5.25$, $p < .001$, $d = 1.2$.

Stimulus materials. In the detection task, participants' accuracy to detect when a realistic virtual character was gazing directly at them was tested. To assess whether participants were generally able to perform the cognitive processes to solve the task, they also completed a non-social control task with a geometric stimulus.

As the stimulus in the social condition, the same male virtual character from the pilot study (see Figure 9a) was used. The virtual character was moving its eyes and fixated at different positions. Fixations included direct gaze, slightly averted gaze, and clearly averted gaze. To reduce the time of the experiment, a shorter gaze sequence as opposed to the pilot study was used with 5 hits and 14 distractors. At the beginning of each trial, the virtual character showed clearly averted gaze (15° – 25° on the y-axis). After that, it moved his eyes to a fixation point of either 0° (hit) or 1-7°

(false alarm). The timing was the same as in the pilot study, with each trial lasting either 4,000 ms or 6,000 ms. Participants used a mouse to signal when the virtual character was showing direct gaze, clicking the left mouse button each time they felt the virtual character was gazing directly at them.

The non-social stimulus was a grey rectangle with a black cross (see Figure 9c). It had approximately the size of one of the virtual character's eyes (height = 1 cm, length = 2.5 cm). The cross was moving inside the rectangle and paused at different positions—directly in the center of the rectangle, slightly off-center, and clearly off-center. At the beginning of each trial, the cross started in a position that was clearly off-center (15° – 25° on the y-axis) and then moved to a point of either 0° (hit) or 1-7° (false alarm). The geometric stimulus used the same sequence of hits and distractors as the social stimulus. The timing was the same as in the social task. Again, participants used a mouse to signal when the cross was directly in the center of the rectangle. The design was a 2 (stimulus: social vs. non-social) × 2 (group: HFA vs. control) mixed-design, with stimulus as a within-subjects factor and group as a between-subjects factor. Both experimental tasks were presented in random order. Dependent variables were reaction times, the number of correct responses (hits) and the number of false alarms. In addition, participants rated the pleasantness (1 = *very unpleasant* to 5 = *very pleasant*) and difficulty (1 = *very easy* to 5 = *very difficult*) of every task on a 5-point scale.

In the setting task, participants' accuracy to actively establish direct gaze with a realistic virtual character was tested. Participants also completed a geometric control task to assess their ability to solve the task in a non-social setting. The setting task used the same stimuli as the detection task. However, instead of detecting when the virtual character was gazing at them, participants actively established direct gaze using the mouse. Both the social and the non-social condition of the setting task in the main experiment consisted of 21 trials. In each trial, the eyes of the virtual character appeared in a random position. Participants then used the mouse to move the eyes of the virtual character into a position that matched direct gaze as closely as possible. When the participants had established that position, they pressed a button on the mouse and the eyes of the virtual character moved to a new random position. In contrast to the pilot study, the rotation of the virtual character's head was not varied. There was no time limit for each trial to allow the participants to work as accurately as possible.

The task in the non-social condition was similar to the task in the social condition. However, instead of controlling the virtual character's eyes, participants tried to center the cross in the center of the rectangle. The design was a 2 (stimulus: social vs. non-social) × 2 (group: HFA vs. control) mixed-design, with stimulus as a within-subjects factor and group as a between-subjects factor. Both experimental tasks were presented in random order. Dependent variables were reaction times and the degree of deviation from a perfect eye contact. Additionally, participants evaluated the pleasantness (1 = *very unpleasant* to 5 = *very pleasant*) and difficulty (1 = *very easy* to 5 = *very difficult*) of every task on a 5-point scale.

Procedure. The participants diagnosed with autism were tested at the Department of Psychiatry at the University Hospital of a large city in Western Germany. The control participants were tested at the Department of Psychology of the University in the same city. Prior to the experiment, all participants gave consent to participate. All participants had normal or corrected-to-normal visual acuity. The experimental stimuli were presented on a 17-in. monitor with a resolution of 1280 x 1024 pixels. The experimental procedure was the same as in the pilot study. However, in addition to the detection and the setting task with the realistic virtual character, participants also completed both tasks with the non-social stimulus. Each participant completed all tasks in one single session. To eliminate carry-over-effects, all tasks were presented in random order.

Results and Discussion

Detection task. To compare both groups' ability to detect direct gaze, a sensitivity index *d'* was calculated out of the number of correct responses and the number of false alarms. Sensitivity *d'* was compared in a 2 × 2 mixed ANOVA (group × stimulus). There was a significant main effect of stimulus on *d'*, $F(1, 74) = 57.0$, $p < .001$, $\eta_p^2 = .435$. There was also a significant main effect of group on *d'*, $F(1, 74) = 5.13$, $p = .026$, $\eta_p^2 = .065$. There was no significant interaction effect between stimulus and group, $F(1, 74) = 0.80$, $p = .375$, $\eta_p^2 = .011$. Pairwise comparisons were performed comparing the differences in *d'* of the HFA and the control group in the social and non-social test conditions. These revealed a significant difference in the social condition, $t(74) = -2.27$, $p = .026$, $d = -0.51$. There was no significant difference in the non-social condition, $t(74) = -1.46$, $p = .149$, $d = -0.33$. This suggests that when confronted with the non-social stimulus participants diagnosed with autism ($M = 0.99$, $SD = 0.54$) were as sensitive at detecting when the geometric stimulus was centered as participants in the control group ($M = 1.15$, $SD = 0.41$). When dealing with the virtual character, however, participants in the HFA group ($M = 0.45$, $SD = 0.60$) were less sensitive at detecting direct gaze than participants in the control group ($M = 0.72$, $SD = 0.45$) (see Figure 13a).

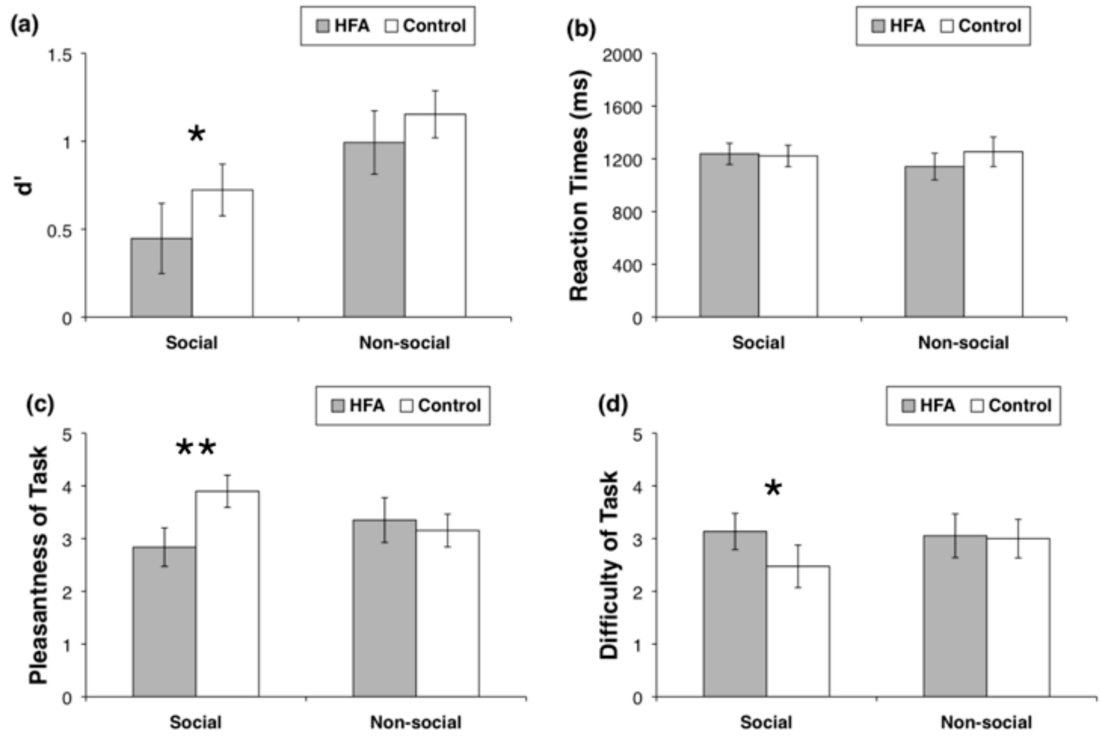

Figure 13. Overall results of the detection task in the main experiment: **(a)** false alarms, **(b)** reaction times, **(c)** pleasantness of the task, and **(d)** difficulty of the task. Error bars represent 95%-CI of the mean. * $p < .05$, ** $p < .001$.

Reaction times were analyzed in a 2 × 2 mixed ANOVA (group × stimulus). There was no significant main effect for stimulus, $F(1, 74) = 0.74$, $p = .392$, $\eta_p^2 = .010$, or group, $F(1, 74) = 0.77$, $p = .383$, $\eta_p^2 = .010$. There was also no significant interaction effect between stimulus and group, $F(1, 74) = 2.87$, $p = .095$, $\eta_p^2 = .037$. Pairwise comparisons revealed no difference between the HFA group (M = 1,238 ms, SD = 242 ms) and the control group (M = 1,222 ms, SD = 250 ms) in the social condition, $t(74) = 0.28$, $p = .777$, $d = 0.07$. There was also no difference between the HFA group (M = 1,142 ms, SD = 303 ms) and the control group (M = 1,253 ms, SD = 347 ms) in the non-social condition, $t(74) = -1.49$, $p = .139$, $d = -0.34$ (see Figure 13b).

Perceived pleasantness of the task was analyzed in a 2 × 2 mixed ANOVA (group × stimulus). There was no significant main effect for stimulus, $F(1, 74) = 0.67$, $p = .415$, $\eta_p^2 = .009$. There was a significant main effect for group, $F(1, 74) = 4.55$, $p = .036$, $\eta_p^2 = .058$. There was also a significant interaction effect between stimulus and group, $F(1, 74) = 20.06$, $p < .001$, $\eta_p^2 = .213$. Pairwise comparisons were performed to compare both groups in the social and non-social condition. Perceived pleasantness did differ significantly between the HFA group ($M = 2.84$, $SD = 1.09$) and the control group ($M = 3.90$, $SD = 0.94$) in the social condition, $t(74) = -4.54$, $p < .001$, $d = -1.04$. This indicates that individuals with autism did perceive the social task to be less pleasant than participants in the control group, whereas in the non-social condition there was no difference in perceived pleasantness of the task between the HFA group ($M = 3.35$, $SD = 1.27$) and the control group ($M = 3.15$, $SD = 0.96$), $t(74) = 0.77$, $p = .446$, $d = 0.18$ (see Figure 13c).

Perceived difficulty of the task was analyzed in a 2 × 2 mixed ANOVA (group × stimulus). There was no significant main effect for stimulus, $F(1, 73) = 1.74$, $p = .192$, $\eta_p^2 = .023$. There was also no significant main effect for group, $F(1, 73) = 2.98$, $p = .089$, $\eta_p^2 = .039$. The interaction between stimulus and group was not significant, $F(1, 73) = 3.23$, $p = .076$, $\eta_p^2 = .042$. Pairwise comparisons revealed a significant difference between the HFA group ($M = 3.14$, $SD = 1.03$) and the control group ($M = 2.47$, $SD = 1.22$) in the social condition, $t(73) = 2.53$, $p = .014$, $d = 0.59$, suggesting that participants diagnosed with autism did perceive the social task to be more difficult than participants in the control group. The HFA group ($M = 3.10$, $SD = 1.25$) and the control group ($M = 3.00$, $SD = 1.12$) did not differ in their perception of the difficulty of the non-social task, $t(73) = 0.20$, $p = .844$, $d = 0.08$ (see Figure 13d).

Setting task. Degrees of deviation from a perfect direct gaze were analyzed in a 2 × 2 mixed ANOVA (group × stimulus). There was a significant main effect for stimulus, $F(1, 74) = 246.45$, $p < .001$, $\eta_p^2 = .769$. There was no significant main effect for group, $F(1, 74) = 2.39$, $p = .126$, $\eta_p^2 = .031$. There was no significant interaction effect between stimulus and group, $F(1, 74) = 0.07$, $p = .745$, $\eta_p^2 = .001$. Pairwise comparisons revealed no significant difference between the HFA group ($M = 4.95$, $SD = 1.68$) and the control group ($M = 4.62$, $SD = 1.39$) in the social condition, $t(74) = 0.94$, $p = .353$, $d = 0.21$. There was also no significant difference between the HFA group ($M = 2.46$, $SD = 1.16$) and the control group ($M = 2.02$, $SD = 0.77$) in the non-social condition, $t(74) = 1.94$, $p = .056$, $d = 0.45$. This indicates that there was no difference in accuracy between participants diagnosed with autism and participants in the control group on either the social or non-social setting task (see Figure 14a).

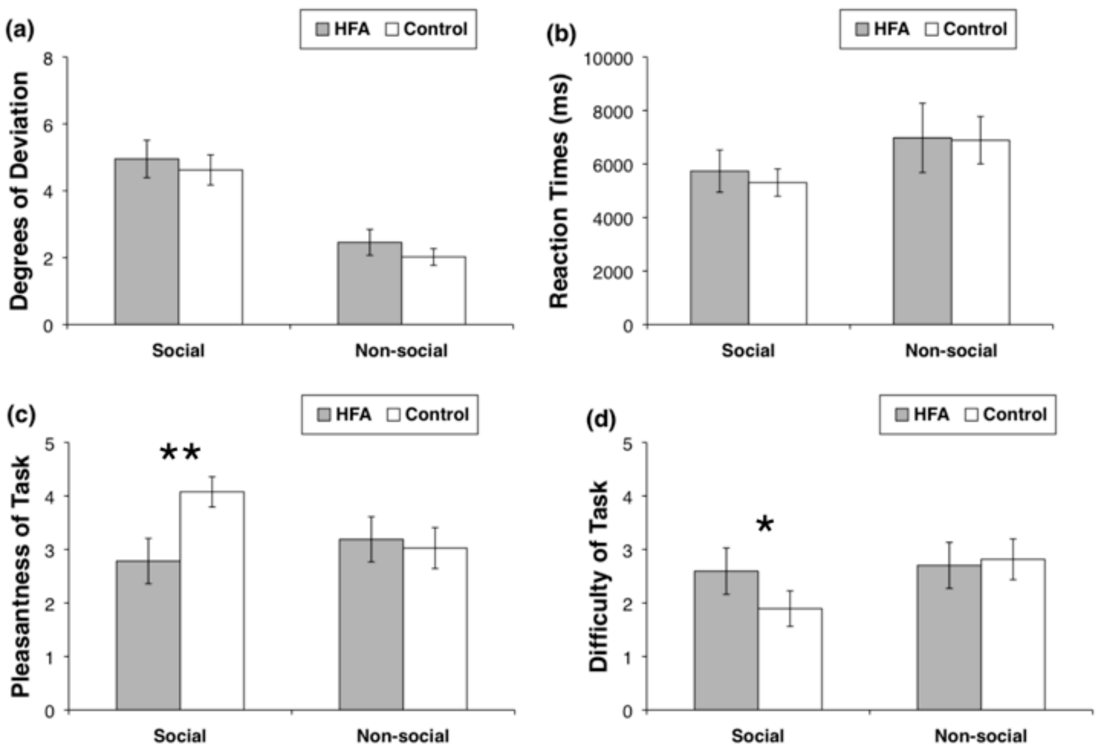

Figure 14. Overall results of the setting task in the main experiment: **(a)** degrees of deviation, **(b)** reaction times, **(c)** pleasantness of the task, and **(d)** difficulty of the task. Error bars represent 95%-CI of the mean. * $p < .05$, ** $p < .001$.

Reaction times were analyzed in a 2 × 2 mixed ANOVA (group × stimulus). There was a significant main effect for stimulus, $F(1, 74) = 24.86$, $p < .001$, $\eta_p^2 = .251$. There was no significant main effect for group, $F(1, 74) = 0.21$, $p = .648$, $\eta_p^2 = .003$. There was also no significant interaction effect between stimulus and group, $F(1, 74) = 0.36$, $p = .550$, $\eta_p^2 = .005$. Pairwise comparisons revealed no significant difference between the HFA group (M = 5,736 ms, SD = 2,363 ms) and the control group in the social condition (M = 5,308 ms, SD = 1,573 ms), $t(74) = 0.94$, $p = .353$, $d = 0.21$. There was also no significant difference between the HFA group (M = 6,978 ms, SD = 3,882 ms) and the control group (M = 6,889 ms, SD =2,732 ms) in the non-social condition, $t(74) = 0.12$, $p = .908$, $d = 0.03$. This indicates that both groups needed a comparable amount of time to complete the task, and that there was no speed-accuracy trade-off in either of the groups (see Figure 14b)

Perceived pleasantness of the task was analyzed in a 2 × 2 mixed ANOVA (group × stimulus). There was no significant main effect for stimulus, $F(1, 74) = 3.47$, $p = .066$, $\eta_p^2 = .045$. There was a significant main effect for group, $F(1, 74) = 7.91$, $p = .006$, $\eta_p^2 = .097$. There was also a significant interaction effect between stimulus and group, $F(1, 74) = 17.66$, $p < .001$, $\eta_p^2 = .193$. Pairwise comparisons were performed to compare both groups in the social and non-social condition. Perceived pleasantness did differ significantly between the HFA group ($M = 2.78$, $SD = 1.27$) and the control group ($M = 4.08$, $SD = 0.87$) in the social condition, $t(74) = -5.20$, $p < .001$, $d = -1.19$, indicating that individuals diagnosed with autism did perceive the social task to be less pleasant than participants in the control group. There was no difference in perceived pleasantness of the task between the HFA group ($M = 3.19$, $SD = 1.27$) and the control group ($M = 3.03$, $SD = 1.18$) in the non-social condition, $t(74) = 0.58$, $p = .562$, $d = 0.13$ (see Figure 14c).

Perceived difficulty of the task was analyzed in a 2 × 2 mixed ANOVA (group × stimulus). There was a significant main effect for stimulus, $F(1, 73) = 9.27$, $p = .003$, $\eta_p^2 = .113$. There was no significant main effect for group, $F(1, 73) = 1.82$, $p = .182$, $\eta_p^2 = .024$. However, there was a significant interaction effect between stimulus and group, $F(1, 73) = 5.79$, $p = .019$, $\eta_p^2 = .073$. Pairwise comparisons revealed a significant difference between the HFA group ($M = 2.59$, $SD = 1.3$) and the control group ($M = 1.89$, $SD = 1.01$) in the social condition, $t(73) = 2.61$, $p = .011$, $d = 0.60$, meaning that participants diagnosed with autism did perceive the social task to be more difficult than participants in the control group. There was no difference between the HFA group ($M = 2.70$, $SD = 1.29$) and the control group ($M = 2.82$, $SD = 1.16$) in the perception of the difficulty of the task in the non-social condition, $t(73) = -0.40$, $p = .690$, $d = -0.10$ (see Figure 14d).

General Discussion

The main goal of the current experiments was to investigate whether adults diagnosed with HFA are impaired in their ability to detect direct gaze. In addition, it was also tested whether these gaze processing impairments could—at least in part—be due to the uncontrollability of the social stimuli by giving participants control over the gaze of a realistic virtual character.

The results of the detection task with the virtual character in both Experiment 1 and Experiment 2 show that adult individuals with high-functioning autism are impaired in their ability to distinguish between direct and averted gaze. In the current study, this difficulty to distinguish between direct and averted gaze manifested itself in an increased rate of false alarms, which may suggest that when faced with ambiguous social gaze individuals with autism overcompensate by classifying gaze as direct that is in fact averted.

The results of the present study indicate that this finding may not be explained by a general deficit in the processing of moving stimuli because individuals with autism performed as well as participants in the control group on the detection task with the geometric, non-social stimulus in the main experiment. However, it is important to note that the present study did only find significant main effects for group and stimulus and no significant interaction between stimulus and group. Therefore, it is possible that with a larger sample size the present study may have also detected a significant difference between both groups in the non-social condition, which would hint at a deficit in the processing of non-social stimuli. However, this effect in the non-social condition was only small compared to the effect in the social condition and may not explain the impaired performance of individuals with autism in the social condition. Interestingly, both groups performed better on the non-social than on the social task, which suggests that the social condition was more demanding for both individuals with autism and participants in the control group.

These results are in contrast to previous studies (Senju et al., 2005; Senju et al., 2008; Webster & Potter, 2011), which have not shown clear deficits in the detection of direct gaze. This might be in part due to the different methodologies used by the other studies. Whereas Senju et al. (2008) used a visual search task with static stimuli showing only two different kinds of gaze (direct or averted), the present study used a realistic virtual character showing dynamic gaze behavior with fine-tuned variations. Even though Webster and Potter (2011) did use different degrees of averted gaze (20°, 10°, and 5°), they presented their participants with two images side by side, one displaying direct gaze and the other displaying averted gaze, which might have made the task considerably easier than the task used in the present study. The present findings support the idea that the study of gaze behavior in autism can benefit from stimuli that are both dynamic and realistic (Gepner, Deruelle, & Grynfeltt, 2001; Wilms et al., 2010).

Now that it has been established that individuals with autism are impaired in the detection of direct gaze, it is important to determine the processes that lead to less accurate gaze detection. At least two explanations seem possible:

First, the gaze behavior of the virtual character was highly unpredictable, which might have interfered with the autistic participants' ability to detect direct gaze. As has been stated by the empathizing-systemizing theory (Baron-Cohen, Richler, Bisarya, Gurunathan, & Wheelwright, 2003), individuals with autism prefer situations that can be predicted and controlled. Previous studies have shown that children with autism favor predictable environments over unpredictable environments (Ferrara & Hill, 1980; Klin et al., 2009) and show more problematic behaviors, such as self-hits and aggression, in unpredictable environments (Flannery & Horner, 1994). During the detection task in the present study, participants were confronted with the possibility of being gazed at by a virtual character. This possibility of becoming engaged in social interaction may have been aversive for the individuals with autism and could have interfered with their ability to accurately detect direct gaze. This would explain why the individuals with autism did not differ from the typically developing individuals in the setting task, where the virtual character did not show unpredictable gaze behavior and was instead controlled by the participants.

Second, the individuals with autism may have averted their gaze from the virtual character and were therefore not able to accurately detect when the virtual character was gazing at them. As has been shown in previous studies, when confronted with social stimuli, individuals with autism avert their gaze to avoid looking at the stimulus (Corden et al., 2008; Hutt & Ounsted, 1966; Richer & Coss, 1976). Thus, it might be possible that the individuals with autism did not attend to the eye region in the detection task and therefore did not accurately detect direct gaze. In the setting task, individuals with autism performed as accurate as participants in the control group. Therefore, the predictability of the situation may have influenced the gaze behavior of the participants with autism causing them to show less gaze aversion. However, without eye-tracking data it can only be speculated about the relation between gaze aversion, task performance, and predictability. Thus, a replication of the present study should include eye-tracking data to determine which regions of the face individuals with autism fixate on during the different tasks.

As indicated by the results of the questionnaire data in the main experiment, individuals with autism experienced the detection task with the virtual character as less pleasant than participants in the control group. These results are in concordance with previous studies showing that social stimuli, such as faces, are less pleasant for individuals with autism and that they prefer non-social stimuli (Corden et al., 2008; Hutt & Ounsted, 1966; Richer & Coss, 1976).

As for the reaction times of the detection task in both the pilot study and the main experiment, individuals with autism did not differ from control persons in reaction times in any of the detection tasks, suggesting that there was no trade-off between speed and accuracy in any of the two groups.

The results of the setting task with the virtual character in both the pilot study and the main experiment show that individuals with autism are in fact able to establish direct gaze accurately when they have full control over a social stimulus. When looking at the degrees of deviation from direct gaze in the main experiment, it is important to note that in the setting task individuals with autism were able to distinguish more accurately between direct or averted gaze than in the detection task. Whereas in the setting task individuals with autism perceived gaze between 0° and 4.95° to be direct, in the detection task individuals with autism indicated gaze to be direct that was more than 5° averted. Why were the individuals with autism more accurate at detecting direct gaze in the setting task as opposed to the detection task? First, the setting task was more predictable. Whereas in the detection task the virtual character showed unpredictable gaze behavior, the gaze in the setting task was controlled by the participants, making it highly predictable. If the unpredictability of the gaze behavior interfered with the autistic participants' ability to detect direct gaze in the detection task, then in the setting task participants with autism could focus more easily on the gaze of the virtual character.

Second, one important difference between both tasks is that there was no time limit in the setting task. In the setting task in the main experiment, participants in the HFA group took about 5.7 s on average to establish direct gaze with the virtual character. In the detection task, on the other hand, the eyes of the virtual character were moving, so the time to make a decision about the virtual character's gaze was limited. This was also reflected in participants' reaction times, with the average reaction time in the detection task being about 1.2 s. In the pilot study, individuals with autism also did take significantly longer to complete the setting task than participants in the control group. However, in the main experiment there was no difference in reaction times between both groups in the setting task. Thus, another important aspect for a future study might be to introduce also a time limit for the setting task to investigate whether individuals with autism can still benefit from control over the social stimulus if the time is limited.

Furthermore, even though individuals with autism were as accurate as participants in the control group at establishing direct gaze in the setting task, it is not clear which strategy they used. Previous studies have argued that individuals with autism use a geometric strategy to process gaze (Ristic et al., 2005). Because the setting task in the main study used a virtual character directly facing the participants, it is possible that the individuals with autism used geometric cues, such as the symmetry of the pupils, to solve the task. To further investigate this possibility, a future study might also use a virtual character with a slightly averted head position, a technique that has been used by other studies to interfere with geometric processing of gaze (Senju et al., 2008; Ashwin, Ashwin, Rhydderch, Howells, & Baron-Cohen, 2009).

With regard to the subjective experience during the tasks, individuals with autism did perceive both the setting and the detection task with the realistic virtual character to be less pleasant than participants in the control group. Thus, the controllability of the setting task did not cause individuals with autism to rate it as more pleasant. In addition, individuals with autism also did perceive both tasks as more difficult than participants in the control group. Taken together, both results indicate that the perceived pleasantness and difficulty of the tasks was more influenced by the social stimulus per se than by the characteristics of the tasks, suggesting a strong influence of the social nature of the task on experience.

An important question for future studies is whether individuals with autism might be able to generalize skills gained in the setting task to other situations. A simple test would be to use the same paradigm as in the present study, but to ask participants to complete the setting task before the detection task. It might be possible that in this case participants with autism would be more accurate at detecting direct gaze. In the present study, the order of tasks was randomized so such training effects could not be investigated.

On the whole, the results of the setting task show that individuals with autism might be able to improve their ability to detect direct gaze by giving them control over or providing a handle for social stimuli. Previous studies have begun to focus on developing training programs for individuals with autism: For instance, Golan et al. (2009) created an animated series of tank engines with faces to teach children with autism to recognize emotions. Children with autism that had watched the series for four weeks significantly improved in their ability to detect emotions compared to those who did not watch the series. In another study, Faja, Aylward, Bernier, and Dawson (2008) used a face training program to help autistic children improve their ability to recognize faces. After eight training sessions, the ability to detect faces improved. However, both studies did not make use of controllable stimuli. The results suggest that presenting individuals with autism with a virtual character that can be made responsive and be controlled might also prove useful in the teaching of social skills (Wilms et al., 2010).

As for further clinical implications of the present study, the results of the detection task show a tendency by the participants diagnosed with autism to classify averted gaze as direct. This tendency may be a source for misunderstandings in social interactions because some communicative actions, such as turn-taking, rely on accurately interpreting the gaze cues of the interaction partner. If individuals with autism mistake averted gaze for direct gaze, they may misinterpret that direct gaze for a turn-taking signal and interrupt their interaction partner. Because there is some evidence that turn-taking in autism may be abnormal (Baron-Cohen, 1988), based on the present results the role of gaze in turn-taking in autism needs to be further investigated.

Several limitations apply to the present study. First, the study focused only on individuals with high-functioning autism. Even though it might be concluded that non-high-functioning individuals with autism are also impaired in the detection of direct gaze, this has to be investigated in a separate study. Furthermore, it is not clear if non-high-functioning individuals with autism would be similarly able to use a task that gives them full control over a social stimulus to train their ability to detect direct gaze. Second, the present study focused only on adults. As other authors have mentioned (Nation & Penny, 2008), a developmental perspective is important for the understanding of autism to establish at which stage of development a certain deficit occurs. Third, the non-social, geometric stimulus was less visually complex than the virtual character. Therefore, a replication of the study should use a non-social stimulus that is more visually complex to better compare the performance on the social and non-social task.

Taken together, the current study investigated whether individuals with autism are impaired in their ability to detect direct gaze. In addition, it was tested whether individuals with autism can overcome such an impairment when given the control over the gaze of a naturalistic virtual character. The results suggest that individuals with autism are impaired in their ability to distinguish between direct and averted gaze. This underlines the importance of using animate, realistic stimuli in experiments with individuals with autism, because the impairments specific to autism might not be fully detected using only static stimuli (Schilbach et al., 2011). Furthermore, when given the control over the gaze of a realistic virtual character, individuals with autism were as accurate as participants in the control group at establishing direct gaze, which suggests that giving individuals with autism control over a virtual character, or social stimuli in general, may be used to help them overcome some of their deficits in the processing of gaze.

Study 4: Perception of Attractiveness in Autism

Introduction

Attractiveness has a strong influence on our perception and judgment. We subscribe more positive attributes to people with attractive faces (Eagly, Ashmore, Makhijani, & Longo, 1991), treat them more positively (Langlois et al., 2000), and value objects more that are paired with attractive faces (Strick, Holland, & van Knippenberg, 2008). Furthermore, attractiveness plays an important role in sexual selection and mate choices (Grammer & Thornhill, 1994).

Also, judgments of attractiveness and gaze are strongly linked (Shimojo, Simion, Shimojo, & Scheier, 2003). We tend to spend more time looking at attractive faces (Maner et al., 2003), and when doing so reward centers in the brain are activated, such as the nucleus accumbens (NAcc) and the orbitofrontal cortex (OFC) (Aharon et al., 2001; O'Doherty et al., 2003; Cloutier, Heatherton, Whalen, & Kelley, 2008; Liang, Zebrowitz, & Zhang, 2010; Winston, O'Doherty, Kilner, Perrett, & Dolan, 2007). Even human infants prefer to look more at attractive faces than at unattractive faces (Langlois et al., 1987).

Because adults with autism spectrum disorders show some deficits in the processing of faces (e.g., Dawson, Webb, & McPartland, 2005; Jemel, Mottron, & Dawson, 2006) and in some cases atypical gaze to faces (e.g., Dalton et al., 2005), it is possible that the ability to detect attractiveness may be impaired in individuals with autism. The present study therefore investigated whether the perception of attractiveness is impaired in autism by having high-functioning adults with autism choose the more attractive face out of pairs of faces while their gaze behavior was being recorded. Additionally, it was investigated whether a systematic gaze bias towards preferred faces—the gaze cascade effect (Shimojo et al., 2003)—could be replicated in individuals with autism.

Perception of Attractiveness in Autism

Two studies, so far, have tested the ability of individuals with autism to detect attractiveness: Da Fonseca et al. (2011) investigated whether children diagnosed with autism perceive the same faces as attractive as typically developing children and whether they also subscribe more positive attributes to those faces ("beauty is good"-stereotype; Eagly, Ashmore, Makhijani, & Longo, 1991). Their results support the notion that the perception of attractiveness is preserved in autism. Children with autism did not differ in their ratings of the attractive faces from typically developing children, and they also rated the attractive faces to be more positive than the unattractive faces, showing that children with autism also use the "beauty is good"-stereotype in their judgments.

However, contrary to the findings by da Fonseca et al. (2011), White et al. (2006) did find evidence for impaired perception of attractiveness in adults diagnosed with Asperger Syndrome. In their study, adults with Asperger Syndrome and controls rated the attractiveness of male and female faces. The results showed that participants with Asperger Syndrome were impaired in their ability to judge the attractiveness of faces but only of their own gender.

Given these contradictory findings, it is important to further investigate whether high-functioning adults with autism are impaired in their ability to detect attractiveness.

Visual Attention to Faces in Autism

One possible explanation for an impaired perception of attractiveness in autism might be that individuals with autism differ in their visual attention to parts of the face that are important for judging attractiveness, and therefore come to different judgments than typically developing individuals. There is a large body of studies that has investigated the visual attention of individuals with autism to different facial features, such as the eyes, nose, and mouth (for a review see Falck-Ytter & von Hofsten, 2011). Even though there are some studies that have found differences between individuals with autism and controls (e.g., Jones, Carr, & Klin, 2008; Dalton et al., 2005), there are also studies that have found no differences (e.g., McPartland, Webb, Keehn, & Dawson, 2011; Georgescu et al., 2013; Rutherford & Towns, 2008). For instance, Louwerse et al. (2013) found no difference in the time spent fixating at the eye-region between individuals with autism and controls. It therefore remains an open question whether visual attention to different facial features differs in high-functioning individuals with autism from typically developing adults and whether these differences might help to explain potential deficits in the detection of attractiveness.

Gaze Cascade Effect

The gaze cascade model was proposed by Shimojo et al. (2003) to explain how human gaze and preferences are linked. According to the model, our gaze does not only reflect our preferences ("I am gazing at what I prefer.") but also influences *what* we prefer ("I prefer what I am gazing at."). In their original study, Shimojo et al. (2003) presented their participants with pairs of faces and asked them to choose the more attractive face. At the end of each trial, they found an increased bias to gaze at the chosen face. Shimojo et al. (2003) proposed two processes that might cause this gaze bias: mere exposure (Zajonc, 2001) and preferential looking (Birch, Shimojo, & Held, 1985). According to the mere exposure effect, stimuli are preferred simply because we are exposed to them. Preferential looking describes the phenomenon that humans increase exposure to preferred stimuli. According to Shimojo et al. (2003) these two processes work together in a positive feedback-loop to produce the gaze cascade effect: At the beginning of each trial there is no preference for any of the two faces, and both faces are each likely to be gazed at. However, due to random factors one of the two faces will be gazed at more. Due to the mere exposure effect, the valence of this face will increase. At which point, preferential looking will lead to more gaze to this face, which will then again increase the valence of this face. This feedback-loop will continue up to the point where a conscious decision is made.

According to the gaze cascade model, the strongest gaze bias will occur if the two stimuli are very similar in attractiveness. Under these circumstances, gaze can play a larger role in forming the decision. If both stimuli differ in attractiveness, the influence of gaze behavior on the decision will be weaker because the decision will be more strongly influenced by features inherent in the stimulus. To test this aspect, Shimojo et al. (2003) presented their participants with three different tasks: In the first task, the two faces in each pair had a large difference in attractiveness, with one face being very attractive and the other face being very unattractive. In the second task, the two faces were of equal attractiveness. In the third task, Shimojo et al. (2003) presented participants with abstract shapes that were supposed to be equally attractive. In line with the gaze cascade model, the strongest gaze bias was found in task two and three, in which both stimuli in each pair were similar in attractiveness. To further support the idea that gaze behavior actively influences the decision process, Shimojo et al. (2003) actively manipulated participants gaze by alternatively flashing the two faces and increasing the duration of one face and decreasing the duration of the other face—thus, mimicking participants' gaze behavior found in the previous study. The results showed that participants were more likely to choose the face to which their gaze had been guided. To rule out the possibility that mere exposure alone accounted for the results and that the decision was not driven by the gaze shifts of the eyes, participants in one control condition simply focused on a fixation cross between both faces. In this control condition without gaze shifts, participants did not show a bias to choose the longer presented stimulus.

To rule out the possibility that participants simply gazed at the chosen alternative, Shimojo et al. (2003) asked participants in two control conditions to choose the rounder face or the more unattractive face. In these control conditions, the likelihood to gaze at the chosen face before the decision was much lower (only around 55–65%). These results show that there is a slight tendency to gaze at the chosen alternative; however, the likelihood to gaze at the chosen face in the preference condition was much higher (around 80%), showing that in preference judgments gaze and choice are more strongly linked and this link cannot be explained by a simple tendency to gaze at the chosen stimulus.

To further support the gaze cascade effect, Simion and Shimojo (2007) again presented participants with two faces; however, in some trials the faces disappeared before participants could make a choice. The gaze behavior of the participants showed a bias toward the chosen face—even if that face was not visible anymore, further supporting the link between gaze and decision.

However, the gaze cascade effect has also been criticized. For instance, even though Glaholt and Reingold (2009) were able to replicate the gaze cascade effect, they were not able to actively produce a gaze bias towards certain stimuli by increasing their exposure. These results are in line with a study by Nittono and Wada (2009) who also replicated the gaze cascade effect but were not able to actively produce a gaze bias in participants towards certain stimuli. However, both studies

(Glaholt & Reingold, 2009; Nittono & Wada, 2009) did differ from the original study by Shimojo et al. (2003) in their use of stimuli. Whereas Shimojo et al. (2003) used faces, the studies by Glaholt and Reingold (2009) and Nittono and Wada (2009) used abstract shapes or photographs. Therefore, Nittono and Wada (2009) argued that the gaze cascade effect may be limited to faces.

Now that the gaze cascade effect has been established as a phenomenon in preference decisions, the question remains whether it can also be found in individuals with autism. However, for the gaze cascade effect to work, some form of self-monitoring of one's own gaze—conscious or unconscious—is needed. At least one study hints at the fact that individuals with autism may be impaired in their sense of agency with regard to their own gaze. In a study using a gaze-contingent lense, Grynszpan et al. (2012) could show that individuals with autism had difficulty at adapting their gaze when they controlled a gaze-contingent lense. In addition, only one participant in the HFA group noticed that the lense could be controlled with the gaze compared to 50% in the control group. Based on these findings Grynszpan et al. (2012) concluded that self-monitoring of gaze may be impaired in individuals with autism. Thus, if the self-monitoring of gaze in high-functioning autism is impaired, it is questionable if the gaze behavior can have an impact on decision making as proposed by the gaze cascade model.

Up until now, one study has investigated the gaze cascade effect in individuals with autism. Using the original paradigm by Shimojo et al. (2003), Gharib et al. (2011) were able to replicate the gaze cascade effect in a sample of individuals with autism. Additionally, Gharib et al. (2011) could show an increased bias towards the chosen face in the participants with autism during the last 40 ms before the decision. This study, however, suffered from two major shortcomings: First, the sample size was very small (HFA: 4, Control: 3), which may have resulted in an increased influence of outliers on the results. Second, the stimuli in the study were not validated in a prestudy, so it was not possible to test whether individuals with autism are generally impaired in the perception of attractiveness. Both shortcomings were addressed in the present study by using a larger sample size and by validating the stimuli in two prestudies.

In sum, the gaze cascade effect describes the interplay between mere exposure and preferential looking that leads to an increased gaze bias towards one of two faces in a two-alternative forced choice preference decision. In this study, the goal was to replicate this effect in individuals with high-functioning autism.

To summarize, the aim of the present study was twofold: First, it was investigated whether individuals with autism are impaired in their ability to detect attractiveness in faces. Second, the present study examined whether the findings on the relation between gaze and preference decisions found in typically developing individuals (e.g., Shimojo et al., 2003; Glaholt & Reingold, 2011) could be replicated in individuals with autism.

Methods

Participants. Thirty adults with high-functioning autism (20 male, 10 female; M_{age} = 43.3, age range: 23–55 years) were recruited at the Adult Autism Outpatient Clinic of the Department of Psychiatry at the University Hospital of a large city in Western Germany. All participants with high-functioning autism were diagnosed by two independent physicians according to ICD-10 criteria. Thirty control participants (20 male, 10 female; M_{age} = 41.2, age range: 33–63 years) were recruited at the Department of Psychology at the University of a large city in Western Germany. Both groups were matched for age, gender, and intelligence (measured with the German version of the Wechsler Adult Intelligence Scale-III (WIE – Wechsler Intelligenztest für Erwachsene; Von Aster et al., 2006). Participants also completed the Beck Depression Inventory (BDI; Beck & Steer, 1987; Hautzinger, 1995), the Autism Spectrum Quotient (Baron-Cohen et al., 2006), the Empathy Quotient (Baron-Cohen & Wheelwright, 2004), and the Systemizing Quotient (Wheelwright et al., 2006). The study protocol had been approved by the local ethics committee.

On average, there was no difference in age between participants diagnosed with autism (M = 43.3, SD = 8.1) and participants in the control group (M = 41.2, SD = 7.7), $t(58) = 1.03$, $p = .307$., $d = 0.27$. See Table 6 for an overview of demographic and psychopathological variables for both groups.

Table 6

Demographic, Psychopathological, and IQ results for the Main Experiment

Variable	HFA (n = 30) *M*	HFA (n = 30) *SD*	Control (n = 30) *M*	Control (n = 30) *SD*	*df*	*t*	*p*	Cohen's d
Age (years)	43.3	8.1	41.2	7.7	58	1.03	.307	0.27
WIE verbal IQ_a	117.1	16.3	110.7	12.3	55	1.70	.096	0.44
WIE performance IQ_a	107.3	18.1	104.3	16.0	55	0.66	.515	0.18
WIE IQ $(total)_a$	113.9	17.5	108.4	13.3	55	1.33	.191	0.35
BDI_b	15.1	10.9	7.8	5.9	57	3.21	.002	0.83
AQ	41.7	4.3	17.5	6.2	58	17.68	<.001	4.53
EQ_c	15.6	8.3	41.9	12.1	55	-9.50	<.001	-2.53
SQ_c	41.3	15.8	26.2	10.0	55	4.35	<.001	1.14

Note. WIE = Wechsler Intelligenztest für Erwachsene, BDI = Beck Depression Inventory, AQ = Autism Spectrum Quotient, EQ = Empathizing Quotient, SQ = Systemizing Quotient

[a]Three participants diagnosed with autism completed an older version of the intelligence test (WAIS-R, German version; Tewes, 1994) and were thus excluded from these analyses.

[b]One participant in the control group did not complete the BDI and was thus excluded from the analyses.

[c]Three participants diagnosed with autism did not complete the EQ and SQ and were thus excluded from the analyses.

Furthermore, there was no difference in intelligence between participants diagnosed with autism (M = 113.9, SD = 17.5) and participants in the control group (M = 108.4, SD = 13.3), $t(55) = 1.33$, $p = .191$, $d = 0.35$.

Participants diagnosed with autism had higher BDI scores (Beck & Steer, 1987; Hautzinger, 1995) (M = 15.1, SD = 10.9) than participants in the control group (M = 7.8, SD = 5.9), $t(57) = 3.21$, $p = .002$, $d = 0.83$.

As expected, participants diagnosed with autism scored higher on the AQ (HFA: M = 41.7, SD = 4.3; Control: M = 17.5, SD = 6.2), $t(58) = 17.68$, $p < .001$, d = 4.53, lower on the EQ (HFA: M = 15.6, SD = 8.3; Control: M = 41.9, SD = 12.1), $t(55) = -9.50$, $p < .001$., $d = -2.53$, and higher on the SQ (HFA: M = 41.3, SD = 15.8; Control: M = 26.2, SD = 10.0), $t(55) = 4.35$, $p < .001$, $d = 1.14$.

Stimulus materials and design.

Experimental task. Experiments on the relation between gaze and preference generally use a forced-choice paradigm (e.g., Shimojo et al., 2003; Glaholt & Reingold, 2011): Participants are presented with two or more stimuli from the same category (e.g., faces, shapes, or objects) and have to choose the preferred one while their gaze behavior is being recorded. Viewing time for each pair of stimuli is usually not restricted. The present study used the same basic paradigm by Shimojo et al. (2003) with three different tasks.

The goal of the first task (face-attractiveness-detection task) was to investigate participants' ability to detect attractiveness by presenting them with pairs of faces, consisting of one attractive and one unattractive face, which had been selected based on a prestudy. Participants' task was to select the more attractive face.

The goal of the second task (face-gaze-preference task) was to examine the relation between participants' gaze behavior and their preference choices. As Shimojo et al. (2003) could show, gaze has a stronger influence on decision making when two stimuli are similar in attractiveness because in this case the decision is not so strongly influenced by features of the stimulus. In this task, participants were presented with pairs of faces, consisting of two faces that were similar with regard to attractiveness. Again, participants' task was to select the more attractive face.

The goal of the third task (shape-gaze-preference task) was also to examine the relation between participants' gaze behavior and their preference choices. However, instead of faces participants were presented with pairs of abstract shapes that were similar with regard to attractiveness. This was done for two reasons: First, several studies have shown that cognitive impairments in autism are primarily found with social stimuli (e.g., Dichter & Belger, 2007). It is therefore important to include a non-social control condition to investigate whether impairments are limited to social stimuli or even extend to non-social domains. Second, in their original study, Shimojo et al. (2003) found the strongest gaze cascade effect in the non-social condition and it is thus important to check if this finding can be replicated.

To summarize, the basic structure of the experiment was as follows: In each of the three tasks, participants were presented with 20 stimulus pairs (faces or shapes) and their task was to choose the more attractive stimulus. In each trial, participants' choice, reaction time, and gaze behavior were recorded. The position of each stimulus in each trial was randomized. Additionally, the order of all three tasks was also randomized.

If the detection of attractiveness is impaired in individuals with autism, a task demanding participants to detect attractiveness might be perceived as difficult. Therefore, participants were also asked to rate how easy it was for them to judge the attractiveness of the faces/shapes on a 7-point scale (1 = *very easy* to 7 = *very difficult*).

Face stimuli. For the face-attractiveness-detection task the goal was to create 20 pairs of avatars that had a large difference in attractiveness, with one avatar in each pair being more attractive than the other avatar. For the face-gaze-preference task the goal was to create 20 pairs of avatars that had a small difference in attractiveness, with both avatars in each pair being about equally attractive.

To select appropriate stimuli for both the face-attractiveness-detection task and the face-gaze-preference task, 52 male and 59 female avatars were created using the software FaceGen (Singular Inversions, 2011).

To gather attractiveness ratings for the facial stimuli, a prestudy with 52 participants (12 male, 41 female; M_{age} = 24.1; SD_{age} = 3.1) was conducted. In the prestudy, each participant was presented with half of the avatars and rated the attractiveness of each avatar on a scale ranging from 1 (*very unattractive*) to 7 (*very attractive*). Based on the mean attractiveness ratings, avatars were paired together. Avatars in each pair were matched for gender, so that each pair included either two male or two female avatars. Two separate sets of avatar pairs were created. For the face-attractiveness-detection task, avatars were paired so that there was a large difference in the attractiveness ratings between the two avatars in each pair. For the face-gaze-preference task, avatars were paired so that both avatars were very similar in their attractiveness ratings.

To further validate the stimulus selection, a second prestudy with 103 participants (21 male, 82 female; M_{age} = 23.9; SD_{age} = 3.4) was conducted. In this prestudy, participants were presented with the 53 pairs of avatars that had been created based on the results of the first prestudy. Participants' task was to select the more attractive avatar in each pair. For each pair, the percentage that the more attractive avatar was chosen was calculated. For the face-attractiveness-detection task, the 10 male and 10 female pairs of avatars in which the attractive avatar was chosen more often than the unattractive avatar were selected. In the 20 final pairs (10 male, 10 female) selected for the study, the more attractive avatar was chosen on average 80.75 percent of the time. With regard to the attractiveness ratings, the average difference between the attractive and the unattractive avatar in each of the pairs was 1.07. In addition, the attractive avatars had significantly higher attractiveness ratings than the unattractive avatars, $t(38) = 4.87$, $p < .001$, $d = 1.53$.

For the face-gaze-preference task, 10 male and 10 female pairs of avatars in which the attractive avatar was not chosen more often than the unattractive avatar were selected. In the 20 final pairs (10 male, 10 female), the more attractive avatar was chosen on average only 48.2 percent of the time. With regard to the attractiveness ratings, the average difference between the attractive and the unattractive avatar in each of the pairs was 0.25. In addition, the attractive avatars had only slightly higher attractiveness ratings than the unattractive avatars, $t(37) = 2.03$, $p = .049$, $d = 0.64$.

In sum, both sets of stimuli matched the selection criteria. The 20 pairs for the face-attractiveness-detection task had a large difference in attractiveness, with one avatar in each pair being more attractive than the other avatar. The 20 pairs for the face-gaze-preference task had a small difference in attractiveness, with both avatars in each pair being about equally attractive. Figure 15 shows one pair of faces from the face-attractiveness detection task.

Figure 15. Pair of faces from the face-attractiveness-detection task. The more attractive face is on the right.

Shape stimuli. For the shape-gaze-preference task the goal was to create 20 pairs of abstract shapes that had a small difference in attractiveness, with both shapes in each pair being about equally attractive.

To select appropriate stimuli for the shape-gaze-preference task, 40 different shape stimuli were created using the algorhythm by Gielis et al. (2003), which has been used in several psychophysiological experiments (e.g., Suchow & Alvarez, 2011). The goal was to create 20 pairs of shapes in which there would be no clear preference for each one of the two shapes. Therefore, the forty shape stimuli were randomly paired to create 20 pairs of two shapes. The twenty pairs were also included in the second prestudy with the facial pairs. Thus, 103 participants (21 male, 82 female; M_{age} = 23.9; SD_{age} = 3.4) were presented with the pairs and chose the more attractive shape. Unlike for the facial stimuli, there were no attractiveness ratings for the shapes that could have been used to determine a more attractive shape in each pair. Therefore, the percentage for each pair that the left shape was chosen was calculated. The left shape was chosen on average 51.75 percent of the time. Thus, there was no preference to select one shape over the other in each pair, and the shape stimuli also matched the selection criteria. Figure 16 shows one pair of shapes from the shape-gaze preference task.

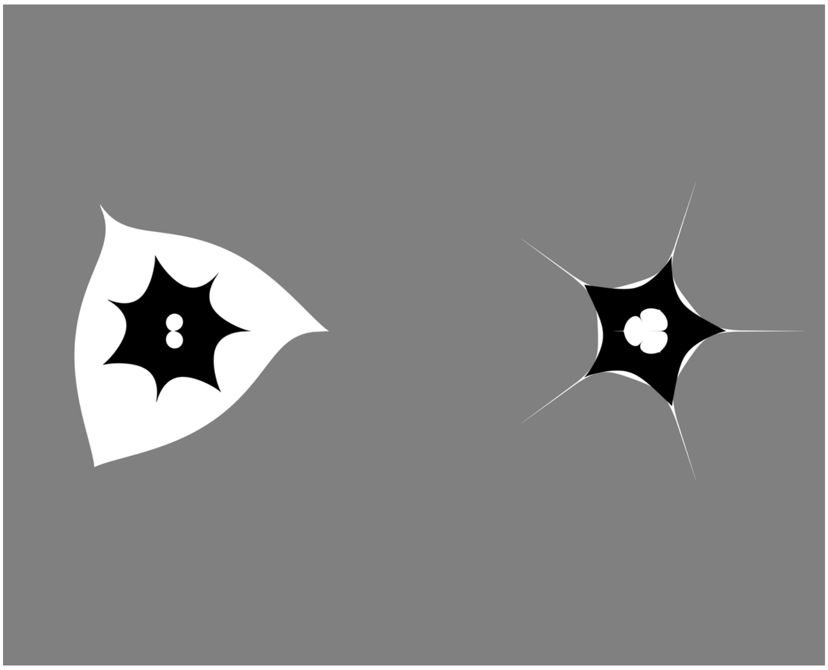

Figure 16. Pair of shapes from the shape-gaze-preference task.

Procedure. The participants diagnosed with autism were tested at the Department of Psychiatry at the University Hospital of a large city in Western Germany. The control participants were tested at the Department of Psychology of the University in the same city. Prior to the experiment, all participants gave consent to participate. All participants had normal or corrected-to-normal visual acuity. The experimental stimuli were presented on Tobii eye-tracker with a 17-in. monitor and a resolution of 1280 x 1024 pixels. Eye movements were recorded at a frequency of 50 Hz. At the beginning of the experiment, the experimenter greeted the participants and told them that they were going to take part in an experiment on social perception. After that, they sat down in a chair approximately 50 cm from the computer. They positioned their head so that the eyes were looking at the center of the screen, and the calibration of the eye-tracker was run. Participants then saw a black screen with white instructions explaining the first task. Participants could read the instructions at their own pace and then proceeded to the actual task. After they had completed the first task, participants proceeded to the instructions for the second task. To eliminate carry-over-effects, all tasks were presented in random order. Participants rated the difficulty of the task on a paper questionnaire next to the computer. The experiment ended after participants had completed all three tasks. After the experiment, participants completed the BDI (Beck & Steer, 1987; Hautzinger, 1995), the AQ (Baron-Cohen et al., 2006), the WAIS-R (Tewes, 1994).

Results

Behavioral data.

Face-attractiveness-detection task. To compare both groups' ability to detect attractiveness, I calculated in how many of the 20 trials participants' correctly chose the more attractive avatar. There was a significant difference between the HFA group (M = 14.45, SD = 2.73) and the control group (M = 15.90, SD = 2.12), $t(59) = -2.28$, $p = .026$, $d = -0.59$, indicating that the participants in the HFA group were impaired in their ability to choose the more attractive avatar in each pair (see Figure 17). One participant diagnosed with autism was excluded from this analysis because he/she admitted to having answered randomly on the task without looking at the faces. However, even with this participant included, the difference between the HFA group (M = 14.30, SD = 2.81) and the control group (M = 15.90, SD = 2.12) was still significant, $t(58) = 2.49$, $p = .016$, $d = -0.64$.

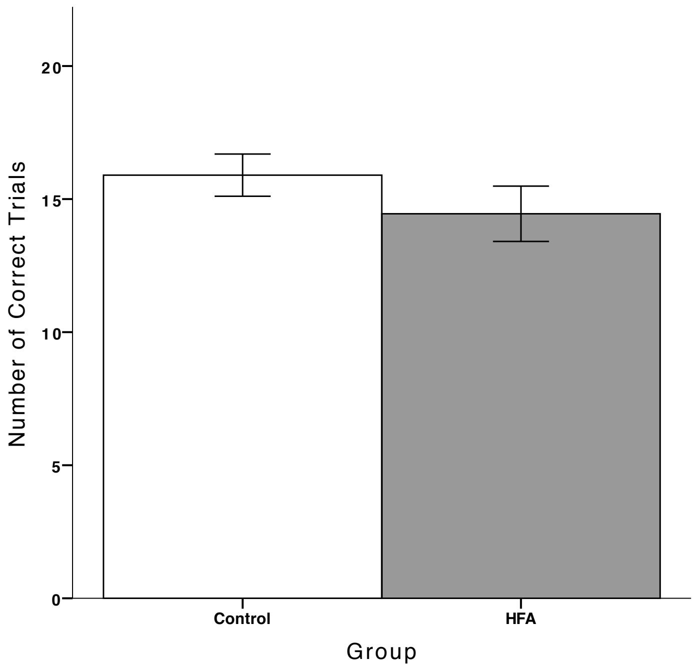

Figure 17. Results of the face-attractiveness-detection task. Error bars represent 95 %-CI of the mean.

With regard to reaction times, there was no significant difference between the HFA group (M = 7,033 ms, SD = 4,805 ms) and the control group (M = 5,869 ms, SD = 3,217 ms), $t(57) = -1.10$, $p = .278$, $d = 0.28$.

Perceived difficulty did differ significantly between the HFA group (M = 4.90, SD = 1.30) and the control group (M = 3.50, SD = 1.61), $t(58) = 3.71$, $p < .001$, d = 0.96, meaning that participants with autism did find it more difficult than participants in the control group to choose the more attractive face.

Because White et al. (2006) found that individuals with autism were less able to judge the attractiveness of faces of their own gender, I analyzed the number of correct trials, taking into account participants' gender and the gender of the stimuli. It is important to note that there were only ten females in both the HFA group (20 male, 10 female) and the control group (20 male, 10 female). Therefore, the results should only be interpreted carefully. However, White et al. (2006) had an even smaller number of females in their study (HFA: 10 male, 6 female; Control: 12 male, 10 female). I performed a 2 × 2 × 2 mixed ANOVA (group × participant gender × avatar gender), with group (HFA vs. Control) and participant gender (Male vs. Female) as between-subjects factors and avatar gender (Male vs. Female) as a within-subjects factor.

There was a significant main effect of group, $F(1, 55) = 4.62$, $p = .036$, $\eta^2_p = .077$, indicating that participants diagnosed with autism were impaired in their ability to detect attractiveness. There was also a significant main effect of participant gender, $F(1, 55) = 7.69$, $p = .008$, $\eta^2_p = .123$, in the direction that males generally performed better than females. All other effects were not significant, all $F < 1.89$. Most importantly, the three-way interaction between group, avatar gender, and participant gender was not significant, $F(1, 55) = 0.20$, $p = .655$, $\eta^2_p = .004$. Taken together, the results show that participants in the HFA group generally performed worse on the face-attractiveness-detection task than participants in the control group. However, contrary to the findings by White et al. (2006) this impairment was not limited to faces of the own gender.

Face-gaze-preference task. Pairs of avatars in the face-gaze-preference task were selected so that there was only a small difference in attractiveness between them, with both avatars in each pair being about equally attractive. Based on the results of the prestudy, I expected participants to choose the more attractive avatar only at chance level. In addition, I did not expect participants in the HFA group and the control group to differ on this task. As expected, there was no significant difference between the HFA group ($M = 9.34$, $SD = 2.71$) and the control group ($M = 9.53$, $SD = 2.55$), $t(57) = -0.28$, $p = .785$, $d = -0.07$.

With regard to reaction times, there was again no significant difference between the HFA group (M = 6,747 ms, SD = 4,266 ms) and the control group (M = 6,773 ms, SD = 3,673 ms), $t(57) = -0.25$, $p = .980$, $d = -0.01$.

Shape-gaze-preference task. Pairs of shapes in the shape-gaze-preference task were selected so that there was only a small difference in attractiveness between them, with both shapes in each pair being about equally attractive. Again, based on the results of the prestudy, I did not expect participants to choose one of the two shapes in each pair over the other. In addition, I did not expect participants in the HFA group and the control group to differ on this task. As expected, there was no significant difference between the HFA group (M = 10.27, SD = 2.39) and the control group (M = 10.37, SD = 2.59), $t(58) = -0.16$, $p = .877$, $d = -0.04$.

With regard to reaction times, there was no significant difference between the HFA group (M = 4,857 ms, SD = 3,111 ms) and the control group (M = 4,687 ms, SD = 2,674 ms), $t(58) = 0.25$, $p = .821$, $d = 0.06$.

Perceived difficulty did not differ significantly between the HFA group (M = 3.28, SD = 1.62) and the control group (M = 3.07, SD = 1.36), $t(57) = 0.54$, $p = .594$, $d = 0.14$, meaning that participants with autism did not find it more difficult than participants in the control group to choose the more attractive shape.

Eye-tracking data.

Preliminary analyses. To evaluate the quality of the eye-tracking data, the percentage of missing data was calculated for each participant. Fixations were calculated using a centroid method (Falkmer, Dahlman, Dukic, Bjällmark, & Larsson, 2008). In line with previous studies (Manor & Gordon, 2003; McPartland et al., 2011), fixations were defined as follows: For gaze to be classified as a fixation, participants' gaze had to be focused within a circular region on the screen with a radius of 30 pixels, which corresponds to about one degree of visual angle, for a duration of 100 ms or longer. Gaze data from some participants had to be excluded because they had more than 20% of missing data and fixations could not be reliably calculated. Based on this criterion, data from five participants diagnosed with autism had to be excluded in each of the three experimental tasks. In addition, one participant with autism did not complete the face-gaze-preference task.

Gaze cascade effect. Based on the procedure by Shimojo et al. (2003), I performed a gaze likelihood analysis for all three tasks. In line with Shimojo et al. (2003), only the last 2,500 ms before each participants' decision were analyzed, which is the mean reaction time for one trial minus one standard deviation. The gaze of the participants was tracked at a frequency of 50 MHz. Every tracking point at which the participants were gazing at the chosen face/shape was coded with a value of 1; every tracking point at which participants were gazing at the face/shape that they did not chose was coded with a value of 0. All tracking points were aligned backwards from the moment of the decision. To obtain the gaze likelihood curves for each task, I averaged for each tracking point across trials and participants. The resulting gaze likelihood curves show for each tracking point the probability that the chosen face/shape was gazed at (see Figure 18).

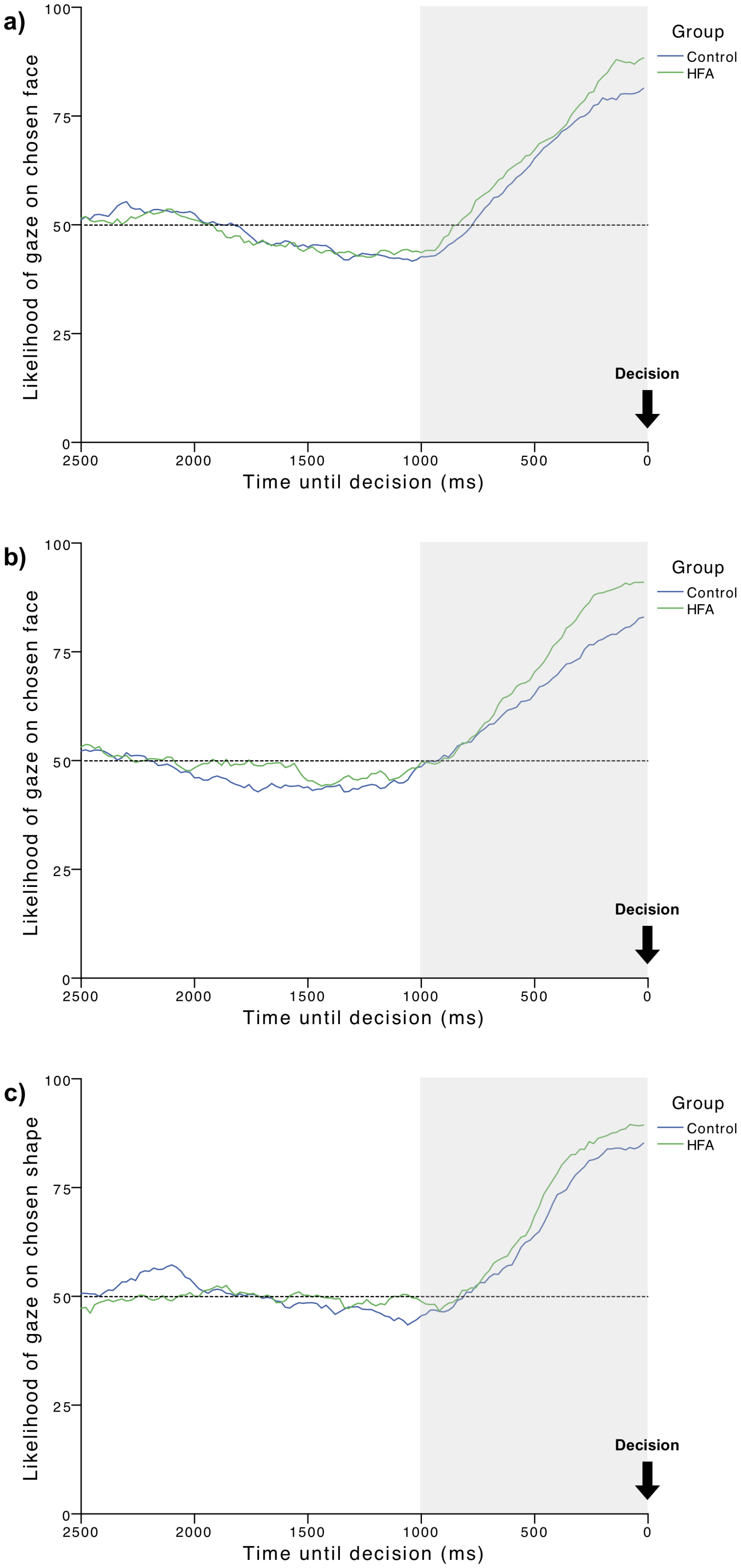
a)
Likelihood of gaze on chosen face
100
75
50
25
0
2500
2000
1500
1000
500
0
Time until decision (ms)
Group
Control
HFA
Decision
b)
Likelihood of gaze on chosen face
Time until decision (ms)
Group
Control
HFA
Decision
c)
Likelihood of gaze on chosen shape
Time until decision (ms)
Group
Control
HFA
Decision

Figure 18. Gaze likelihood curves for participants in the HFA and control group. The gaze likelihood curves show for each tracking point the probability that participants were gazing at the chosen face/shape. (a) Face-attractiveness-detection task, (b) face-gaze-preference task, and (c) shape-gaze-preference task.

As can be seen in the gaze likelihood curves, up until about one second before the decision, there is no bias to look at either the chosen or not chosen face/shape. The probability to look at either of the two stimuli is at chance level (50%), indicating that participants are equally likely to gaze at the chosen or not chosen face/shape. However, at around 1,000 ms before the decision there was a continuous rise in the likelihood to gaze at the chosen face/shape in all three tasks. As can be seen in Figure 18, both participants diagnosed with autism and participants in the control group showed a gaze cascade effect, indicated by an increasing bias to gaze at the chosen face/shape. The likelihood to gaze at the chosen face/shape was averaged for the last 1,000 ms before the decision and compared against chance level (50%) using one-sample *t* tests. The results in Table 7 show that both groups were significantly more likely to gaze at the chosen face/shape during the last 1,000 ms before the decision than at the not chosen face/shape.

Table 7

Likelihood that Gaze was on the Chosen Face/Shape during the last 1,000 ms before the Decision

Task	Group	Likelihood gaze on chosen face/shape	*df*	*t*	*p*
Face-attractiveness-detection	HFA	68.1 (11.3)	24	8.02	< .001
	Control	63.4 (9.1)	29	8.08	< .001
Face-gaze-preference	HFA	70.8 (10.2)	23	10.05	< .001
	Control	65.4 (11.2)	29	7.57	< .001
Face-shape-preference	HFA	70.0 (10.2)	24	9.82	< .001
	Control	65.7 (10.4)	29	8.29	< .001

According to the gaze cascade model (Shimojo et al., 2003), the gaze cascade effect should be stronger if the two stimuli are very similar in attractiveness, which was the case in the face-gaze-preference task and the shape-gaze-preference task, because the gaze behavior has a stronger influence on the decision when it is not driven by stimulus features. In line with the original results by Shimojo et al. (2003), it was thus predicted that the likelihood to gaze at the chosen alternative would be higher in both the face-gaze-preference task and the shape-gaze-preference task than in the face-attractiveness-detection task. Additionally, it should be investigated if the gaze cascade effect differed between participants diagnosed with autism and participants in the control group. Therefore, the likelihood to gaze at the chosen face/shape during the last 1,000 ms before the decision was analyzed in a 2 × 2 mixed ANOVA (group × task), with group (HFA vs. Control) as between-subjects factors and task (face-attractiveness-detection task vs. face-gaze-preference task vs. shape-gaze-preference task) as a within-subjects factor. There was no significant main effect for group, $F(1, 50) = 2.45$, $p = .124$, $\eta_p^2 = .047$, indicating that there was no difference between both groups in the likelihood to gaze at the chosen face/shape. There was an almost significant main effect for task, $F(2, 100) = 3.04$, $p = .052$, $\eta_p^2 = .057$. The interaction between group and task was not significant, $F(2, 100) = 0.43$, $p = .652$, $\eta_p^2 = .009$. Because of the predicted differences between the three tasks (Shimojo et al., 2003) and the almost significant main effect for task, planned contrasts were performed comparing the likelihood to gaze at the chosen face/shape in all three tasks. As the results in Table 8 show, participants were significantly more likely to gaze at the chosen face in the face-gaze-preference task ($M = 67.98$, $SD = 10.82$) as compared to the face-attractiveness-detection task ($M = 64.93$, $SD = 9.55$). In addition, there was a marginally significant difference ($p = .067$) between the face-attractiveness-detection task ($M = 64.93$, $SD = 9.55$) and the shape-gaze-preference task ($M = 67.15$, $SD = 9.43$). The likelihood to gaze at the chosen face did not differ between the face-gaze-preference task ($M = 67.98$, $SD = 10.82$) and the shape-gaze-preference task ($M = 67.15$, $SD = 9.43$).

Table 8

Planned Contrasts comparing all three Tasks

Task	$F(1, 50)$	p	η_p^2
Face-attractiveness-detection vs. Face-gaze-preference	5.60	.022	.101
Face-attractiveness-detection vs. Shape-gaze-preference	3.51	.067	.066
Face-gaze-preference vs. Shape-gaze-preference	0.37	.544	.007

These results partially support the effect found by Shimojo et al. (2003) that gaze behavior has a stronger influence on decision making when two stimuli are similar in attractiveness.

Because Gharib et al. (2011) found an increased tendency to gaze at the chosen face in individuals with autism at 40 ms before the decision, the data were reanalyzed and only the gaze behavior during this time period was used. The likelihood to gaze at the chosen face/shape during the last 40 ms before the decision was analyzed in a 2 × 2 mixed ANOVA (group × task), with group (HFA vs. Control) as a between-subjects factor and task (face-attractiveness-detection task vs. face-gaze-preference task vs. shape-gaze-preference task) as a within-subjects factor. There was a significant main effect for group, $F(1, 50) = 4.54$, $p = .038$, $\eta_p^2 = .083$, indicating that individuals with autism were more likely to gaze at the chosen face/stimulus during the last 40 ms before the decision than participants in the control group. All other effects were not significant, $F < 2.03$.

Taken together, the results of the gaze likelihood analysis show that the gaze cascade effect could be successfully replicated in both participants in the control group and participants with autism. Starting at around 1,000 ms before the decision, both groups showed a continuous increase in the likelihood to gaze at the chosen face/shape. Additionally, in line with the findings by Shimojo et al. (2003), the gaze cascade effect was stronger when the two faces were very similar in attractiveness. Also in line with the findings by Shimojo et al. (2003) was the trend that the gaze cascade effect was stronger when the two shapes were very similar in attractiveness. Supporting the previous findings by Gharib et al. (2011), individuals with autism showed an even stronger bias to gaze at the chosen face/shape at 40 ms before the decision than control participants.

Visual attention to regions of interest. To measure participants' visual attention to specific parts of the face, each face was divided into four regions of interest (ROIs): the eyes, the nose, the mouth, and the rest of the face not covered by the other regions of interest. Regions of interest were only defined for the faces in the face-attractiveness-detection task and the face-gaze-preference task. In the shape-gaze-preference task, no regions of interest could be specified. Because the avatar faces differed in size, regions of interest were fitted to each individual face. Figure 19 shows the regions of interest for one sample face.

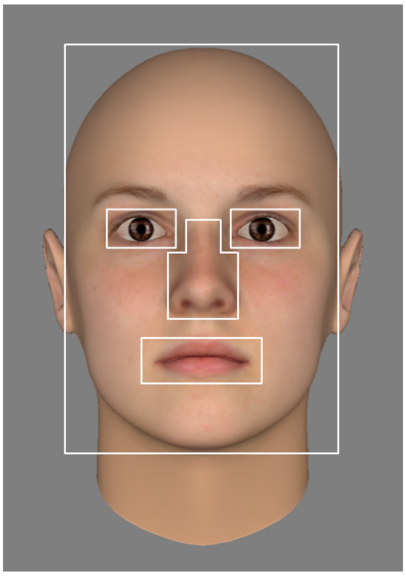

Figure 19. Sample face with the four regions of interest used for analyses: the eyes, the nose, the mouth, and the rest of the face.

As the dependent variables, the proportion of fixations (in %) and the total duration of fixation time (in ms)—the amount of time participants fixated at the different regions of interest in each trial—were calculated for each region of interest. Results were calculated separately for the face-attractiveness-detection task and the face-gaze-preference-task. Table 9 and Table 10 show the results for the face-attractiveness-detection task.

Table 9

Mean Proportion of Fixations to ROIs for the Face-attractiveness-detection Task

	HFA (n = 25)		Control (n = 30)					
Region of Interest	*M*	*SD*	*M*	*SD*	*df*	*t*	*p*	Cohen's *d*
Eyes	24.2	15.5	19.2	15.4	53	1.20	.235	0.32
Nose	22.3	11.2	25.8	17.0	53	-0.88	.383	-0.24
Mouth	13.3	12.3	9.7	10.1	53	1.17	.247	0.32
Face	40.2	10.0	45.3	16.4	53	-1.35	.184	-0.38

Table 10

Mean Total Duration of Fixation Time (in ms) to ROIs for the Face-attractiveness-detection Task

	HFA (n = 25)		Control (n = 30)					
Region of Interest	*M*	*SD*	*M*	*SD*	*df*	*t*	*p*	Cohen's *d*
Eyes	1367	1837	1156	1302	53	0.50	.235	0.13
Nose	1196	1232	1122	833	53	0.26	.383	0.07
Mouth	931	1203	571	373	53	1.55	.247	0.40
Face	1732	1925	1735	1396	53	-0.001	.184	< -0.01

As the results show, visual attention to the different regions of interest did not differ significantly between participants diagnosed with autism and participants in the control group (see *t* tests in Table 9 and Table 10). With regard to the proportion of fixations, both groups spent the most amount of attention to the face region—the part of the face not covered by the other regions of interest—and the least amount of attention to the mouth. However, this is not surprising because the face region was also the largest of the four regions of interest. The results for the mean total duration of fixation time mirror the results for the proportion of fixations, with participants fixating the longest on the face area and the shortest on the mouth.

Table 11 and Table 12 show the results for the face-gaze-preference task.

Table 11

Mean Proportion of Fixations to ROIs for the Face-gaze-preference Task

	HFA (n = 24)		Control (n = 30)					
Region of Interest	*M*	*SD*	*M*	*SD*	*df*	*t*	*p*	Cohen's *d*
Eyes	25.9	17.9	17.8	13.4	52	1.91	.062	0.51
Nose	20.7	13.7	23.2	10.9	52	-0.76	.449	-0.20
Mouth	12.9	14.8	11.2	12.1	52	0.45	.652	0.13
Face	40.6	13.6	47.8	15.3	52	-1.81	.077	-0.50

Table 12

Mean Total Duration of Fixation Time (in ms) to ROIs for the Face-gaze-preference Task

	HFA (n = 24)		Control (n = 30)					
Region of Interest	*M*	*SD*	*M*	*SD*	*df*	*t*	*p*	Cohen's *d*
Eyes	1558	2171	1375	2174	52	0.31	.760	0.08
Nose	1132	867	1219	986	52	-0.34	.738	-0.09
Mouth	785	949	583	405	52	1.05	.298	0.28
Face	1681	1554	2075	1529	52	-0.94	.354	-0.26

The results for the face-gaze-preference task were similar to the results for the face-attractiveness-detection task. Overall, visual attention to the different regions of interest did not differ significantly between participants diagnosed with autism and participants in the control group (see *t* tests in Table 11 and Table 12). This was true for both the proportion of fixations and the total duration of fixation time. Again, most of the attention was focused on the face region as opposed to the mouth.

Because Swanson and Siller (2013) did find a difference in the duration of the first fixation to the eyes, with individuals with autism making shorter first fixations to the eyes, I also compared the duration of the first fixation to the eyes between both groups. There was no significant difference between individuals diagnosed with autism (*M* = 208 ms, *SD* = 52 ms) and participants in the control group (*M* = 219 ms, *SD* = 53 ms) in the face-attractiveness-detection task, $t(52) = -0.47$, $p = .474$, $d = -0.19$. There was also no significant difference between the HFA group (*M* = 221 ms, *SD* = 70 ms) and the control group (*M* = 210 ms, *SD* = 55 ms) in the face-gaze-preference task, $t(51) = 0.68$, $p = .500$, $d = 0.18$.

To sum up, the analysis of participants' visual attention to the different regions of the faces revealed no differences between participants diagnosed with autism and participants in the control group for both the face-attractiveness-detection task and the face-gaze-preference task.

Discussion

The main goal of this study was to investigate whether adults diagnosed with high-functioning autism are impaired in their ability to detect attractiveness. Furthermore, the present study investigated whether the gaze cascade effect (Shimojo et al., 2003) could be replicated in individuals with autism.

As the behavioral results show, individuals with high-functioning autism are impaired in their ability to detect attractiveness. When presented with two faces—one attractive, one unattractive—individuals with autism chose the attractive face less often than participants in the control group. In line with these results, individuals with autism also rated the task to be more difficult than participants in the control group.

With regard to reaction times, there was no difference between individuals with autism and participants in the control group.

These results differ from a previous study by da Fonseca et al. (2011) who found that children with autism did not differ in their ratings of attractive faces from typically developing children. In a different study, White et al. (2006) found impaired performance of individuals with autism in the detection of attractiveness in faces of the own gender. Compared to White et al. (2006), the results from the present study support the notion that the ability to detect attractiveness in faces might be generally impaired in individuals with autism and not limited to faces of the own gender. The difference in results between the study by White et al. (2006) and the present study can in part be explained by the small sample size of the study by White et al. (2006) (HFA: 10 male, 6 female; Control: 12 male, 10 female), which had probably too little statistical power to detect a general deficit.

In order to investigate if the impairment in the detection of attractiveness could be due to autistic participants focusing on other parts of the face than typically developing participants and therefore missing parts of the face that might be crucial in judging attractiveness, the present study analyzed the visual attention directed at different parts of the face. The results from the present study show no differences between participants with autism and participants in the control group, and are therefore in line with previous studies that have also found no differences (e.g., McPartland et al., 2011; Georgescu et al., 2013; Rutherford & Towns, 2008). The highest amount of attention was focused on the face region—without the other regions of interest—followed by the eyes and the nose; the least amount of attention was focused on the mouth. This distribution of attention was an adaptive strategy with regard to the task that participants had to perform—namely, to choose the more attractive face—because the face area contained the most information to make that decision. Because several other studies have found a tendency in individuals with autism to avoid the gaze of faces (e.g., Jones et al., 2008; Dalton et al., 2005), the present study also used a micromeasure proposed by Swanson and Siller (2013)—duration of the first fixation to the eyes—to test if individuals with autism fixate less at the eyes. However, contrary to Swanson and Siller (2013), individuals diagnosed with autism did fixate at the eyes during the first fixation as long as participants in the control group. Thus, the results of the present study do not support the idea that individuals with autism avert their gaze to avoid looking at the eyes of other persons. Even though gaze aversion is considered to be one of the prominent signs auf autism and has been documented in several studies (e.g., Spezio, Adolphs, Hurley, & Piven, 2007; Jones et al., 2008; Richer & Coss, 1976; Hutt & Ounsted, 1966), there are also studies that have failed to show gaze aversion in social interactions (e.g., Doherty-Sneddon, Whittle, & Riby, 2012; García-Pérez, Lee, & Hobson, 2007). One possible explanation for these diverging results may be the different levels of functioning of the participants with autism in these studies. The participants in the present study were adults with high-functioning autism and may therefore have developed compensatory strategies for social interactions. However, the duration of the first fixation to the eyes is a measure that lies out of conscious control und may rather represent an automatic orientation. Taking this into account, the results rather support the idea that high-functioning individuals with autism automatically fixate the eyes as typically developing individuals.

In sum, the results show that the impaired performance of individuals with autism in detecting attractiveness cannot be explained by a different distribution of attention to parts of the face because individuals with autism did not differ from participants in the control group. Because the gaze behavior does not reveal a possible explanation for the impairment, other methodologies should be used to test alternative explanations. For instance, future fMRI-studies could investigate the different neural responses to attractive and unattractive faces in individuals with autism. Previous fMRI-studies have shown that attractive faces activate reward centers in the brain, such as the nucleus accumbens (NAcc) and the orbitofrontal cortex (OFC) (Aharon et al., 2001; O'Doherty et al., 2003; Cloutier et al., 2008; Liang et al., 2010; Winston et al., 2007). It should therefore be investigated if individuals with autism show a similar response to attractive faces in these areas. In a first study in this direction, Dichter et al. (2012) could show that faces produced an altered response in the reward circuitry in individuals with autism as compared to typically developing individuals, hinting at a possible explanation for the impairment in the perception of attractiveness that should be explored in future studies.

Another possible explanation for autistic participants' impaired performance in the perception of attractiveness may be that individuals with autism do not have access to the same standards typically developing individuals use for judging attractiveness. Several studies have shown that the averageness of faces has a strong influence on the perception of attractiveness—average faces are judged to be more attractive (e.g., Langlois & Roggman, 1990). Because individuals with autism may have less exposure to faces than typically developing individuals, it may be harder for them to judge the averageness of faces.

With regard to the relation between participants gaze behavior and their decisions, the present study successfully replicated the gaze cascade effect in individuals with autism. At around 1,000 ms both groups showed a continuous increase in the likelihood to gaze at the chosen face/shape. Additionally, in line with the findings by Shimojo et al. (2003) the gaze cascade effect was stronger when the two faces were very similar in attractiveness. Also in line with the findings by Shimojo et al. (2003) was the trend that the gaze cascade effect was also stronger when the two shapes were very similar in attractiveness. When analyzing the last 1,000 ms before the decision—about the time period when the likelihood to gaze at the chosen face starts to rise above 50%—individuals with autism did not differ in their likelihood to gaze at the chosen from participants in the control group. Only when analyzing the last 40 ms before the decision, such as Gharib et al. (2011) did in their study, did I find an increased likelihood in individuals with autism to gaze at the chosen face/shape compared to typically developing individuals. This difference during such a short time period before the decision is difficult to interpret. Gharib et al. (2011) did not give a theory-driven explanation in their study for selecting this time period for analysis.

Interestingly, participants with autism did show a gaze bias toward both faces and abstract shapes. Even though some studies have shown that facial stimuli may be aversive (e.g., Corden et al., 2008) for individuals with autism, this did not influence the gaze bias towards them.

The results of the present study may also be conceptually in contrast to a study by Grynszpan et al. (2012) that found impaired self-monitoring of gaze in individuals with autism. Because the gaze cascade effect relies on a proposed connection between gaze behavior and decision making, the successful replication of the gaze cascade effect in individuals with autism may hint at the fact that gaze behavior may also feed into decision processes in individuals with autism. However, future studies are needed to further entangle the relation between self-monitoring of gaze and decision making in autism. Overall, the results hint at the possibility that mere exposure (Zajonc, 2001) and preferential looking (Birch et al., 1985) might work in a similar way together in individuals with autism to produce an increasing gaze bias towards the chosen face/shape. It is important to note, however, that it is not possible to draw this conclusion solely based on the results of this study because it did not include an active manipulation of participants' gaze. A further study with an active gaze manipulation is needed to show that the gaze behavior itself did influence participants' decisions. Still, the results of the gaze likelihood analysis show that gaze and decision-making are related in a similar way in individuals with autism as in typically developing individuals.

There are some limitations to the present study that should be considered: First, the present study only included a small number of female participants (HFA: 20 male, 10 female; Control: 20 male, 10 female). It should therefore be further investigated if the perception of attractiveness does differ between males and females in autism. Interestingly, the present study did find a significant main effect of participants' gender on the perception of attractiveness, with males generally performing better than females. This was true in both the HFA group and the control group. On the whole, female participants with autism performed worst on the attractiveness detection task. In light of the background of theories proposing that symptom severity in autism is stronger in autistic females than in autistic males (Holtmann, Bölte, & Poustka, 2007; Crespi & Badcock, 2008), this trend should be further investigated in future studies with more female autistic participants.

Second, even though participants' position in front of the monitor was standardized, gaze data from five participants diagnosed with autism had to be excluded because they had more than 20% of missing data and fixations could not be reliably calculated. Therefore, a future replication of the study should include a larger sample of individuals with autism to account for possible tracking problems.

To conclude, the present study aimed at investigating whether individuals with autism are impaired in their ability to detect attractiveness. Taking into account participants' gaze behavior, the study further investigated whether this impairment may be explained by different patterns of visual attention to parts of the face in individuals with autism. Additionally, the study investigated whether an orienting bias towards the chosen face—the gaze cascade effect (Shimojo et al., 2003)—could be replicated in individuals with autism. The results suggest that individuals with autism are impaired in their ability to distinguish between attractive and unattractive faces. However, patterns of visual attention did not differ between individuals with autism and participants in the control group. Finally, the gaze cascade effect could be successfully replicated in individuals with autism. Thus, even though individuals with autism were impaired in their ability to detect attractiveness, gaze behavior was comparable to typical adults.

General Discussion

The aim of the four studies in this thesis was to advance knowledge in two major fields of social cognition—cross-cultural psychology and high-functioning autism—by using virtual avatars:

Study 1 assessed the validity of avatars by investigating whether people trust avatars in a similar way as they trust photographs in an e-commerce setting. Both positive reputation and trustworthy avatars led to higher purchase rates than negative reputation and untrustworthy avatars. However, participants' responses to seller avatars were influenced by the uncertainty inherent in the reputation scores. When the uncertainty that the seller would ship the product was high, participants were more susceptible to the influence of the seller avatars. As the results show, there are subtle differences in the effects of virtual avatars as opposed to photographs.

Study 2 focused on cross-cultural differences in trust and investigated whether individualistic and collectivistic cultures differ in their use of two different sources of information—factual information and tacit cues—in building trust. As the results of the trust game show, participants from Germany, a more individualistic culture, and the United Arab Emirates, a more collectivistic culture, were influenced by both reputation scores and seller avatars. However, both cultures did differ in the effect of reputation on purchase decisions. German participants bought significantly less often than Arab participants when the reputation of the seller was low, hinting at a possible higher susceptibility to low reputation of members of individualistic cultures.

Study 3 focused on the question whether adults diagnosed with HFA are impaired in their ability to detect direct gaze. In addition, it was also tested whether these gaze processing impairments could—at least in part—be due to the uncontrollability of the social stimuli by giving participants control over the gaze of a realistic virtual character. As the results show, adult individuals with high-functioning autism are impaired in their ability to distinguish between direct and averted gaze. Results of the setting task with the virtual character show that individuals with autism are in fact able to establish direct gaze accurately when they have full control over a social stimulus, hinting at a possible training mechanism that may help individuals with autism to improve at the detection of direct gaze.

Study 4 investigated whether adults diagnosed with high-functioning autism are impaired in their ability to detect attractiveness. Also, by monitoring participants' gaze while they made their decisions, it was tested whether atypical patterns of visual attention can in part explain a possible deficit. As the behavioral results show, individuals with high-functioning autism are impaired in their ability to detect attractiveness. However, gaze behavior was comparable to typical adults.

Limitations

It is important to consider that each of the four studies has its own specific limitations:

Study 1 did use only photorealistic avatars. The results with regard to the trust building effect of avatars can therefore not easily be generalized to other forms of avatars that are less realistic. Besides, participants in the study were informed that the avatars were created on the basis of photographs from real humans. This knowledge about the origins of the avatars may have further increased their trustworthiness. Therefore, other studies using avatars should be careful at assuming a general effect of avatars on producing behavioral trust.

Study 2 investigated only two very specific cultures: Germany and the United Arab Emirates. Even though both cultures differ on the individualism/collectivism dimension (Hofstede et al., 2010), there are also numerous other differences between the two cultures that could potentially account for the differences found in the study. For instance, both cultures also differ with regard to power distance.

Study 3 only focused on adults with high-functioning autism and therefore does not allow for a developmental perspective on autistic deficits in gaze detection. Additionally, the non-social, geometric stimulus was less visually complex than the social stimulus. Therefore, it is possible that a more complex non-social, geometric stimulus may have uncovered a general deficit in the processing of moving stimuli by individuals with autism.

Study 4 included only a small number of female participants in both the HFA and the control group. The results did show a tendency that female participants with high-functioning autism were particularly impaired in the perception of attractiveness. However, this should be replicated in a larger sample of females.

Future Research Directions

A future replication of **Study 1** should investigate if avatars that are not photorealistic can elicit similar levels of behavioral trust. Moreover, it should be further investigated if the knowledge that the avatars were based on real individuals did increase their trustworthiness.

Study 2 should be replicated with a larger group of different individualistic and collectivistic cultures. Special attention should hereby be given to the other cultural dimensions and their potential influence on the likelihood to trust.

Because **Study 3** has shown that individuals with autism can in part overcome their deficit in the detection of direct gaze if given active control over the gaze of the avatar, it should be investigated if a training program utilizing active gaze control can produce a lasting improvement in direct gaze detection in autism. Additionally, it should be tested if such improvements in direct gaze detection may generalize to improvements in other areas of social interaction.

Even though **Study 4** did find an impairment in the perception of attractiveness in individuals with autism, eye-tracking results did not reveal cues to an underlying mechanism that may account for the impairment. Based on the finding that faces (Dichter et al., 2012) produced an altered response in the reward circuitry in individuals with autism, future fMRI-studies may help to investigate the underlying differences in the neural responses to attractive faces.

Conclusion

In sum, this thesis investigated how the unique advantages of avatars can help uncover new findings in social cognition: Avatars can produce similar behavioral responses as real humans, as shown by the willingness to trust avatars in a similar way as other humans. Individualistic and collectivistic cultures are similar in their tendency to rely both on factual information and tacit, social cues. However, negative factual information may have special salience in individualistic cultures. With regard to psychopathology, individuals with autism may be impaired in two basic areas of social perception—the detection of direct gaze and the perception of attractiveness. However, with the help of interactive paradigms individuals with autism may learn to overcome some of their deficits.

Bibliography

Adams, R. B. & Kleck, R. E. (2005). Effects of direct and averted gaze on the perception of facially communicated emotion. *Emotion*, *5*(1), 3-11. doi:10.1037/1528-3542.5.1.3

Aharon, I., Etcoff, N., Ariely, D., Chabris, C. F., O'Connor, E., & Breiter, H. C. (2001). Beautiful faces have variable reward value: FMRI and behavioral evidence. *Neuron*, *32*(3), 537-551.

Akechi, H., Senju, A., Kikuchi, Y., Tojo, Y., Osanai, H., & Hasegawa, T. (2010). The effect of gaze direction on the processing of facial expressions in children with autism spectrum disorder: An ERP study. *Neuropsychologia*, *48*(10), 2841-2851. doi:10.1016/j.neuropsychologia.2010.05.026

Ashwin, E., Ashwin, C., Rhydderch, D., Howells, J., & Baron-Cohen, S. (2009). Eagle-eyed visual acuity: An experimental investigation of enhanced perception in autism. *Biological Psychiatry*, *65*(1), 17-21. doi:10.1016/j.biopsych.2008.06.012

Ayres, I., Banaji, M., & Jolls, C. (2011). Race effects on ebay. *The Selected Works of Ian Ayres.* Retrieved from http://works.bepress.com/ian_ayres/23

Ballew, C. C. & Todorov, A. (2007). Predicting political elections from rapid and unreflective face judgments. *Proceedings of the National Academy of Sciences*, *104*(46), 17948-17953.

Baron-Cohen, S. (1988). Social and pragmatic deficits in autism: Cognitive or affective?. *Journal of Autism and Developmental Disorders*, *18*(3), 379-402.

Baron-Cohen, S. (2009). Autism: The empathizing-systemizing (ES) theory. *Annals of the New York Academy of Sciences*, *1156*(1), The Year in Cognitive Neuroscience 2009), 68-80. doi:10.1111/j.1749-6632.2009.04467.x

Baron-Cohen, S. & Wheelwright, S. (2004). The empathy quotient: An investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, *34*(2), 163-75.

Baron-Cohen, S., Ashwin, E., Ashwin, C., Tavassoli, T., & Chakrabarti, B. (2009). Talent in autism: Hyper-systemizing, hyper-attention to detail and sensory hypersensitivity. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1522), 1377-1383. doi:10.1098/rstb.2008.0337

Baron-Cohen, S., Hoekstra, R. A., Knickmeyer, R., & Wheelwright, S. (2006). The autism-spectrum quotient (AQ)--adolescent version. *Journal of Autism and Developmental Disorders*, *36*(3), 343-50. doi:10.1007/s10803-006-0073-6

Baron-Cohen, S., Richler, J., Bisarya, D., Gurunathan, N., & Wheelwright, S. (2003). The systemizing quotient: An investigation of adults with Asperger syndrome or high-functioning autism, and normal sex differences. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *358*(1430), 361-374. doi:10.1098/rstb.2002.1206

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "reading the mind in the eyes" test revised version: A study with normal adults, and adults with asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *42*(2), 241-251.

Beck, A. T. & Steer, R. A. (1987). *Beck depression inventory - manual.* San Antonio: The Psychological Association.

Bente, G., Baptist, O., & Leuschner, H. (2012). To buy or not to buy: Influence of seller photos and reputation on buyer trust and purchase behavior.

International Journal of Human-Computer Studies, *70*(1), 1-13. doi:10.1016/j.ijhcs.2011.08.005

Bente, G., Donaghy, W. C., & Suwelack, D. (1998). Sex differences in body movement and visual attention: An integrated analysis of movement and gaze in mixed-sex dyads. *Journal of Nonverbal Behavior*, *22*(1), 31-58.

Bente, G., Eschenburg, F., & Krämer, N. (2007). Virtual gaze. A pilot study on the effects of computer simulated gaze in avatar-based conversations. *Virtual Reality, HCI 2007*, 185-194.

Bente, G., Krämer, N. C., Petersen, A., & de Ruiter, J. P. (2001). Computer animated movement and person perception: Methodological advances in nonverbal behavior research. *Journal of Nonverbal Behavior*, *25*(3), 151-166.

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

Beugelsdijk, S., De Groot, H. L., & Van Schaik, A. B. (2004). Trust and economic growth: A robustness analysis. *Oxford Economic Papers*, *56*(1), 118-134.

Birch, E. E., Shimojo, S., & Held, R. (1985). Preferential-looking assessment of fusion and stereopsis in infants aged 1-6 months. *Investigative Ophthalmology & Visual Science*, *26*(3), 366-370.

Bohnet, I., Herrmann, B., & Zeckhauser, R. (2010). Trust and the reference points for trustworthiness in gulf and western countries. *The Quarterly Journal of Economics*, *125*(2), 811-828.

Bolton, G. E., Katok, E., & Ockenfels, A. (2004). Trust among internet traders: A behavioral economics approach. *Analyse Und Kritik*, *26*(1), 185-202.

Branzei, O., Vertinsky, I., & Camp, R. D. (2007). Culture-contingent signs of trust in emergent relationships. *Organizational Behavior and Human Decision Processes*, *104*(1), 61-82. doi:10.1016/j.obhdp.2006.11.002

Chua, H. F., Boland, J. E., & Nisbett, R. E. (2005). Cultural variation in eye movements during scene perception. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(35), 12629-33. doi:10.1073/pnas.0506162102

Cloutier, J., Heatherton, T. F., Whalen, P. J., & Kelley, W. M. (2008). Are attractive people rewarding? Sex differences in the neural substrates of facial attractiveness. *Journal of Cognitive Neuroscience*, *20*(6), 941-951.

Corden, B., Chilvers, R., & Skuse, D. (2008). Avoidance of emotionally arousing stimuli predicts social-perceptual impairment in Asperger's syndrome. *Neuropsychologia*, *46*(1), 137-47. doi:10.1016/j.neuropsychologia.2007.08.005

Crespi, B. & Badcock, C. (2008). Psychosis and autism as diametrical disorders of the social brain. *The Behavioral and Brain Sciences*, *31*(3), 241-61; discussion 261-320. doi:10.1017/S0140525X08004214

Da Fonseca, D., Santos, A., Rosset, D., & Deruelle, C. (2011). The “beauty is good” for children with autism spectrum disorders too. *Research in Autism Spectrum Disorders*, *5*(1), 299-304. doi:10.1016/j.rasd.2010.04.012

Dakhli, M. (2009). Investigating the effects of individualism-collectivism on trust and cooperation. *Psychology Journal*, *6*(3), 90-99.

Dalton, K. M., Nacewicz, B. M., Johnstone, T., Schaefer, H. S., Gernsbacher, M. A., Goldsmith, H. H., et al. (2005). Gaze fixation and the neural circuitry of face processing in autism. *Nature Neuroscience*, *8*(4), 519-26. doi:10.1038/nn1421

Dawson, G., Meltzoff, A. N., Osterling, J., Rinaldi, J., & Brown, E. (1998). Children with autism fail to orient to naturally occurring social stimuli. *Journal of Autism and Developmental Disorders*, *28*(6), 479-85.

Dawson, G., Webb, S. J., & McPartland, J. (2005). Understanding the nature of face processing impairment in autism: Insights from behavioral and electrophysiological studies. *Developmental Neuropsychology*, *27*(3), 403-24. doi:10.1207/s15326942dn2703_6

Dichter, G. S. & Belger, A. (2007). Social stimuli interfere with cognitive control in autism. *NeuroImage*, *35*(3), 1219-30. doi:10.1016/j.neuroimage.2006.12.038

Dichter, G. S., Richey, J. A., Rittenberg, A. M., Sabatino, A., & Bodfish, J. W. (2012). Reward circuitry function in autism during face anticipation and outcomes. *Journal of Autism and Developmental Disorders*, *42*(2), 147-60. doi:10.1007/s10803-011-1221-1

Diekmann, A., Jann, B., & Wyder, D. (2007). Trust and reputation in internet auctions. In K. S. Cook, C. Snijders, V. Buskens, & C. Cheshire (Eds.), *ETrust. Forming relationships in the online world.* (pp. 139-65). New York, NY: Russel Sage Foundation.

Doherty-Sneddon, G., Whittle, L., & Riby, D. M. (2012). Gaze aversion during social style interactions in autism spectrum disorder and williams syndrome. *Research in Developmental Disabilities*, *34*(1), 616-626. doi:10.1016/j.ridd.2012.09.022

Doleac, J. L. & Stein, L. C. (2010). The visible hand: Race and online market outcomes. *Social Science Research Network.* Retrieved from http://dx.doi.org/10.2139/ssrn.1615149

Eagly, A. H., Ashmore, R. D., Makhijani, M. G., & Longo, L. C. (1991). What is beautiful is good, but…: A meta-analytic review of research on the physical attractiveness stereotype. *Psychological Bulletin*, *110*(1), 109-128.

Elsabbagh, M., Volein, A., Csibra, G., Holmboe, K., Garwood, H., Tucker, L., et al. (2009). Neural correlates of eye gaze processing in the infant broader autism phenotype. *Biological Psychiatry*, *65*(1), 31-8. doi:10.1016/j.biopsych.2008.09.034

FaceGen. (2012). *Singular Inversions Inc.*

Faja, S., Aylward, E., Bernier, R., & Dawson, G. (2008). Becoming a face expert: A computerized face-training program for high-functioning individuals with autism spectrum disorders. *Developmental Neuropsychology*, *33*(1), 1-24. doi:10.1080/87565640701729573

Falck-Ytter, T. & von Hofsten, C. (2011). How special is social looking in ASD: A review. *Progress in Brain Research*, *189*, 209.

Falkmer, T., Dahlman, J., Dukic, T., Bjällmark, A., & Larsson, M. (2008). Fixation identification in centroid versus start-point modes using eye-tracking data. *Perceptual and Motor Skills*, *106*(3), 710-724.

Farroni, T., Csibra, G., Simion, F., & Johnson, M. H. (2002). Eye contact detection in humans from birth. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(14), 9602-9605. doi:10.1073/pnas.152159999

Ferrara, C. & Hill, S. D. (1980). The responsiveness of autistic children to the predictability of social and nonsocial toys. *Journal of Autism and Developmental Disorders*, *10*(1), 51-57.

Flannery, K. & Horner, R. H. (1994). The relationship between predictability and problem behavior for students with severe disabilities. *Journal of Behavioral Education*, *4*(2), 157-176.

Frith, U. & de Vignemont, F. (2005). Egocentrism, allocentrism, and Asperger syndrome. *Consciousness and Cognition*, *14*(4), 719-38. doi:10.1016/j.concog.2005.04.006

Gamer, M. & Hecht, H. (2007). Are you looking at me? Measuring the cone of gaze. *Journal of Experimental Psychology. Human Perception and Performance*, *33*(3), 705-715. doi:10.1037/0096-1523.33.3.705

García-Pérez, R. M., Lee, A., & Hobson, R. P. (2007). On intersubjective engagement in autism: A controlled study of nonverbal aspects of conversation. *Journal of Autism and Developmental Disorders*, *37*(7), 1310-22. doi:10.1007/s10803-006-0276-x

Gächter, S., Herrmann, B., & Thöni, C. (2010). Culture and cooperation. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *365*(1553), 2651-61. doi:10.1098/rstb.2010.0135

Georgescu, A. L., Kuzmanovic, B., Santos, N. S., Tepest, R., Bente, G., Tittgemeyer, M., et al. (2013). Perceiving nonverbal behavior: Neural correlates of processing movement fluency and contingency in dyadic interactions. *Human Brain Mapping.* doi:10.1002/hbm.22259

Gepner, B., de Gelder, B., & de Schonen, S. (1996). Face processing in autistics: Evidence for a generalised deficit?. *Child Neuropsychology*, *2*(2), 123-139.

Gepner, B., Deruelle, C., & Grynfeltt, S. (2001). Motion and emotion: A novel approach to the study of face processing by young autistic children. *Journal of Autism and Developmental Disorders*, *31*(1), 37-45.

Gharib, A., Mier, D., Adolphs, R., & Shimojo, S. (2011). Gaze and preference decision making in autism. *Journal of Vision*, *11*(11).

Gheorghiu, M. A., Vignoles, V. L., & Smith, P. B. (2009). Beyond the united states and japan: Testing yamagishi's emancipation theory of trust across 31 nations. *Social Psychology Quarterly*, *72*(4), 365-383.

Gielis, J. (2003). A generic geometric transformation that unifies a wide range of natural and abstract shapes. *American Journal of Botany*, *90*(3), 333-338.

Glaholt, M. G. & Reingold, E. M. (2009). Stimulus exposure and gaze bias: A further test of the gaze cascade model. *Attention, Perception & Psychophysics*, *71*(3), 445-50. doi:10.3758/APP.71.3.445

Glaholt, M. G. & Reingold, E. M. (2011). Eye movement monitoring as a process tracing methodology in decision making research. *Journal of Neuroscience, Psychology, and Economics*, *4*(2), 125-146. doi:10.1037/a0020692

Golan, O., Ashwin, E., Granader, Y., McClintock, S., Day, K., Leggett, V., et al. (2009). Enhancing emotion recognition in children with autism spectrum conditions: An intervention using animated vehicles with real emotional faces. *Journal of Autism and Developmental Disorders*, *40*, 269-279. Retrieved October 10, 2009, from http://www.springerlink.com/content/y4164n56855305p0/ doi:10.1007/s10803-009-0862-9

Gomot, M. & Wicker, B. (2011). A challenging, unpredictable world for people with autism spectrum disorder. *International Journal of Psychophysiology.* doi:10.1016/j.ijpsycho.2011.09.017

Grammer, K. & Thornhill, R. (1994). Human (homo sapiens) facial attractiveness and sexual selection: The role of symmetry and averageness. *Journal of Comparative Psychology*, *108*(3), 233-242.

Gregg, D. G. & Scott, J. E. (2006). The role of reputation systems in reducing on-line auction fraud. *International Journal of Electronic Commerce*, *10*(3), 95-120.

Grice, S. J., Halit, H., Farroni, T., Baron-Cohen, S., Bolton, P., & Johnson, M. H. (2005). Neural correlates of eye-gaze detection in young children with autism. *Cortex*, *41*(3), 342-53.

Grynszpan, O., Nadel, J., Martin, J. C., Simonin, J., Bailleul, P., Wang, Y., et al. (2012). Self-Monitoring of gaze in high functioning autism. *Journal of Autism and Developmental Disorders, 42*(8), 1642-1650.

Hautzinger, M. (1995). *Beck-Depressions-Inventar (BDI)*. Bern: Huber.

Herrmann, B., Thöni, C., & Gächter, S. (2008). Antisocial punishment across societies. *Science (New York, N.Y.)*, *319*(5868), 1362-7. doi:10.1126/science.1153808

Hofstede, G., Hofstede, G. J., & Minkov, M. (2010). *Cultures and organisations: Software of the mind.* New York: McGraw-Hill.

Holtmann, M., Bölte, S., & Poustka, F. (2007). Autism spectrum disorders: Sex differences in autistic behaviour domains and coexisting psychopathology. *Developmental Medicine & Child Neurology*, *49*(5), 361-366.

Howard, M. A., Cowell, P. E., Boucher, J., Broks, P., Mayes, A., Farrant, A., et al. (2000). Convergent neuroanatomical and behavioural evidence of an amygdala hypothesis of autism. *Neuroreport*, *11*(13), 2931-35.

Hudson, J. (2006). Institutional trust and subjective well-being across the EU. *Kyklos*, *59*(1), 43-62.

Huff, L. & Kelley, L. (2003). Levels of organizational trust in individualist versus collectivist societies: A seven-nation study. *Organization Science*, *14*(1), 81-90.

Hutt, C. & Ounsted, C. (1966). The biological significance of gaze aversion with particular reference to the syndrome of infantile autism. *Behavioral Science*, *11*(5), 346-356.

Jamal, A. (2007). When is social trust a desirable outcome?: Examining levels of trust in the Arab world. *Comparative Political Studies, 40*(11), 1328-1349. doi:10.1177/001041400629183

Jemel, B., Mottron, L., & Dawson, M. (2006). Impaired face processing in autism: Fact or artifact?. *Journal of Autism and Developmental Disorders*, *36*(1), 91-106. doi:10.1007/s10803-005-0050-5

Jones, E., Moore, J. N., Stanaland, A. J., & Wyatt, R. A. (1998). Salesperson race and gender and the access and legitimacy paradigm: Does difference make a difference?. *The Journal of Personal Selling and Sales Management*, 71-88.

Jones, W., Carr, K., & Klin, A. (2008). Absence of preferential looking to the eyes of approaching adults predicts level of social disability in 2-year-old toddlers with autism spectrum disorder. *Archives of General Psychiatry*, *65*(8), 946.

Joseph, R. M., Ehrman, K., McNally, R., & Keehn, B. (2008). Affective response to eye contact and face recognition ability in children with ASD. *Journal of the International Neuropsychological Society*, *14*(06), 947-955. doi:10.10170S1355617708081344

Klin, A., Lin, D. J., Gorrindo, P., Ramsay, G., & Jones, W. (2009). Two-year-olds with autism orient to non-social contingencies rather than biological motion. *Nature*, *459*(7244), 257-61. doi:10.1038/nature07868

Knight, J. (2001). Social norms and the rule of law: Fostering trust in a socially diverse society. In K. Cook (Ed.), *Trust in society.* (pp. 354-73). New York, NY: Russel Sage Foundation.

Kuzmanovic, B., Georgescu, A. L., Eickhoff, S. B., Shah, N. J., Bente, G., Fink, G. R., et al. (2009). Duration matters: Dissociating neural correlates of detection and evaluation of social gaze. *NeuroImage*, *46*(4), 1154-63. doi:10.1016/j.neuroimage.2009.03.037

Kylliäinen, A. & Hietanen, J. K. (2006). Skin conductance responses to another person's gaze in children with autism. *Journal of Autism and Developmental Disorders*, *36*(4), 517-25. doi:10.1007/s10803-006-0091-4

Langlois, J. H. & Roggman, L. A. (1990). Attractive faces are only average. *Psychological Science*, *1*(2), 115-121. doi:10.1111/j.1467-9280.1990.tb00079.x

Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin*, *126*(3), 390-423.

Langlois, J. H., Roggman, L. A., Casey, R. J., Ritter, J. M., Rieser-Danner, L. A., & Jenkins, V. Y. (1987). Infant preferences for attractive faces: Rudiments of a stereotype?. *Developmental Psychology*, *23*(3), 363-369.

Lee, F., Hallahan, M., & Herzog, T. (1996). Explaining real-life events: How culture and domain shape attributions. *Personality and Social Psychology Bulletin*, *22*(7), 732-741. doi:10.1177/0146167296227007

Lehnhardt, F. G., Gawronski, A., Volpert, K., Schilbach, L., Tepest, R., Huff, W., et al. (2011). Autismus-Spektrum-Sto□rungen im Erwachsenenalter: Klinische und neuropsychologische Befunde spa□tdiagnostizierter Asperger-Syndrome. *Fortschritte Der Neurologie Und Psychiatrie*, *79*, 290-297.

Liang, X., Zebrowitz, L. A., & Zhang, Y. (2010). Neural activation in the “reward circuit” shows a nonlinear response to facial attractiveness. *Social Neuroscience*, *5*(3), 320-334.

Loth, E., Happé, F., & Gómez, J. C. (2010). Variety is not the spice of life for people with autism spectrum disorders: Frequency ratings of central, variable and inappropriate aspects of common real-life events. *Journal of Autism and Developmental Disorders*, *40*(6), 730-42. doi:10.1007/s10803-009-0929-7

Louwerse, A., van der Geest, N., Tulen, H. M., van der Ende, J., Van Gool, R., Verhulst, C., et al. (2013). Effects of eye gaze directions of facial images on looking behaviour and autonomic responses in adolescents with autism spectrum disorders. *Research in Autism Spectrum Disorders*, *7*(9), 1043-1053. doi:10.1016/j.rasd.2013.04.013

Maner, J. K., Kenrick, D. T., Becker, D. V., Delton, A. W., Hofer, B., Wilbur, C. J., et al. (2003). Sexually selective cognition: Beauty captures the mind of the beholder. *Journal of Personality and Social Psychology*, *85*(6), 1107-20. doi:10.1037/0022-3514.85.6.1107

Manor, B. R. & Gordon, E. (2003). Defining the temporal threshold for ocular fixation in free-viewing visuocognitive tasks. *Journal of Neuroscience Methods*, *128*(1-2), 85-93. doi:10.1016/S0165-0270(03)00151-1

Markus, H. R. & Kitayama, S. (1991). Culture and the self: Implications for cognition, motivation, and emotion. *Psychological Review*, *98*(2), 224-253.

Martin, W. & Rovira, M. (1982). Response biases in eye-gaze perception. *Journal of Psychology: Interdisciplinary and Applied*, *110*(2), 203-209.

McPartland, J. C., Webb, S. J., Keehn, B., & Dawson, G. (2011). Patterns of visual attention to faces and objects in autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *41*(2), 148-57. doi:10.1007/s10803-010-1033-8

Moser, E., Derntl, B., Robinson, S., Fink, B., Gur, R. C., & Grammer, K. (2007). Amygdala activation at 3T in response to human and avatar facial expressions of emotions. *Journal of Neuroscience Methods*, *161*(1), 126-33. doi:10.1016/j.jneumeth.2006.10.016

Myowa-Yamakoshi, M., Tomonaga, M., Tanaka, M., & Matsuzawa, T. (2003). Preference for human direct gaze in infant chimpanzees (pan troglodytes). *Cognition*, *89*(2), 53-64.

Nation, K. & Penny, S. (2008). Sensitivity to eye gaze in autism: Is it normal? Is it automatic? Is it social?. *Development and Psychopathology*, *20*(1), 79-97. doi:10.1017/S0954579408000047

Nisbett, R. E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic versus analytic cognition. *Psychological Review*, *108*(2), 291-310.

Nittono, H. & Wada, Y. (2009). Gaze shifts do not affect preference judgements of graphic patterns. *Perceptual and Motor Skills*, *109*, 79-94.

O'Doherty, J., Winston, J., Critchley, H., Perrett, D., Burt, D. M., & Dolan, R. J. (2003). Beauty in a smile: The role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia*, *41*(2), 147-155.

Petty, R. E. & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.), *Advances in experimental social psychology.* (pp. 123-205). New York: Academic Press.

Pitskel, N. B., Bolling, D. Z., Hudac, C. M., Lantz, S. D., Minshew, N. J., Vander Wyk, B. C., et al. (2011). Brain mechanisms for processing direct and averted gaze in individuals with autism. *Journal of Autism and Developmental Disorders.* doi:10.1007/s10803-011-1197-x

Qian, N. & Lipkin, R. M. (2011). A learning-style theory for understanding autistic behaviors. *Frontiers in Human Neuroscience*, *5*, 1-17. doi:10.3389/fnhum.2011.00077

Richer, J. M. & Coss, R. G. (1976). Gaze aversion in autistic and normal children. *Acta Psychiatrica Scandinavica*, *53*(3), 193-210.

Riedl, R., Mohr, P., Kenning, P., Davis, F. D., & Heekeren, H. (2011). Trusting humans and avatars: Behavioral and neural evidence. In *Proceedings of the 32nd international conference on information systems (ICIS).*

Ristic, J., Mottron, L., Friesen, C. K., Iarocci, G., Burack, J. A., & Kingstone, A. (2005). Eyes are special but not for everyone: The case of autism. *Brain Research. Cognitive Brain Research*, *24*(3), 715-718. doi:10.1016/j.cogbrainres.2005.02.007

Rutherford, M. D. & Towns, A. M. (2008). Scan path differences and similarities during emotion perception in those with and without autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *38*(7), 1371-81. doi:10.1007/s10803-007-0525-7

Sanefuji, W. & Ohgami, H. (2011). Imitative behaviors facilitate communicative gaze in children with autism. *Infant Mental Health Journal*, *32*(1), 134-142. doi:10.1002/imhj.20287

Schilbach, L., Eickhoff, S. B., Cieslik, E. C., Kuzmanovic, B., & Vogeley, K. (2012). Shall we do this together? Social gaze influences action control in a comparison group, but not in individuals with high-functioning autism. *Autism, 16*(2), 151-162. doi:10.1177/1362361311409258

Senju, A. & Johnson, M. H. (2009). The eye contact effect: Mechanisms and development. *Trends in Cognitive Sciences*, *13*(3), 127-34. doi:10.1016/j.tics.2008.11.009

Senju, A., Hasegawa, T., & Tojo, Y. (2005). Does perceived direct gaze boost detection in adults and children with and without autism? The stare-in-the-crowd effect revisited. *Visual Cognition*, *12*(8), 1474-1496. doi:10.1080/13506280444000797

Senju, A., Kikuchi, Y., Hasegawa, T., Tojo, Y., & Osanai, H. (2008). Is anyone looking at me? Direct gaze detection in children with and without autism. *Brain and Cognition*, *67*(2), 127-139. doi:10.1016/j.bandc.2007.12.001

Senju, A., Yaguchi, K., Tojo, Y., & Hasegawa, T. (2003). Eye contact does not facilitate detection in children with autism. *Cognition*, *89*(1), 43-51.

Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience*, *6*(12), 1317-22. doi:10.1038/nn1150

Simion, C. & Shimojo, S. (2007). Interrupting the cascade: Orienting contributes to decision making even in the absence of visual stimulation. *Perception & Psychophysics*, *69*(4), 591-5.

Spezio, M. L., Adolphs, R., Hurley, R. S., & Piven, J. (2007). Analysis of face gaze in autism using "bubbles". *Neuropsychologia*, *45*(1), 144-51. doi:10.1016/j.neuropsychologia.2006.04.027

Stewart, M. E., Barnard, L., Pearson, J., Hasan, R., & O'Brien, G. (2006). Presentation of depression in autism and Asperger syndrome: A review. *Autism*, *10*(1), 103-116. doi:10.1177/1362361306062013

Strick, M., Holland, R. W., & van Knippenberg, A. (2008). Seductive eyes: Attractiveness and direct gaze increase desire for associated objects. *Cognition*, *106*(3), 1487-96. doi:10.1016/j.cognition.2007.05.008

Suchow, J. W. & Alvarez, G. A. (2011). Motion silences awareness of visual change. *Current Biology*, *21*(2), 140-143. doi:10.1016/j.cub.2010.12.019

Swanson, M. R. & Siller, M. (2013). Patterns of gaze behavior during an eye-tracking measure of joint attention in typically developing children and children with autism spectrum disorder. *Research in Autism Spectrum Disorders*, *7*(9), 1087-1096.

Swettenham, J., Plaisted, K., Milne, E., Campbell, R., Rosenfelder, D., & Coleman, M. (2001). Children with autism show subtle difficulties detecting eye direction. *Paper Presented at the Annula Meeting of the Cognitive Neuroscience Society, New York, NY.*

Tewes, U. (1994). *Hamburg-Wechsler-Intelligenztest für Erwachsene, Revision 1991.* Bern: Huber.

Theodoridou, A., Rowe, A. C., Penton-Voak, I. S., & Rogers, P. J. (2009). Oxytocin and social perception: Oxytocin increases perceived facial trustworthiness and attractiveness. *Horm Behav*, *56*(1), 128-32. doi:10.1016/j.yhbeh.2009.03.019

Vishwanath, A. (2003). Comparing online information effects: A cross-cultural comparison of online information and uncertainty avoidance. *Communication Research*, *30*(6), 579-598. doi:10.1177/0093650203257838

Vishwanath, A. (2004). Manifestations of interpersonal trust in online interaction A cross-cultural study comparing the differential utilization of seller ratings by ebay participants in canada, france, and germany. *New Media & Society*, *6*(2), 219-234.

Vogeley, K. & Bente, G. (2010). "Artificial humans": Psychology and neuroscience perspectives on embodiment and nonverbal communication. *Neural Networks, 23*(8-9), 1077-1090. doi:10.1016/j.neunet.2010.06.003

Von Aster, M., Neubauer, A., & Horn, R. (2006). *Wechsler Intelligenztest für Erwachsene (WIE). Deutschsprachige bearbeitung und adaption des WAIS-III von david wechsler*. Frankfurt: Harcourt Test Services.

Wallace, S., Coleman, M., Pascalis, O., & Bailey, A. (2006). A study of impaired judgment of eye-gaze direction and related face-processing deficits in autism spectrum disorders. *Perception*, *35*(12), 1651-1664. doi:10.1068/p5442

Webster, S. & Potter, D. D. (2011). Gaze perception develops atypically in children with autism. *Child Development Research*, *2011,* 1-9.

Wheelwright, S., Baron-Cohen, S., Goldenfeld, N., Delaney, J., Fine, D., Smith, R., et al. (2006). Predicting autism spectrum quotient (AQ) from the systemizing quotient-revised (SQ-R) and empathy quotient (EQ). *Brain Research*, *1079*(1), 47-56. doi:10.1016/j.brainres.2006.01.012

White, S., Hill, E., Winston, J., & Frith, U. (2006). An islet of social ability in Asperger syndrome: Judging social attributes from faces. *Brain and Cognition*, *61*(1), 69-77. doi:10.1016/j.bandc.2005.12.007

Willis, J. & Todorov, A. (2006). First impressions. *Psychological Science*, *17*(7), 592-598. doi:10.1111/j.1467-9280.2006.01750.x

Wilms, M., Schilbach, L., Pfeiffer, U., Bente, G., Fink, G. R., & Vogeley, K. (2010). It's in your eyes--using gaze-contingent stimuli to create truly interactive paradigms for social cognitive and affective neuroscience. *Social Cognitive and Affective Neuroscience*, *5*(1), 98-107. doi:10.1093/scan/nsq024

Winston, J. S., O'Doherty, J., Kilner, J. M., Perrett, D. I., & Dolan, R. J. (2007). Brain systems for assessing facial attractiveness. *Neuropsychologia*, *45*(1), 195-206. doi:10.1016/j.neuropsychologia.2006.05.009

Yamagishi, T. & Yamagishi, M. (1994). Trust and commitment in the united states and japan. *Motivation and Emotion*, *18*(2), 129-166.

Zajonc, R. B. (2001). Mere exposure: A gateway to the subliminal. *Current Directions in Psychological Science*, *10*(6), 224. doi:10.1111/1467-8721.00154