SOCIAL NEUROSCIENCE, 2016
http://dx.doi.org/10.1080/17470919.2016.1160953

Routledge
Taylor & Francis Group

Beliefs about human agency influence the neural processing of gaze during joint attention

Nathan Caruana[a,b,c], Peter de Lissa[a,b,d] and Genevieve McArthur[a,b]

[a]Department of Cognitive Science, Macquarie University, Sydney, Australia; [b]ARC Centre of Excellence in Cognition and its Disorders, Sydney, Australia; [c]Perception in Action Research Centre, Sydney, Australia; [d]Department of Psychology, Macquarie University, Sydney, Australia

ABSTRACT

The current study measured adults' P350 and N170 ERPs while they interacted with a character in a virtual reality paradigm. Some participants believed the character was controlled by a human ("avatar" condition, n = 19); others believed it was controlled by a computer program ("agent" condition, n = 19). In each trial, participants initiated joint attention in order to direct the character's gaze toward a target. In 50% of trials, the character gazed toward the target (congruent responses), and in 50% of trials the character gazed to a different location (incongruent response). In the avatar condition, the character's incongruent gaze responses generated significantly larger P350 peaks at centro-parietal sites than congruent gaze responses. In the agent condition, the P350 effect was strikingly absent. Left occipitotemporal N170 responses were significantly smaller in the agent condition compared to the avatar condition for both congruent and incongruent gaze shifts. These data suggest that beliefs about human agency may recruit mechanisms that discriminate the social outcome of a gaze shift after approximately 350 ms, and that these mechanisms may modulate the early perceptual processing of gaze. These findings also suggest that the ecologically valid measurement of social cognition may depend upon paradigms that simulate genuine social interactions.

ARTICLE HISTORY
Received 7 September 2015
Revised 26 February 2016
Published online
21 March 2016

KEYWORDS
Joint attention; eye gaze; virtual reality; social interaction; agency

Humans are skilled in extracting social signals conveyed by another's gaze during interactions. We use gaze to understand the emotions and intentions of others, and to coordinate joint attention experiences with them (i.e., a common focus of attention). Joint attention is critical for the development of language and social learning (Adamson, Bakeman, Deckner, & Romski, 2009; Baron-Cohen, 1995; Charman, 2003; Mundy, Sullivan, & Mastergeorge, 2009; Murray et al., 2008; Tomasello, 1995); and impairments in joint attention constitute one of the most homogenous characteristics of autism (Pelphrey, Shultz, Hudac, & Vander Wyk, 2011). Thus, understanding the neural processing of gaze during joint attention interactions is important for understanding both typical and atypical development.

Unfortunately, this understanding has been hindered by the challenge of developing neurophysiological paradigms that can simulate an ecologically valid interaction whilst simultaneously ensuring tight experimental control (Schilbach et al., 2013). Many social neuroscientists had tackled this challenge by integrating virtual reality characters with functional magnetic resonance imaging (fMRI; Caruana, Brock, & Woolgar, 2015; Pfeiffer et al., 2014; Schilbach et al., 2010) and event-related potentials (ERPs; Caruana, de Lissa, & McArthur, 2015) because virtual characters can be fully controlled and realistically convey anthropomorphic form and behavior (see Georgescu, Kuzmanovic, Roth, Bente, & Vogeley, 2014 for review).

The increased use of virtual characters in social neuroscience raises an important empirical and methodological question: *Is it important for participants to believe that a virtual character is being controlled by a real human?* This is typically achieved by deceiving participants, which introduces practical and ethical issues into an experiment. In order to determine if such issues are justified (i.e., to inform best practice), the current study investigated the effect that beliefs of human agency have on gaze-related neural processes during joint attention interactions with a virtual partner.

The influence of agency beliefs on subjective experience

Agency beliefs refer to the extent to which an individual believes a stimulus to represent the online

CONTACT Nathan Caruana nathan.caruana@mq.edu.au Level 3, 16 University Avenue, Macquarie University, 2109 Sydney, NSW, Australia

behavior and intentions of another human. Virtual reality studies have begun to distinguish between virtual characters that are believed to be computer-controlled "agents" or human-controlled "avatars." This distinction was first made by von der Pütten, Krämer, Gratch, and Kang (2010), who investigated the influence of agency beliefs on subjective experience and self-disclosure during one-way conversational interactions with agents and avatars. During these interactions, participants were asked to respond to a series of questions asked by a virtual character. The presence of contingent head-nods made by the virtual character resulted in an increase in participants' self-disclosure, and a reduction of low-dominance negative feelings (e.g., weak, shy) measured using the Positive and Negative Affect Schedule (Watson, Clark, & Tellegen, 1988). Participants also reported less negative low-dominance feelings (e.g., scared, ashamed) when they believed the virtual character to be an avatar rather than an agent. Nevertheless, the authors concluded that participants' experience with the virtual character was mostly influenced by the perceived realism of the stimulus rather than beliefs about whether the character was controlled by another human. This interpretation seems at odds with the data since (1) the belief of whether the virtual character was an avatar or agent *was* found to influence subjective experience on one outcome measure, and (2) perceived realism was measured offline after the virtual interaction was over using self-disclosure, which is heavily influenced by personality traits (e.g., extraversion; Peter, Valkenburg, & Schouten, 2005). Furthermore, these data do not elucidate whether agency beliefs influence gaze-related neural processes during joint attention interactions.

The influence of agency beliefs on gaze processing

Pfeiffer et al. (2014) were the first group to provide evidence that human agency beliefs influence the neural processing of social stimuli during virtual interactions. In their fMRI study, they asked participants to interact with an anthropomorphic virtual character. Participants were instructed that in some trials the virtual character's gaze would be controlled by a computer program, and in other trials, it would be controlled by another human using the online recordings of their eye movements. On each trial, participants initiated a joint attention bid by fixating on one of two squares located on either side of the virtual character's face. The virtual character averted his gaze to look at the same square (a congruent response that achieved joint attention) or at the alternate square (an incongruent response that avoided joint attention). Each block comprised five trials. The "congruency" of each block was manipulated by adjusting the proportion of congruent trials and incongruent trials. At the end of each block, participants decided whether they believed that the virtual character was an avatar or agent. The authors reported that participants were more likely to believe that the avatar was controlled by a human in blocks where the virtual character responded congruently more often to achieve joint attention. Blocks in which participants believed that they interacted with a human were associated with greater activation of the ventral striatum than blocks in which participants believed that they were interacting with a computer-controlled agent. Schilbach et al. (2010) suggested that this effect reflected the hedonic experience of achieving a self-initiated joint attention bid. However, given that beliefs about human agency were clearly influenced by the congruency of the virtual character's response, this cannot be interpreted as direct evidence for the influence of human agency beliefs on the neural processing of gaze during joint attention interactions.

More direct evidence for the effect of human agency beliefs on the neural processing of gaze comes from two ERP studies. Wykowska, Wiese, Prosser, Müller, and Hamed (2014) asked participants to complete a gaze-cueing task in which they were presented with a pair of eyes imbedded in a robot face. On each trial, participants were asked to use a button box to identify the location of a target presented to the left or right side of the robot's face. The target was preceded by a valid gaze cue (i.e., the robot shifted its gaze toward the target location) or an invalid gaze cue (i.e., the robot looked in the opposite direction to the target location). On some trials, participants were told that the robot was controlled by a human; on other trials, they were told that the robot was pre-programmed.

P1 ERP responses were measured at posterior-occipital sites 100–140 ms after the onset of the target. The P1 is a positive ERP peak that is believed to reflect neural processes associated with visual attention (Itier & Taylor, 2004a). The authors reported that P1 responses to targets were significantly larger following the presentation of valid gaze cues than invalid gaze cues. However, this effect was only observed when participants believed the robot to be controlled by a human, and not when they believed it to be pre-programmed. This outcome is particularly striking given that (1) agency beliefs were only manipulated via instruction (i.e., the tasks were identical), and (2) this instruction was irrelevant to the task. The authors explained this effect within the Intentional Stance Model of Social Cognition, suggesting that the perception of agency recruits the neural mechanisms that support mentalizing (i.e., the cognitive ability to understand the mental states of others). These mentalizing mechanisms are argued to have a top–down "sensory gain" effect on visual processes (Wykowska et al., 2014).

While Wykowska et al.'s (2014) findings certainly support the idea that agency beliefs influence gaze-related effects on visual attention, they do not elucidate whether agency beliefs influence the perceptual processing of gaze shifts specifically. The P1 was time-locked to the presentation of the target that appeared a long time after the presentation of the gaze cue (i.e., 600 ms). It is more likely that the P1 reflected processing of the target than the processing of the preceding gaze cue. Thus, the P1 effect reported by Wykowska et al. may not provide a direct measure of the influence that agency beliefs have on the neural processing of gaze during social interactions.

The occipitotemporal N170, a negative brain potential peaking approximately 170 ms after stimulus presentation, is believed to provide a more sensitive measure of gaze processing than the P1 component. It has been found to be most sensitive to faces and eyes in comparison to inanimate objects (Itier & Taylor, 2004a), and it is thought to reflect the earliest structural processing of faces (Ganis et al., 2012). Gaze processing studies have found that the amplitude of the N170 response is influenced by whether gaze is averted or directed at participants, although the direction of this effect is inconsistent (see Itier & Batty, 2009 for review). Interestingly, Pönkänen, Alhoniemi, Leppänen, and Hietanen (2010) found that viewing direct gaze in a live-viewing condition elicited a larger N170 response than averted gaze or closed eyes. This effect was not found when the same faces were viewed as photographs on a computer screen. Pönkänen et al. suggested that gaze may be processed more "intensely" when it is believed to convey the current perspective, intentions, and agency of another person in real time (p. 486). However, this study did not employ a joint attention paradigm, and hence provides little insight into the effect of agency beliefs on the neural processing of gaze during joint attention interactions.

Carrick, Thompson, Epling, and Puce (2007) also examined the neural processing of gaze using ERPs. Participants were presented with trials that comprised three horizontally aligned faces (a central face and two flanker faces). The gaze of both flanker faces were directed either to the left or to the right. The gaze of the central face, which was initially directed toward the participant, was updated to match the flanker faces (the "group" condition), to face toward one flanker face (the "mutual" condition), or to gaze upwards away from both flanker faces (the "avoid" condition). The onset of the updated central face generated N170 responses measured at occipitotemporal sites. These responses were not modulated by the social significance of the gaze-shift. In contrast, gaze shifts in the group and mutual conditions generated earlier P350 and smaller P500 peaks relative to the avoid condition. Carrick et al. concluded that the P350 and P500 peaks reflected the integration of the spatial properties of gaze in order to evaluate its social significance within the depicted social interaction.

The influence of agency beliefs on gaze processing during social interactions

While previous ERP studies have done well to employ sensitive measures of the neural processing of gaze, they were not designed to investigate the effect of agency beliefs on processing another person's gaze in the context of a social interaction that involves the participant. To this end, we developed a novel virtual reality paradigm in a previous study to investigate the time course of neural processes associated with evaluating self-initiated joint attention bids (Caruana, de Lissa, et al., 2015). We used this paradigm to measure participants' ERPs while they interacted with a virtual character whom they believed was an avatar controlled by a human in a nearby laboratory via live infrared eye-tracking (the "social" condition). Unbeknownst to participants, the virtual character was controlled by a gaze-contingent algorithm. On each trial, participants initiated joint attention toward a task-relevant target. The virtual partner responded by gazing congruently toward the target (achieving joint attention) or incongruently toward one of the remaining onscreen targets (avoiding joint attention). The ERP data revealed that incongruent gaze shifts made by the virtual partner elicited a significantly larger mean centro-parietal P350 ERP than congruent gaze shifts. The same effect was not observed in a "non-social" control condition that superimposed computer-controlled arrows over the closed eyes of the virtual character. Additionally, this effect was not observed in the N170 data. These data, which are consistent with the findings of Carrick et al. (2007), suggest that the P350 ERP is triggered by neural processes associated with evaluating the social outcome of a gaze cue—in this case—whether or not joint attention has been achieved. Specifically, we believe the larger P350 ERPs observed following incongruent gaze shifts may reflect the additional neural effort required to process a social partner's current focus of attention when they gaze toward an unexpected location. However, it is not clear from this study if the absence of the P350 effect in the non-social control condition resulted from perceptual differences between eyes and arrows, or from participants' lack of belief that they were interacting with another human.

In the current study, we investigated whether the P350 effect identified by Caruana, de Lissa, et al. (2015) was present when participants believed that the virtual character was a computer-programmed agent rather than a human-controlled avatar. Data from Caruana, de Lissa, et al. (2015) social condition became the "avatar" condition in the current study, and we recruited a new group of individuals to participate in an "agent" condition. Consistent with the claim that the P350 ERP represents a process of evaluating the social significance of a gaze shift (e.g., whether joint attention has been achieved), we anticipated that participants in the agent condition would show a significantly reduced P350 effect (i.e., a larger P350 to incongruent gaze shifts than congruent gaze shifts) in the agent condition compared to the avatar condition. We also predicted that smaller occipitotemporal N170 responses would be elicited when participants observed gaze shifts believed to be controlled by a computer agent than a human-operated avatar, irrespective of stimulus congruency. This is consistent with previous studies that have found the gaze-related N170 to be sensitive to human agency beliefs (Pönkänen et al., 2010) but not the social outcome of a gaze shift (Caruana, de Lissa, et al., 2015).

Method

The methods used in this study were approved by the Macquarie University Human Research Ethics Committee.

Participants

This study used an independent-groups design that included two group conditions ("avatar" versus "agent") in which participants responded to two conditions of stimuli ("congruent" versus "incongruent"). Participants volunteered or received course credit for their time and provided written consent before participating.

In the avatar condition, 24 individuals completed the task under the instruction that the virtual character was being controlled by a human partner named "Alan." Participants were instructed that Alan would be interacting with them from a nearby eye-tracking laboratory. The data from two individuals could not be used due to unreliable eye tracking calibration. Another two participants did not believe that the virtual character was being controlled by a human. The behavioral data from a fifth participant also indicated that they were not attending to the virtual character's gaze shifts. This participant was also excluded from analyses (see Behavioral data in Results). This resulted in a final sample of 19 participants (3 male, M_{age} = 20.95, SD = 5.78) for the avatar condition.

In the agent condition, a separate group of 19 individuals (3 male, M_{age} = 23.21, SD = 6.49) completed the same task except that they were instructed that the virtual character was a computer-controlled agent. No participants were excluded from the analyses given that reliable behavioral and eye-tracking data was obtained for all individuals.

Stimuli

An anthropomorphic virtual character was animated using *FaceGen* (Singular Inversions, 2008). The animated face subtended 8 × 12 degree visual angle and was presented in the center of the screen (a 60 × 34 cm Samsung SynchMaster SA950 HD LED monitor with a refresh rate of 120 Hz) at a distance of 65 cm from the participant. Five face stimuli were generated in which the eyes were either directed at the participant or toward the top-left, top-right, bottom-left, or bottom-right corner of the screen. Each corner of the screen contained a cartoon building. These buildings were identical and animated using *GIMP-2* (Kimball & Mattis, 1995). Each building subtended 11° visual angle. There were 15° visual angle separating the virtual character's eyes and each building. *Experiment Builder* 1.10.165 (SR Research, 2004) was used to program the gaze-contingent algorithm and present the stimuli.

Stimulus conditions

We employed the same virtual reality paradigm developed and used in a previous study (Caruana, de Lissa, et al., 2015). A gaze-contingent algorithm was used to simulate a live interaction between the participant and an onscreen virtual character. Participants believed that the virtual character was controlled by a human partner (avatar condition) or a computer program (agent condition). The tasks completed by participants in the avatar and agent conditions were identical.

Participants were instructed to play a cooperative game with their virtual partner called "Catch the Prisoner." The task was to catch a prisoner who, on each trial, attempted to escape from one of the four prison exits. Participants were told that they would play the role of "watch person" while their virtual partner would play the "guard." The watch person's task was to monitor the outside of the prison while the guard's task was to monitor the inside of the prison. The watch person was required to inform the guard if a prison exit was breached via initiating joint attention with the guard. A prisoner would be caught if the guard responded congruently to this joint attention bid. However, participants were told that sometimes the guard would respond incongruently to a joint attention bid (and hence the prisoner would escape) because inmates fighting inside the prison distracted him.

At the beginning of each trial, a crosshair was presented in the center of the screen subtending 1.4° visual angle. Once the participant (i.e., watch person) fixated for a minimum of 150 ms on the crosshair, it was replaced by an anthropomorphic face of a virtual character (i.e., the guard) with the nasion in the same location as the crosshair. At the same time, four cartoon buildings were displayed in each corner of the screen depicting the prison exits (see Figure 1). After a delay of 200–1000 ms (jittered with a random distribution), a yellow circle (depicting a sensor light that could not be seen by the guard) was presented above one of the exits. The participant was required to look at the spotlight for a minimum of 150 ms. If this was done correctly, after a further delay of 200–1000 ms (jittered with a random distribution), a prisoner appeared at the exit. The participant was then required to initiate joint attention with the guard. To this end, the participant was required to fixate back on the virtual character's face for at least 150 ms. If this was also done correctly, the guard's gaze shifted after 350–650 ms. This delay provided enough time for an N170 to be generated but was short enough so that the virtual character's response did not appear unrealistically sluggish. On 50% of trials, the guard's gaze shifted to the correct location (i.e., the escaping prisoner) to achieve joint attention (congruent trials). On the remaining trials, the guard shifted his gaze to one of the remaining three locations (incongruent trials).

Participants completed four blocks, each comprising 60 trials. Trials containing congruent and incongruent gaze shifts were presented in random order across blocks. The direction of congruent and incongruent gaze shifts were fully counterbalanced across all trials. Thus, on incongruent trials, the guard was equally likely to gaze toward one of the three buildings not fixated by the participant.

To ensure that participants learned how to engage with the virtual interface appropriately, they received negative feedback (i.e., text reading "Bad Fix" presented in the center of the screen) if they (1) failed to fixate the spotlight, (2) fixated away from the spotlight before the prisoner appeared, (3) did not fixate back on the guard's face within 3000 ms of the prisoner's appearance, or (4) fixated on the guard's face for less than 1000 ms after fixating the target. This also ensured that participants remained fixated on the guard's face during the interval that gaze-related ERPs were being measured.

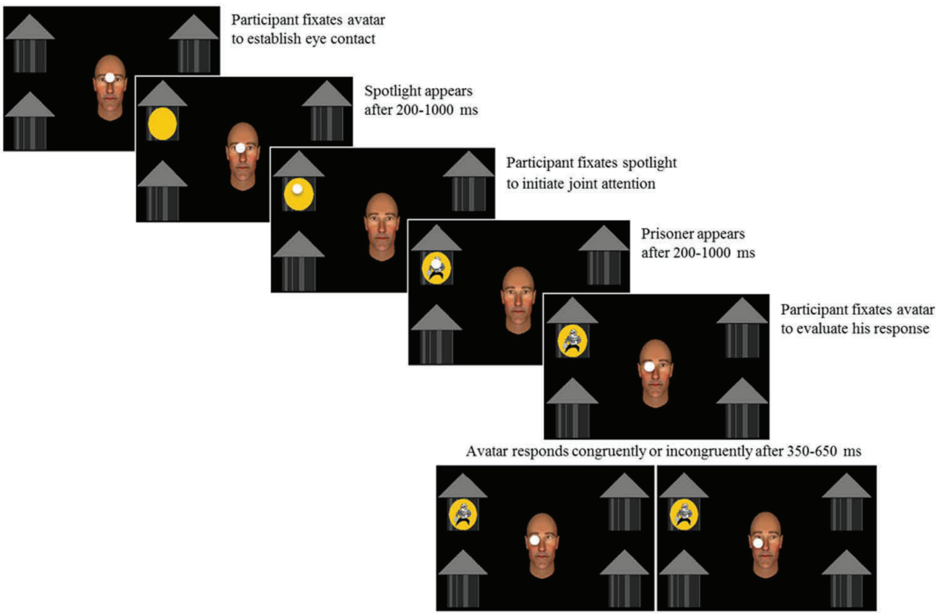

Figure 1. Schematic representation of trial sequence. White circle represents the location of the participant's gaze and was not part of the stimuli visible to the participant.

Eye movement and electroencephalogram (EEG) recording

An EyeLink 1000 monocular tower-mounted eye tracker was used to record the eye movements of each participant's right eye. Heads were stabilized using a chin rest, and eye movements were sampled at 1000 Hz. The online EEG of each participant was recorded using a Synamps II amplifier with a sampling rate 1000 Hz, an online band-pass filter of .05–100 Hz, and a notch filter at 50 Hz. A montage of 29 electrodes were positioned according to the 10–20 system (EasyCap; FP1, FP2, F7, F3, FZ, F4, F8, FT7, FC3, FC4, FT8, T7, C3, CZ, CPZ, C4, T8, TP7, CP3, CP4, TP8, P7, P3, PZ, P4, P8, O1, OZ, O2). Online and offline reference electrodes were attached to the left and right earlobes, respectively. The ground electrode was positioned between FP1, FP2, and FZ. Impedances were maintained below 5 kΩ for all electrodes. Bipolar electrodes were positioned at the outer canthi to measure horizontal electro-ocular activity (HEOG), and above and below the left eye to measure vertical ocular activity (VEOG).

Creating ERPs

Neuroscan 4.5 software was used for the offline processing of the EEG data (Neuroscan, El Paso, Texas, U.S.A.). A standard ocular reduction algorithm was used to remove VEOG activity. Corrected data was then band-pass filtered (0.1–30 Hz) with a 12 dB octave roll-off and segmented into epochs that were time-locked to the onset of the virtual character's congruent or incongruent gaze shifts. Epochs comprised a pre-stimulus baseline of −100 to 0 ms and ended 700 ms after the virtual character's gaze shift. Epochs containing voltages exceeding +/−100 mV were removed from further analysis. All epochs retained in the analysis were baseline corrected using the 100 ms of pre-stimulus electrical activity. Each participant's accepted epochs were averaged to produce congruent and incongruent ERPs.

Measuring ERPs

To ascertain the length of the intervals used to measure each ERP in this study (P350 at CZ and PZ, and N170 at P7 and P8), we visually inspected the ERPs of each individual. This revealed that a clear P350 peak could be identified in both conditions for 12 individuals from the avatar group and 9 individuals from the agent group. It also revealed clear N170 peaks for all individuals and conditions measured at P7 and P8. Thus, the P350 was measured using mean amplitude whilst the N170 was measured using peak amplitude.

A 130 ms interval (310–440 ms) captured each individual's P350 peak in both the congruent and incongruent conditions. Thus, we used 130 ms intervals (65 ms either side of the mean peak) to measure the mean amplitude of P350 (310–440 ms) at CZ and PZ and peak amplitude of the N170 (107–237 ms) recorded at P7 and P8.

Subjective experience questionnaire

At the end of the testing session, participants rated various aspects of their experience on a five-point Likert scale (1 = not at all to 5 = extremely). Participants in both the avatar and agent conditions rated how difficult, intuitive, natural, and pleasant the interactive task felt. Following the debriefing, individuals in the avatar condition were asked to rate how convinced they were that a real person controlled the virtual character. Participants in the agent condition rated the extent to which it felt like they were interacting with a human. They also rated how human-like the virtual character appeared and behaved.

Attention to gaze shifts

At the end of each block, participants were asked to estimate how frequently (expressed as a percentage) they successfully caught the prisoner (i.e., the percentage of congruent trials). This provided a measure of task engagement.

Statistical analysis

To ascertain if parametric or non-parametric analyses should be used, we tested whether (1) data sets within groups for each condition were normally distributed (using the Kolmogorov–Smirnov test), and (2) data sets between groups for each condition were equivariant (using the Levene test). The P350 data passed tests for normality and equivariance. The N170 data failed tests for normality and equivariance. The subjective measures failed tests for normality, and all but one measure passed tests for equivariance.

Since parametric tests are robust to moderate violations of normality when comparing equal samples of this size, the effect of group condition (i.e., avatar, agent) on subjective experience ratings were assessed using independent *t*-tests using statistics that did not assume equal variance. The effect of group condition (i.e., avatar, agent) and stimulus condition (i.e., congruent, incongruent) on each ERP measure was assessed using two-way ANOVAS (Statistical Package for the Social Sciences v19). Main effects of group were assessed using independent *t*-tests with statistics that did not assume equal variance when the assumption of equivariance was violated. An α level of $p = 0.05$ was used for all analyses.

Results

Subjective experience questionnaire

Participants from both the avatar and agent conditions rated the interactive task as easy and intuitive. They also rated the interaction as feeling moderately natural. There were no significant differences between group conditions on these dimensions (independent *t*-tests, all *p*s > 0.15). However, participants in the avatar condition did rate the interactive task as significantly less pleasant than participants in the agent condition ($t(26.77) = 2.21$, $p = .036$). Many participants in the avatar condition also explained that they felt frustrated with their partner when he did not respond to their joint attention bid. Thus, the reduced pleasantness ratings provided by participants in this group may be associated with the frustration they felt toward their partner. Likewise, it is possible that participants in the agent condition found the task less frustrating, and thus more pleasant because they knew they were interacting with a computer-programmed agent whom they might expect to behave in a less predictable or cooperative way. The descriptive statistics for these subjective ratings are summarized in Table 1.

Avatar condition. All participants in the avatar condition provided ratings that confirmed that they were convinced (5 = completely convinced) that the virtual character was an avatar controlled by a real person ($M = 4.89$, SD = 0.32). Those who provided a 4/5 rating claimed that they only momentarily thought it possible that they were interacting with a computer-controlled agent, and took the interaction for granted.

Agent condition. Participants in the agent condition also rated their level of engagement with the virtual character. On average they reported that the interaction felt moderately human-like ($M = 2.53$, SD = 1.02) and that the virtual character appeared ($M = 3.58$, SD = 0.84) and behaved ($M = 3.47$, SD = 0.96) very human-like.

Table 1. Ratings on subjective experience questionnaire.

	Avatar	Agent
Task Aspect	*M* (SD)	*M* (SD)
Difficulty	1.68 (0.75)	1.68 (0.86)
Intuitiveness	4.32 (0.82)	4.32 (0.89)
Naturalness	2.58 (1.26)	3.16 (1.17)
Pleasantness of task*	3.11 (1.66)	4.05 (0.85)

Ratings provided on a 5-point scale (1 = low, 5 = high). *denotes a significant difference between groups (avatar versus agent).

Attention to gaze shifts

At the end of each block, participants estimated the percentage of trials that the virtual character responded congruently. One participant was excluded from the avatar condition (see Participants in Method) because they provided an average estimate that was two standard deviations above the group mean ($M = 87.00$). The average congruency estimates in the final samples for the avatar ($M = 48.33\%$, SD = 11.05) and agent ($M = 48.86\%$, SD = 77.79) conditions accurately reflected the 50% congruency manipulation employed in the current study, suggesting that participants were attending to the virtual character's gaze shifts throughout the task.

ERPs

Summary statistics for the mean amplitude measures are shown in Table 2. Group average waveforms comprising the P350 at CZ and PZ are shown in Figure 2, and for the N170 at P7 and P8 are shown in Figure 3. Topographic maps highlighting differences in electrical activity at the scalp between the congruent and incongruent conditions are depicted for each group condition (i.e., avatar and agent) in Figure 4.

P350 mean amplitude

There was a main effect of group measured at both CZ ($F(1,36) = 9.492$, $p = .004$), and PZ ($F(1,36) = 15.492$, $p < .005$) since the P350 generated in the avatar condition was significantly larger than the agent condition. There was also a main effect of condition measured at CZ ($F(1,36) = 17.605$, $p < .005$), and PZ ($F(1,36) = 7.790$, $p = .008$) because the P350 was significantly larger on incongruent trials than congruent trials. Most importantly, there was a significant interaction between group and condition at CZ ($F(1,36) = 12.739$, $p = .001$) and PZ ($F(1,36) = 7.272$, $p = .001$) because the difference between the P350 in the incongruent and congruent conditions was larger in the avatar condition [CZ: (t (18) = 4.798, $p < .001$; PZ: (t (18) = 3.425, $p = .003$] than in the agent condition [(CZ: (t (18) = .533, $p = .600$); PZ: (t (18) = .079, $p = .938$].

N170 peak amplitude

There was a main effect of group at P7 ($F(1,36) = 5.10$, $p = .030$) since the N170 generated in the avatar

Table 2. Summary statistics for amplitude and latency measures by electrode.

	Congruent	Incongruent	Congruent	Incongruent
	CZ		PZ	
P350 mean amplitude				
Avatar	11.85(4.60)	14.90(5.45)	10.52(4.15)	12.53(4.53)
Agent	8.44(4.52)	8.69(5.23)	6.43(3.48)	6.46(4.31)
	P7		P8	
N170 peak amplitude				
Avatar	−7.42(4.59)	−7.24(4.27)	−9.79(5.48)	−9.89(5.05)
Agent	−4.72(2.27)	−4.93(2.00)	−7.53(5.74)	−7.47(5.55)

Summary statistics are provided in the format of *M*(SD).

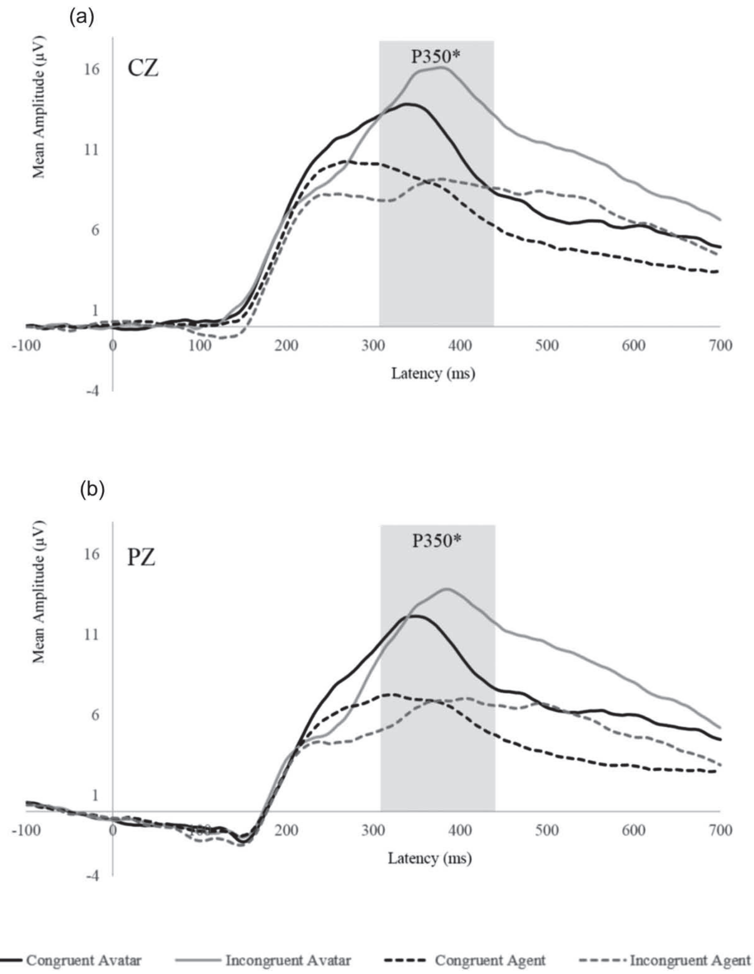

Figure 2. Group average waveforms comprising the P350 at (a) Cz and (b) Pz electrodes. Epochs were time-locked to the onset of the virtual character's gaze shift.

condition was significantly larger than the agent condition. This was also significant when assessed using an independent *t*-test not assuming equal variance ($t(25.63) = 2.258$, $p = .033$). A main effect of group was not significant at P8 ($F(1,36) = 1.763$, $p = .193$). Similarly, there was no significant main effect of condition [(P7: ($F(1,36) = .006$, $p = .937$); P8: ($F(1,36) = .008$, $p = .931$)], and no significant group*condition interaction when measured at either P7 ($F(1,36) = 0.892$, $p = .351$) or P8 ($F(1,36) = 0.160$, $p = .692$).

Discussion

The aim of the current study was to determine if agency beliefs influence neural processes associated with evaluating the achievement of joint attention during gaze-based social interactions. We predicted that the centro-parietal P350 effect, previously identified by Caruana, de Lissa, et al. (2015), would be significantly larger in participants who believed that they were interacting with a human than those who believed that they were

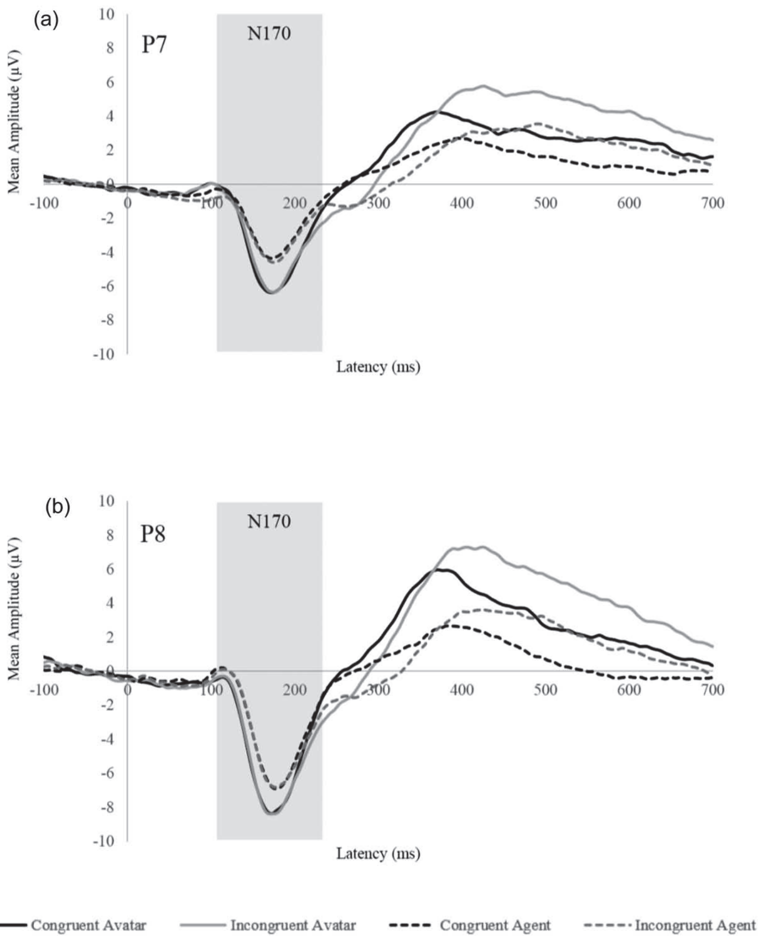

Figure 3. Group average waveforms comprising the N170 at (a) P7 and (b) P8 electrodes. Epochs were time-locked to the onset of the virtual character's gaze shift.

interacting with a computer. We also predicted that a significantly larger N170 would be evoked by gaze shifts believed to be made by a human than a computer, regardless of whether the gaze shift resulted in joint attention or not. Consistent with these predictions, the centro-parietal P350 ERP only differed between congruent and incongruent gaze shifts when individuals believed the virtual character to be controlled by a human avatar. In addition, larger occipitotemporal N170 responses were observed in individuals who believed that they were interacting with a human than individuals who believed that they were interacting with a computer. Taken together, these data suggest that agency beliefs influence the neural processing of social signals conveyed by virtual characters.

The influence of agency beliefs on processing the social outcome of gaze shifts

The P350 ERP effect measured at centro-parietal electrodes was only observed in the avatar condition in participants who believed that a human controlled the virtual character's gaze. This centro-parietal P350 effect has been previously associated with evaluating the social significance of a gaze shift (Carrick et al., 2007). In the current study, it was specifically associated with evaluating whether a gaze shift signaled the achievement of joint attention with another person (i.e., "Is my partner attending the same thing as me?"). If the P350 ERP truly represents the onset of mentalizing during gaze processing, it makes sense that this effect was not present when the participant did not believe the gaze shift to represent the intentional actions of another human. In line with this expectation, the P350 effect was absent in participants who believed that they were interacting with a computer-programmed agent.

This finding provides support for the social-specificity of the P350 effect that we identified in our earlier work (Caruana, de Lissa, et al., 2015). Given that the P350 effect was only observed in the avatar condition, it is unlikely that it represents an effect of gaze

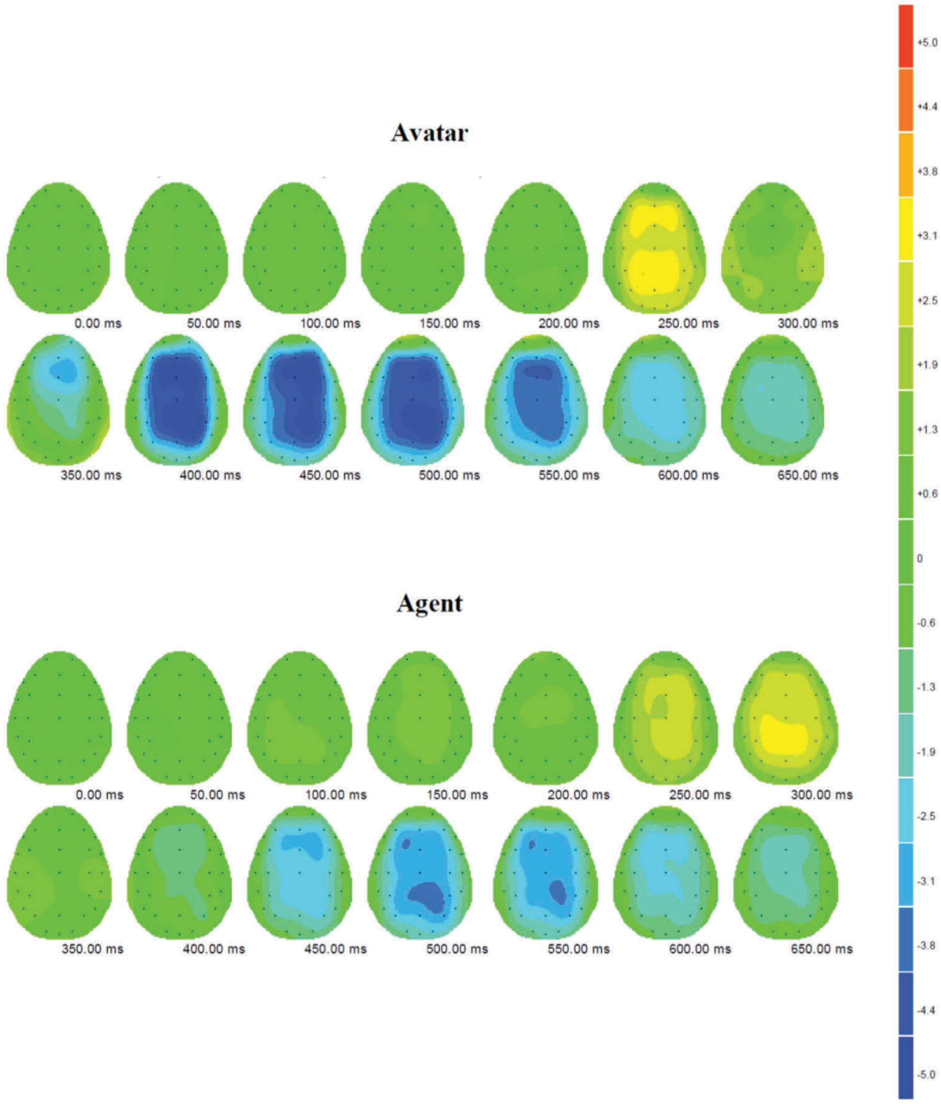

Figure 4. Effect topography maps (Congruent–Incongruent) by group.

congruency on non-social attention mechanisms (e.g., odd-ball, error detection, or attention orienting effects). If this were the case, the P350 effect should have been measured in both the avatar and agent conditions, since both conditions manipulate the spatial properties of gaze in the same way. Thus, the P350 effect seems specific to conditions where participants believe the virtual character is a human-controlled avatar. This may be because the P350 ERP reflects the evaluation of gaze to represent another person's mental perspective.

The influence of agency beliefs on the perceptual processing of gaze shifts

Consistent with previous findings, N170 responses were largest when measured over the right temporoparietal region (see Figure 3; Itier & Taylor, 2004b). Of particular relevance to the current study, and consistent with previous findings using non-interactive paradigms, we also found a significant group effect in which larger N170 responses to gaze shifts were measured in individuals who believed that they were interacting with a human rather than a computer (Pönkänen et al., 2010). Wykowska et al. (2014) have argued that these effects of agency beliefs on the early perceptual processing of gaze-related stimuli may be driven by neural mechanisms of "stimulus gain control." Specifically, neural processing of sensory information may be amplified to increase the signal-to-noise ratio for stimuli that are relevant to the observer's current context. This has been explained using the Intentional Stance Model proposed by Wykowska et al. The authors have argued that the brain takes an "intentional stance" toward stimuli believed to represent a human mind. This involves the recruitment of neural substrates that govern mentalizing processes (e.g., medial prefrontal cortex, and temporoparietal junction). These mentalizing mechanisms may then have a top-down influence on attentional control in the parietal cortex (e.g., intraparietal sulcus) by prioritizing the processing of social stimuli. This results in the enhanced early processing of social stimuli in extrastriate visual areas where the sensory gain effect is measured in occipitotemporal ERPs (e.g., P1 and N170). In the current study, the N170 group effect only reached significance in the left hemisphere. This is consistent with lesion studies that have reported that the left temporoparietal junction is especially important in supporting the ability to evaluate another's mental perspective (Samson, Apperley, Chiavarino, & Humphreys, 2004). There is also evidence that gray matter volume in the left temporoparietal junction is associated with individual differences in anthropomorphism (i.e., the tendency to attribute human agency to non-human phenomenon), suggesting that the region may be involved in supporting beliefs about human agency (Cullen, Kanai, Bahrami, & Rees, 2013).

Whilst the Intentional Stance Model provides a sensible framework for interpreting these converging findings, the evidence supporting the direction of the proposed top–down relationship between mentalizing and early visual perception brain regions remains tentative. Future research integrating neuroimaging techniques that have high temporal (e.g., EEG, MEG) and spatial (e.g., fMRI, PET) resolution are needed to determine whether the Intentional Stance Model provides an accurate account of the mechanisms underlying the influence of agency beliefs on early perceptual processes. Specifically, connectivity analyses and dynamic causal modeling may elucidate whether neural substrates associated with mentalizing modulate occipitotemporal and parietal areas early on in the perceptual processing of gaze shifts. Whilst the current study cannot confirm all of the mechanisms proposed by the Intentional Stance Model, it does corroborate the finding that gaze-related N170 responses are modulated by beliefs of human agency (Pönkänen et al., 2010).

Implications and recommendations

These findings present both methodological and empirical implications for social neuroscience research. First, this study contributes further evidence that beliefs about human agency influence the neural processing of social signals conveyed by virtual characters. This suggests that in order to achieve an ecologically valid simulation of social interactions using virtual characters, participants must believe the virtual character to represent a real human to whom they can attribute mental states. Not only does this match our subjective experience during real social interactions, but this belief is important in engaging the neural processes that support genuine social interactions (e.g., mentalizing). Therefore, our data suggests that the practical and ethical considerations involved in deceiving participants are justified by the importance of this benign deception in supporting the ecological validity of virtual interactions.

Second, the current study provides a new approach for achieving control over the effects that gaze stimuli may have on non-social cognitive processes. Traditionally, gaze-processing studies have relied on arrow stimuli to control for the effects that gaze may have on spatial attention (see Nation & Penny, 2008 for review). However, it is impossible to obtain gaze and arrow stimuli that are perceptually equivalent. This is reflected by the inconsistent cueing effects found in paradigms comparing gaze and arrow cues in behavioral (see Frischen, Bayliss, & Tipper, 2007 for review) and ERP studies (e.g., Feng & Zhang, 2014; Hietanen, Nummenmaa, Nyman, Parkkola, & Hämäläinen, 2006; Holmes, Mogg, Garcia, & Bradley, 2010; Lassalle & Itier, 2013; van Velzen & Eimer, 2003). We suggest that the social-specificity of gaze-related ERP effects can be conservatively determined by manipulating whether the gaze stimuli is believed to be controlled by a human or a computer, rather than manipulating whether spatial information is conveyed by eyes or arrows.

In addition to achieving an ecologically valid measure of the neural processing of gaze shifts, inducing agency beliefs may assist in minimizing the effect that individual differences in anthropomorphism have on gaze-related ERPs. For example, it is possible that the current study observed a P350 effect in the avatar condition but not the agent condition because an agency belief ensures that participants consistently treat the virtual character as a human. Given that individuals differ in their propensity to anthropomorphize non-human stimuli (e.g., Cullen et al., 2013), it is possible that the P350 effect may have been present to different extents in some individuals in the agent condition. That is, some individuals may have been more likely to anthropomorphize the virtual character, resulting in the engagement of spontaneous mentalizing processes. This is consistent with previous findings that individual differences in anthropomorphism are correlated with the size of brain regions associated with mentalizing (Cullen et al.). Whilst future virtual reality studies could employ measures of anthropomorphism as a covariate to account for these individual differences, this source of noise can be effectively minimized by ensuring participants believe the virtual character to be a human-controlled avatar.

Summary

In sum, the current study demonstrates that the neural processing of gaze is sensitive to agency beliefs. This has significant implications for the use of virtual reality as a tool for simulating ecologically valid interactions in social neuroscience research. We found larger left occipitotemporal N170 responses to gaze shifts in individuals who believed the virtual character's gaze shift to be controlled by a human rather than a computer. This suggests that agency beliefs may have a top–down influence on the early perceptual processing of gaze. Furthermore, we found that a centro-parietal peak differentiated gaze shifts that signaled the success or failure of a self-initiated joint attention bid after approximately 350 ms. This P350 effect was only observed in individuals who believed the virtual character to be operated by a human. These data support the claim that the brain decodes information about whether a human's focus of attention is the same or different to our own approximately 350 ms after the observation of a gaze shift. Thus, the P350 may provide a useful neural marker for evaluating the achievement of joint attention, which may be used in future research investigating how gaze is processed by individuals with autism. It would be particularly interesting to investigate the relationship between the P350 and social communication ability, and whether clinical gains in social communication intervention programs are associated with changes in the P350 effect.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Adamson, L. B., Bakeman, R., Deckner, D. F., & Romski, M. A. (2009). Joint engagement and the emergence of language in children with autism and Down syndrome. *Journal of Autism & Developmental Disorders*, *39*(1), 84–96. doi:10.1007/s10803-008-0601-7

Baron-Cohen, S. (1995). *Mindblindness*. Cambridge, MA: MIT Press.

Carrick, O. K., Thompson, J. C., Epling, J. A., & Puce, A. (2007). It's all in the eyes: Neural responses to socially significant gaze shifts. *Neuroreport*, *18*(8), 763–766. doi:10.1097/WNR.0b013e3280ebb44b

Caruana, N., Brock, J., & Woolgar, A. (2015). A frontotemporoparietal network common to initiating and responding to joint attention bids. *NeuroImage*, *108*, 34–46. doi:10.1016/j.neuroimage.2014.12.041

Caruana, N., de Lissa, P., & McArthur, G. (2015). The neural time course of evaluating self-initiated joint attention bids. *Brain and Cognition*, *98*, 43–52. doi:10.1016/j.bandc.2015.06.001

Charman, T. (2003). Why is joint attention a pivotal skill in autism? *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*, 315–324. doi:10.1098/rstb.2002.1199

Cullen, H., Kanai, R., Bahrami, B., & Rees, G. (2013). Individual differences in anthropomorphic attributions and human brain structure. *Social Cognitive and Affective Neuroscience*. doi:10.1093/scan/nst109

Feng, Q., & Zhang, X. (2014). Eye gaze triggers reflexive attention shifts: Evidence from lateralised ERPs. *Brain Research*, *1589*, 37–44. doi:10.1016/j.brainres.2014.09.029

Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: Visual attention, social cognition, and individual differences. *Psychological Bulletin*, *133*(4), 694–724. doi:10.1037/0033-2909.133.4.694

Ganis, G., et al. (2012). The N170, not the P1, indexes the earliest time for categorical perception of faces, regardless of interstimulus variance. *NeuroImage*, *62*(3), 1563–1574. doi:10.1016/j.neuroimage.2012.05.043

Georgescu, A. L., Kuzmanovic, B., Roth, D., Bente, G., & Vogeley, K. (2014). The use of virtual characters to assess and train nonverbal communication in high-functioning autism. *Frontiers in Human Neuroscience*, *8*. doi:10.3389/fnhum.2014.00807

Hietanen, J. K., Nummenmaa, L., Nyman, M. J., Parkkola, R., & Hämäläinen, H. (2006). Automatic attention orienting by social and symbolic cues activates different neural networks: An fMRI study. *NeuroImage*, *33*(1), 406–413. doi:10.1016/j.neuroimage.2006.06.048

Holmes, A., Mogg, K., Garcia, L. M., & Bradley, B. P. (2010). Neural activity associated with attention orienting triggered by gaze cues: A study of lateralized ERPs. *Social Neuroscience*, *5*(3), 285–295. doi:10.1080/17470910903422819

Itier, R. J., & Batty, M. (2009). Neural bases of eye and gaze processing: The core of social cognition. *Neuroscience & Biobehavioral Reviews*, *33*(6), 843–863. doi:10.1016/j.neubiorev.2009.02.004

Itier, R. J., & Taylor, M. J. (2004a). Face recognition memory and configural processing: A developmental ERP study using upright, inverted, and contrast-reversed faces. *Journal of Cognitive Neuroscience*, *16*(3), 487–502. doi:10.1162/089892904322926818

Itier, R. J. C. A., & Taylor, M. J. (2004b). Source analysis of the N170 to faces and objects. *Neuroreport*, *15*(8), 1261–1265. doi:10.1097/01.wnr.0000127827.73576.d8

Kimball, S., & Mattis, P. (1995). *GIMP Image Manipulation Program (Version 2.8.2)*. Berkley.

Lassalle, A., & Itier, R. J. (2013). Fearful, surprised, happy, and angry facial expressions modulate gaze-oriented attention: Behavioral and ERP evidence. *Social Neuroscience*, *8*(6), 583–600. doi:10.1080/17470919.2013.835750

Mundy, P., Sullivan, L., & Mastergeorge, A. M. (2009). A parallel and distributed-processing model of joint attention, social cognition and autism. *Autism Research*, *2*(1), 2–21. doi:10.1002/aur.61

Murray, D. S., Creaghead, N. A., Manning-Courtney, P., Shear, P. K., Bean, J., & Prendeville, J.-A. (2008). The relationship between joint attention and language in children with autism spectrum disorders. *Focus on Autism & Other Developmental Disabilities*, *23*(1), 5–14. doi:10.1177/1088357607311443

Nation, K., & Penny, S. (2008). Sensitivity to eye gaze in autism: Is it normal? Is it automatic? Is it social? *Development and Psychopathology*, *20*(01), 79–97. doi:10.1017/S0954579408000047

Pelphrey, K. A., Shultz, S., Hudac, C. M., & Vander Wyk, B. C. (2011). Research review: Constraining heterogeneity: The social brain and its development in autism spectrum disorder. *Journal of Child Psychology and Psychiatry*, *52*(6), 631–644. doi:10.1111/j.1469-7610.2010.02349.x

Peter, J., Valkenburg, P. M., & Schouten, A. P. (2005). Developing a model of adolescent friendship formation on the internet. *CyberPsychology & Behavior*, *8*(5), 423–430. doi:10.1089/cpb.2005.8.423

Pfeiffer, U. J., Schilbach, L., Timmermans, B., Kuzmanovic, B., Georgescu, A. L., Bente, G., & Vogeley, K. (2014). Why we interact: On the functional role of the striatum in the subjective experience of social interaction. *NeuroImage*, *101*, 124–137. doi:10.1016/j.neuroimage.2014.06.061

Pönkänen, L. M., Alhoniemi, A., Leppänen, J. M., & Hietanen, J. K. (2010). Does it make a difference if I have an eye contact with you or with your picture? An ERP study. *Social Cognitive and Affective Neuroscience*. doi:10.1093/scan/nsq068

Samson, D., Apperley, I. A., Chiavarino, C., & Humphreys, G. W. (2004). Lefttemporoparietal junction is necessary for representing someone else'sbeliefs. *Nature Neuroscience*, *7*, 499–500.

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T., & Vogeley, K. (2013). Toward a second-person neuroscience. *Behavioral and Brain Sciences*, *36*(4), 393–414. doi:10.1017/S0140525X12000660.

Schilbach, L., Wilms, M., Eickhoff, S. B., Romanzetti, S., Tepest, R., Bente, G., … Vogeley, K. (2010). Minds made for sharing: Initiating joint attention recruits reward-related neurocircuitry. *Journal of Cognitive Neuroscience*, *22*(12), 2702–2715. doi:10.1162/jocn.2009.21401.

Singular Inversions. (2008). *FaceGen Modeller (Version 3.3) [Computer Software]*. Toronto, ON: Singular Inversions.

SR Research. (2004). *Experiment Builder (Version 1.10.165)*. Ontario.

Tomasello, M. (1995). Joint Attention as Social Cognition. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development*. Hillsdale: Lawrence Erlbaum Associates.

van Velzen, J., & Eimer, M. (2003). Early posterior ERP components do not reflect the control of attentional shifts toward expected peripheral events. *Psychophysiology*, *40*(5), 827–831. doi:10.1111/1469-8986.00083

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

Watson, D., Clark, A., & Tellegen, L. A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*, 1063–1070. doi:10.1037/0022-3514.54.6.1063

Wykowska, A., Wiese, E., Prosser, A., Müller, H. J., & Hamed, S. B. (2014). Beliefs about the minds of others influence how we process sensory information. *PLoS ONE*, *9*(4), e94339. doi:10.1371/journal.pone.0094339