Social Cognitive and Affective Neuroscience Advance Access published January 22, 2013

doi:10.1093/scan/nss144

Default distrust? An fMRI investigation of the neural development of trust and cooperation

Anne-Kathrin J. Fett,[1,2] Paula M. Gromann,[1,2] Vincent Giampietro,[3] Sukhi S. Shergill,[2] and Lydia Krabbendam[1,2]

[1]Department of Educational Neuroscience, Faculty of Psychology and Education, VU University of Amsterdam, Van der Boechorststraat 1, 1081 BT Amsterdam, The Netherlands, [2]Department of Psychosis Studies, and [3]Brain Imaging Analysis Unit, Centre of Neuroimaging Sciences, Institute of Psychiatry, King's College London, De Crespigny Park, London SE5 8AF, UK

The tendency to trust and to cooperate increases from adolescence to adulthood. This social development has been associated with improved mentalizing and age-related changes in brain function. Thus far, there is limited imaging data investigating these associations. We used two trust games with a trustworthy and an unfair partner to explore the brain mechanisms underlying trust and cooperation in subjects ranging from adolescence to mid-adulthood. Increasing age was associated with higher trust at the onset of social interactions, increased levels of trust during interactions with a trustworthy partner and a stronger decline in trust during interactions with an unfair partner. Our findings demonstrate a behavioural shift towards higher trust and an age-related increase in the sensitivity to others' negative social signals. Increased brain activation in mentalizing regions, i.e. temporo-parietal junction, posterior cingulate and precuneus, supported the behavioural change. Additionally, age was associated with reduced activation in the reward-related orbitofrontal cortex and caudate nucleus during interactions with a trustworthy partner, possibly reflecting stronger expectations of trustworthiness. During unfair interactions, age-related increases in anterior cingulate activation, an area implicated in conflict monitoring, may mirror the necessity to inhibit pro-social tendencies in the face of the partner's actual levels of cooperation.

Keywords: fMRI; perspective taking; theory of mind; trust game; development

INTRODUCTION

Although humans are social by nature, the cognitive abilities that form the basis for successful social interactions are not fully developed at birth, but evolve gradually over time (Hughes, 2004; Ensink and Mayes, 2010). During the transition from adolescence to adulthood, social behaviour becomes increasingly oriented towards others (Steinberg and Sheffield Morris, 2001; Eisenberg *et al.*, 2005; Dumontheil *et al.*, 2010). Improved mentalizing, the sensitivity to the perspective of others, has been suggested to drive increases in trust and cooperation (King-Casas *et al.*, 2005; Sutter and Kocher, 2007; Van den Bos *et al.*, 2010, 2011b). The changes in social cognition and behaviour occur in parallel with structural and functional maturation of the brain. Several studies have investigated the neural correlates of social interactions, but research has only just begun to investigate the brain–behaviour association from a developmental perspective (Van den Bos *et al.*, 2011b).

Over the last decade, combining approaches from neuroscience and economic research has lead to an interest in the neural mechanisms underlying trust and cooperation in adults (King-Casas *et al.*, 2005, 2008; Krueger, 2008). Social cognition in action has been investigated with economic exchange paradigms, such as the trust game (Berg *et al.*, 1995). The trust game requires mentalizing in order to infer intentions from behavioural cues of the other player and to appreciate how own behaviour (e.g. lowering trust in response to trustworthiness) leads to a reputation with others. During the trust game, the first player (investor) receives an amount of money from the experimenter and can choose to cooperate (i.e. share any part of the money) with the second player (trustee) or to defect (i.e. keep the money). The shared amount is multiplied and the trustee decides whether to return any part of this amount or whether to keep the money. The best pay-offs for both players occur when they cooperate. However, the trustee yields the highest pay-off by defecting. Thus, to share money, the investor needs to trust in the good intentions of the trustee. Despite different predictions from classic economic theory, investors typically send a share of 50% or more of their initial endowment; this signal of trust is generally reciprocated by the trustee (Gintis, 2000; Camerer, 2003; Krueger, 2008). In multi-round versions of the trust game, the degree of trustee reciprocity is a strong predictor for subsequent decreases or increases in investor trust (King-Casas *et al.*, 2005), showing that trust or the expectation about future behaviour is modulated by the actual behaviour of the game partner.

Previous functional magnetic resonance imaging (fMRI) research with the trust game demonstrated activation in brain regions important for mentalizing, reward learning, cognitive control and emotional processing (Rilling *et al.*, 2004; Delgado *et al.*, 2005; King-Casas *et al.*, 2005; Krueger *et al.*, 2007; Krueger, 2008; Van den Bos *et al.*, 2011b). It has been postulated that the reward network, extending from the striatum and specific frontal regions, is involved in the motivation to cooperate and that brain networks of cognitive control and social cognition modulate this motivation in response to contextual information (Declerck *et al.*, 2011). Cognitive control is important for the adaptation of behavioural patterns in response to new evidence, e.g. behavioural feedback from interaction partners (Kerns, 2004; Magno *et al.*, 2006). The social cognitive network supports mentalizing, the process of interpreting others' social signals and is important to minimize betrayal (Van Overwalle, 2009; Declerck *et al.*, 2011).

Although cooperation seems to be the preference of adults, developmental studies using trust games suggest a tendency to invest lower amounts and to have less reciprocal interactions in adolescents (Sutter and Kocher, 2007; Van den Bos *et al.*, 2011a, b). This changing quality of social interactions has been attributed to the lower propensity of adolescents to mentalize and a subsequently reduced sensitivity to others' social signals (Dumontheil *et al.*, 2010; Van den Bos *et al.*, 2011b). Studies have begun to elucidate how these differences in social behaviour and cognition are reflected in differential activation of networks subserving social interactions (Van den Bos *et al.*, 2011b).

Received 10 May 2012; Accepted 26 November 2012

Correspondence should be addressed to Anne-Kathrin J. Fett, Department of Educational Neuroscience, Faculty of Psychology and Education, VU University Amsterdam; Prof. E.M. Meijerslaan 2, 1183 BT Amstelveen, The Netherlands. Tel.: +31 20 598 3462; Fax: +31 20 598 8745. E-mail: A.J.Fett@vu.nl

In the study by Van den Bos *et al.* (2011b), a sample of 62 participants between 12 and 22 years took the trustee role in a two-choice trust game with a generally trustworthy investor. Within this sample, age was unrelated to the degree of reciprocity towards the investor. Yet, indicative of increased mentaliszing, the sensitivity to the degree of risk that the investor took during decision making significantly increased with age. The neuroimaging results showed age-related increases in brain activation in the left temporo-parietal junction (TPJ) and the right dorsolateral prefrontal cortex in response to investor trust. The TPJ has been suggested to play a role in identifying goals and intentions behind others' behaviour (Mitchell *et al.*, 2005; Van Overwalle, 2009). Previous research showed age to be associated with higher TPJ activity during simple mentalizing tasks (Blakemore *et al.*, 2007) and self-referential processing (Pfeifer *et al.*, 2007). The involvement of the dorsolateral prefrontal cortex was hypothesized to reflect the age-related regulation of selfish responding. Age was also associated with decreases in activation in the anterior medial prefrontal cortex during defection, when compared with reciprocation of investor trust. This area has also been hypothesized to underlie mentalizing processes (Krueger, 2008), specifically those that include thinking about how oneself is perceived by others (Amodio and Frith, 2006; Frith and Frith, 2008). However, previous developmental research also found this area more active during self-related processing when compared with social processing (Blakemore *et al.*, 2007; Pfeifer *et al.*, 2007).

The previously described study presents initial behavioural evidence in support of age-related increases in the sensitivity to others' perspectives, as indicated by an increased degree of reciprocity when the interaction partner made high-risk investment decisions. However, it remains unclear whether the sensitivity to the other person's behavioural cues also increases with age and whether this underlies the age-related development towards a behavioural default of trust and cooperation. To elucidate this question, this study investigated age-related changes in investor behaviour as a function of partner reciprocity. Participants played two multi-round trust games with anonymous hypothetical game partners, one with a cooperative and one with an unfair decision-making style. If sensitivity to the other person's behavioural cues increases with age, more pronounced increases or decreases in the levels of trust with age should occur in response to cooperation or unfair behaviour by the trustee, respectively.

There is limited research examining social cognitive and behavioural changes during social interactions across broader age ranges. However, previous research indicates that social (cognitive) processes continue to change into adulthood (Dumontheil *et al.*, 2010). Therefore, this study included a sample of participants ranging from adolescence to mid-adulthood. We hypothesized that age would be associated with higher trust and an increased sensitivity to social signals of others and that this would be reflected in (i) higher initial investments and (ii) higher investments throughout interactions with a cooperative, but (iii) lower investments towards an unfair game partner, as any increased sensitivity to the behavioural cues of the other person will likely involve better mentalizing skills and/or better social reward learning. We expected at the neural level (i) age-related increases in brain activation within frontal and temporo-parietal brain areas implicated in mentalizing (e.g. medial prefrontal cortex, TPJ and precuneus) and (ii) age-related decreases in activation in social reward-related areas (e.g. ventromedial prefrontal and orbitofrontal cortex and caudate nucleus) during interactions with a cooperative trustee as a consequence of age-related increases in expectations of the trustworthiness of others. During interactions with an unfair trustee, we similarly expected increases in brain areas implicated in mentalizing, but in addition (iii) age-related increases in activation in cognitive control regions (e.g. anterior cingulate cortex and dorsolateral prefrontal cortex) because of the need to suppress the default intention to invest.

METHOD

Participants

Forty-five healthy right-handed males between the ages of 13 and 49 years (mean age = 23.6 years; SD = 9.76) participated in this study. Participants were recruited at local schools, via colleagues and through a community volunteer database 'Mindsearch' (http://mindsearch.iop.kcl.ac.uk). All participants had a good command of the English language. There was no history of neurological disorder, current psychiatric diagnosis or psychotropic medication. Informed consent was obtained from all participants and their parents/guardians if they were under the age of 16 years. The study was approved by the local research ethics committees [London-Surrey Borders (10/H0806/38) and Barking and Havering REC (08/H0702/83)].

Design

Measures

The vocabulary subtest of the Wechsler Abbreviated Scale of Intelligence (WASI) was used as indicator of general cognitive ability in adolescents [13–18 years (Wechsler, 1999)]. The vocabulary subtest of the Wechsler Adult Intelligence Scale (WAIS) was used in adults [19–49 years (Wechsler, 1981)]. To exclude systematic differences in IQ between the two samples, *T*-scores of the WASI were converted to WAIS-scaled scores for comparability (Deutsch Lezak *et al.*, 2012). The mean scaled score was 11.56 (SD = 2.9). There was no significant age effect on the WAIS/WASI scores ($b = -0.71$, $P = 0.11$ and 95% CI = −0.16/0.02).

Trust game

Participants were told that they would play two trust games with anonymous human counterparts. However, in reality, two probabilistic computer algorithms were used to model the game partners' behaviour, one reflecting a trustworthy, cooperative and one reflecting an unfair decision-making style. Participants were in the investor role throughout the games. In each round, they were asked to transfer an (integer) amount between £0 and £10 to the trustee. The transferred amount was tripled. The subsequent trustee repayment depended on the previous investments of the investor and on the computer algorithm (Supplementary Material S1). At the beginning of a new round, they received £10 again, i.e. there were no cumulative totals. The order of the trustees (cooperative/unfair) was counterbalanced between subjects. Each trust game consisted of 20 trust game rounds and 20 randomly interspaced control rounds. Each round started with an investment cue of £10 (2 s). The following investment period required the subject to move a cursor with their index fingers in order to select a number from 0 to 10 (max. 4 s). The invested amount was shown (2 s), followed by a waiting period with a bar slowly filling itself with dots (2–4 s) and a fixation cross (500 ms). The trustee's response was then displayed (3 s), followed by the totals (3–5 s, depending on the trustee's response). The trial ended with a fixation cross (500 ms). In total, each trial lasted 18.5 s (Supplementary Figure S1). Control rounds consisted of the same timings but during the investment phase, subjects moved the cursor to a randomly placed target, which was displayed below one of the numbers. During the repayment and outcome phase, participants saw two columns as in the real trials.

Procedure

All participants/primary caregivers read the information material and gave written informed consent before the testing procedure. The testing sessions were held individually in a quiet room at the Institute of Psychiatry/Centre of Neuroimaging Sciences. First, participants were assessed with the WASI/WAIS vocabulary subtest. Then, they completed 10 practice trust game rounds on a laptop before the MRI scan. Participants were told that their game partners were in a different location and that they were connected via the internet. After the session, the participants answered a short questionnaire, which was used as a manipulation check and examined their individual perceptions of the games and their game partners. The earnings from one randomly selected round of the trust game were paid to each participant in addition to a fixed payment for the participation.

fMRI image acquisition

Imaging data were acquired using a 3 Tesla GE Signa Neuro-optimised MR System. A quadrature birdcage head coil was used for radio frequency transmission and reception. For each game, 370 T_2^*-weighted whole-brain echo-planar images depicting the blood oxygen-level-dependent (BOLD) contrast were acquired with the following parameters: slice thickness = 2.4 mm; interslice gap = 1 mm; TR = 2000 ms; TE = 25 ms; flip angle = 75°; in-plane voxel dimension = 3.4 mm; number of slices = 38; dummy acquisitions = 4 and matrix = 64 × 64. For anatomical reference, a whole-brain high-resolution gradient-echo image of 43 slices was acquired with the following parameters: slice thickness = 3 mm; interslice gap = 0.3 mm; TR = 3000 ms; TE = 30 ms; flip angle = 90°; in-plane voxel size = 1.9 mm and matrix = 128 × 128. Foam padding was placed around the head in the coil to minimize head movement and the participants were provided with ear protectors.

Data analysis

Behavioural data

The statistical analysis of the behavioural data was conducted in STATA 11.0 (Statacorp, 2009). For the trust game, we used regression analysis to examine (i) the effect of condition (cooperative *vs* unfair) on the mean investments across the whole game and (ii) the associations between age and the first investments in the two games (basic trust). To control for multiple observations within subjects, multi-level random regression was used to investigate the associations between age and the evolution of investments towards the two game partners (relation-specific trust) across four sequential blocks of five game rounds.

fMRI data

The fMRI data were analysed with software developed at King's College London (XBAM, cf. http://brainmap.it). XBAM uses a non-parametric median-based strategy to minimize the assumptions of normal theory-based inference, which are difficult to establish in MRI data (Thirion *et al.*, 2007). Images were pre-processed and corrected for motion, global intensity and spin excitation history (Bullmore *et al.*, 1996). Following realignment, the data were smoothed using a Gaussian filter (FWHM 8.8 mm). Time series analysis for each individual subject was based on wavelet-based data resampling methods (Bullmore *et al.*, 1999). Responses to the experimental paradigms were detected by first convolving each component of the experimental design with each of two gamma variate functions (peak responses at 4 and 8 s, respectively). The best fit between the weighted sum of these convolutions and the time series at each voxel was computed using the constrained BOLD effect model (Friman *et al.*, 2003). Following computation of the model fit, a goodness-of-fit statistic was computed. This consisted of the ratio of the sum of squares of deviations from the mean image intensity (over the whole-time series) due to the model to the sum of squares of deviations due to the residuals [sum of squares ratio (SSQratio)]. The observed and permuted SSQratio maps for each individual, as well as the BOLD effect size maps, were normalized into Talairach space using a two stage warping procedure (Brammer *et al.*, 1997).

Whole-brain linear correlations between age and brain activation during the investments (active condition *vs* implicit baseline) were computed for each condition. To investigate the interaction between condition and age, we analysed the correlation between age and the difference images of the cooperative minus the unfair condition. A positive number indicates higher activation in the cooperative than in the unfair condition, while a negative number indicates higher activation in the unfair than in the cooperative condition (see Figure 3). We also computed correlations between brain activation and investments, to examine whether differences in brain activation could be explained by differences in the invested amounts (Supplementary Table S2). First, the Pearson product-moment correlation coefficient was computed at each voxel in standard space between age and signal change over all subjects. The correlation coefficients were recalculated after randomly permuting the age values (or investments) between subjects. Repeating this step 50 times per voxel, then combining the 'random' coefficients across all voxels, gives the distribution of correlation coefficients under the null hypothesis that there is no association between age (or investments) and brain activity. This probability distribution is used to assess the probability of any particular correlation coefficient under the null hypothesis. The critical value of the correlation coefficient at any desired type I error level in the original (non-permuted) data was determined by reference to this distribution. The analysis was extended to cluster level as described in Supplementary Material S2 (GBAM).

RESULTS

Behavioural data

Participants made significantly higher investments in the cooperative than in the unfair condition ($b = -3.09$, $P < 0.01$ and 95% CI = −3.35/−2.83; Table 1). Age was significantly and positively associated with the initial investments ($b = 0.06$, $P = 0.05$ and 95%CI = 0.001/0.12), i.e. basic trust towards an anonymous interaction partner increased with age. To investigate the development of investments over interactions with the two game partners, we analysed the change in investments across four blocks of five game rounds. During cooperative interactions, there was a significant effect of age ($b = 0.08$, $P < 0.01$ and 95% CI = 0.02/0.13) and block ($b = 0.26$, $P < 0.01$ and 95% CI = 0.09/0.43). The interaction between age and block was not significant ($b = -0.18$, $P = 0.14$ and 95% CI = −0.43/0.06). Thus, while all participants increased their investments in response to cooperation, older people continued to make higher investments throughout the course of repeated cooperative interactions. In the unfair condition, there was no significant main effect of age ($b = 0.04$, $P = 0.18$ and 95% CI = −0.02/0.10). A higher block number was significantly associated with lower investments ($b = -0.53$, $P < 0.01$ and 95% CI = −0.75/−0.32), i.e. all participants decreased their levels of trust in response to unfair behaviour. The interaction was marginally significant ($b = -0.27$, $P = 0.07$ and 95% CI = −0.57/0.03). In the first block, there was a significant positive association between age and investments ($b = 0.06$, $P = 0.04$ and 95% CI = 0.002/0.12; all other

Table 1 Mean investments by condition and block

Condition	Block 1 Investment £ *M* (SD)	Block 2 Investment £ *M* (SD)	Block 3 Investment £ *M* (SD)	Block 4 Investment £ *M* (SD)	Overall Investment £ *M* (SD)
Cooperative	5.99 (2.77)	6.38 (2.99)	6.56 (3.03)	6.70 (2.82)	7.34 (2.73)
Unfair	4.66 (2.92)	3.72 (2.96)	3.12 (2.76)	2.51 (2.34)	4.35 (3.35)

$P > 0.53$) showing that this effect was driven by a stronger decline in initial trust in older individuals (Figure 1). Regression analyses show that the degree of the decline in investments over the first block (round 1–5) is significantly associated with the first investment $b = 0.52$, $P < 0.01$ and 95% CI = 0.15/0.90. The higher the initial trust, the stronger the decline in trust. This pattern may reflect stronger tendencies to punish unfair behaviour by the game partner.

Imaging data

Correlations between age and BOLD signal during investments were analysed by condition (Table 2). In the cooperative condition, age was positively associated with increasing brain activation in foci in the left TPJ, extending into the inferior parietal lobule. There was also activation evident in the bilateral middle frontal gyri and right precentral gyri (Figure 2a). A negative correlation between brain activation and age was present in the orbitofrontal cortex (Figure 2b), the left and right caudate nucleus (Figure 2c) and the bilateral dorsomedial prefrontal cortex. In the unfair condition, increasing age was correlated with increasing activation in the left TPJ including the inferior parietal lobule (Figure 2d) and the mid-cingulate gyrus. Increasing age was also associated with decreasing signal in the left posterior cingulate gyrus, thalamus and the bilateral dorsomedial prefrontal cortex.

An interaction between age and condition was present in the posterior cingulate gyrus and precuneus (Figure 3a) and within foci in the lingual gyrus. With increasing age, these structures were more sensitive to cooperation. An opposite activation pattern was present for the anterior cingulate gyrus (Figures 2e and 3b), i.e. with increasing age, the anterior cingulate became more active in response to unfair behaviour.

DISCUSSION

This study examined age-related changes in the neural correlates of trust and cooperation during social interactions. Participants played

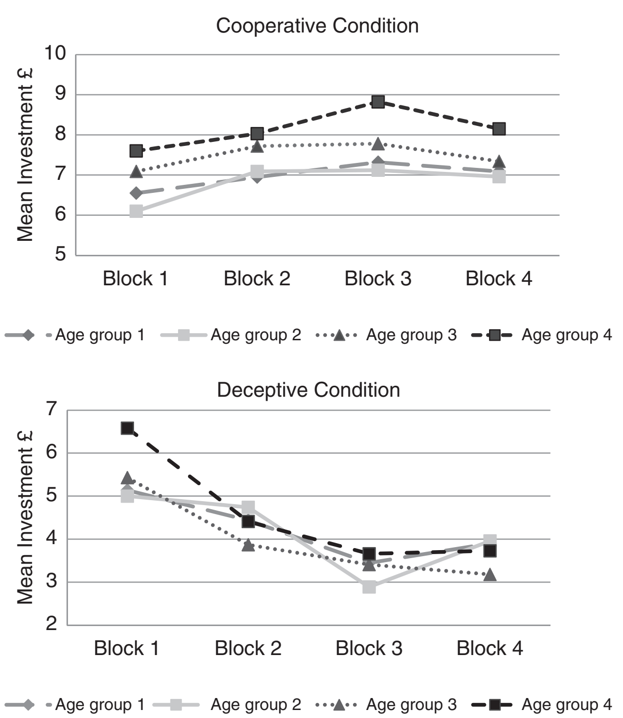

Fig. 1 Mean investments by age group and block number. To display the interaction between condition and age, we grouped age into four groups. Age groups in years: 1 = 13–16, 2 = 17–19, 3 = 20–27 and 4 = 28–49 years. Each block consists of five trust game rounds.

trust games with two hypothetical partners with a cooperative and an unfair decision-making style. Age was associated with higher levels of trust at the onset of social interactions and throughout interactions with a cooperative partner, but also with a steeper decline in the levels of trust throughout interactions with an unfair partner. Associations between age and increased activation in the left TPJ were present during investment decisions towards both game partners. During cooperative interactions, activity in the orbitofrontal cortex and caudate nucleus decreased with age, but increased in the posterior cingulate/precuneus region. The anterior cingulate became increasingly responsive to unfair behaviour as age increased. The findings may reflect a shift from negative to positive expectations of trustworthiness and suggest that differential neural activation patterns in brain areas associated with mentalizing, reward learning and cognitive control may underlie age-related increases in the sensitivity to others' social signals.

In the trust game, mentalizing and social reinforcement learning are important for strategic reasoning about the game partner's intentions and to infer how the game partner perceives one's own behaviour (King-Casas *et al.*, 2005; Sanfey, 2007). As hypothesized, age was not only associated with higher initial trust but also with higher trust during interactions with a cooperative partner and with a steeper decline in trust during interactions with an unfair partner. Still, all participants showed a learning effect regardless of age. They increased their levels of trust in response to cooperation and decreased their levels of trust in response to unfair behaviour. The current imaging findings are in line with hypothesized increases in mentalizing. A better understanding of how the game partner will react and interpret one's own behaviour may support and initially higher investment, a move that signals the intention to cooperate. Age-related increases in activity in the left TPJ were present regardless of the nature of the game partner. The TPJ has been described as part of the 'mentalizing system' (Fletcher, 1995; Ruby, 2004; Gobbini *et al.*, 2007; Van Overwalle, 2009) and trust game research found this area activated when the investor's decision was revealed to the trustee (Van den Bos *et al.*, 2011b). We found increased activation within the TPJ during investment decisions. Possibly, this area plays a role in mentalizing during decisions about how much to trust that are made while predicting the game partner's behaviour. The current data also showed a specific age-related activation during cooperation in the right posterior cingulate/precuneus region which also has been described as part of the mentalizing network (Van Overwalle, 2009; Wolf *et al.*, 2010). Our results show that different areas of the mentalizing network are differentially activated during social interactions. The TPJ is involved regardless of the nature of the social interaction and its activation increases with age. However, the posterior cingulate/precuneus region shows specific activation in response to trustworthiness. In contrast to Van den Bos *et al.* (2011b), current age-related changes in the mentalizing network were predominantly located in the posterior and not the medial prefrontal brain areas. This discrepancy could be explained by the differential nature of the tasks, the specified contrasts or age differences in the samples (i.e. 12–22 *vs* 13–49 years).

In order to make sensible decisions to trust, humans need to learn the associations between their behaviour and the feedback that they receive from others in response. For example, if a certain person answers trust with betrayal, we are less likely to trust that person a second time. However, if the person proves to be trustworthy, we are more likely to trust this person in future (King-Casas *et al.*, 2005). Thus, reward learning shapes behaviour towards optimal decision making. In this study, age was associated with higher levels of trust at the onset of anonymous social interactions. This suggests an increased expectation of benevolence of others and offers a suitable explanation for why feedback learning during cooperation becomes less important as individuals get older. Where cooperation is anticipated, experiencing

Table 2 Correlations between age and brain activation

Talairach coordinates[a]			P-Value	Hemisphere	Clustersize	Cerebral region	BA
x	y	z					
Cooperative condition							
Increasing signal with age							
4	−8	56	0.00002	R	132	Medial frontal gyrus	6
−22	−14	59	0.00004	L	81	Middle frontal gyrus	6
22	−11	56	0.0001	R	83	Middle frontal gyrus	6
50	−15	33	0.0032	R	43	Precentral gyrus	6
−40	−37	40	0.0001	L	121	Temporo-parietal junction	40
−29	63	37	0.001	L	24	Precuneus	19
18	−37	−20	0.0002	R	47	Culmen	
Decreasing signal with age							
0	41	−10	0.0005	L	50	Orbitofrontal cortex	11
−3	44	43	0.0002	L	71	Dorsomedial prefrontal gyrus	8
18	15	40	0.004	R	20	Cingulate	32
14	30	46	0.001	R	24	Dorsomedial prefrontal cortex	8
−11	22	−3	0.0005	L	67	Caudate	
29	−41	16	0.0002	R	73	Caudate	
11	−33	10	0.0002	R	78	Thalamus	
Unfair condition							
Increasing signal with age							
0	4	56	0.0006	R	48	Superior frontal gyrus	6
47	22	36	0.0005	R	11	Precentral gyrus	9
−43	−37	40	0.0002	L	78	Temporo-parietal junction	40
−22	−19	46	0.0007	L	65	Mid cingulate	24
Decreasing signal with age							
−4	−41	30	0.0001	L	46	Posterior cingulate	31
Interaction age × condition							
Higher activation during cooperation with age							
12	−67	−4	0.0002	R	70	Lingual gyrus	18
−28	−70	20	0.0027	L	41	Posterior cingulate/precuneus	18/31
Higher activation during unfair behaviour with age							
7	44	−3	0.0006	R	90	Anterior cingulate	32

Notes: [a]Talairach coordinates where x = left (−) *vs* right (+); y = anterior (+) *vs* posterior (−) and z = ventral (−) *vs* dorsal (+). Maps are indexed by mean SSQ ratio and thresholded to less than one false-positive cluster. BA = Brodmann area.

cooperative behavioural feedback from the trustee matches the prediction. Therefore, learning of new behaviour–outcome associations is unnecessary. fMRI research showed that the orbitofrontal cortex is important for feedback learning and subsequent goal-directed behaviour. It is thought to signal reward value during decision making and is sensitive to value changes, i.e. becomes less responsive when stimuli are less novel or expected (Rolls, 2000; O'Doherty *et al.*, 2001; Rolls and Grabenhorst, 2008). The caudate also plays a role in the anticipation and reception of rewards and has been found to decrease signalling when certain predictions match an outcome (Delgado *et al.*, 2005; Schiffer and Schubotz, 2011). Accordingly, we expected that with increasing age, brain areas that are important for reward learning (e.g. ventromedial prefrontal and orbitofrontal cortex and the striatum) would be recruited to a lesser extent. During cooperative interactions, we found age-related decreases in orbitofrontal cortex and caudate nucleus activation. We did not find the reverse pattern during interactions with an unfair partner. This is in line with previous trust game research of Phan *et al.* (2010) who found that only positive reciprocity by the partner engages the ventral striatum and orbitofrontal cortex. Also, King-Casas *et al.* (2005) found the caudate specifically active during cooperation.

As indicated by increased levels of trust, the anticipation of benevolent behaviour by others increases with age and responding to deviations of trust requires a stronger correction in response to an unfair interaction partner. Hence, the interaction is more demanding for older individuals, as it requires them to adapt their beliefs and behaviour. Younger people need to make fewer adjustments to their investments because of their more distrusting mindset. In line with this reasoning, we found age to be associated with a stronger decrease in trust during the interactions with an unfair partner. During these interactions, expectations of cooperation will have resulted in cognitive conflict in the face of the actual returns. The need to reduce the levels of trust was not reflected in increases in activation in brain areas associated with reward learning, but activation in the dorsal anterior cingulate was modulated by age and the nature of the game partner. This area of the anterior cingulate is known to play a role in conflict monitoring (Botvinick *et al.*, 1999, 2004; Bush *et al.*, 2000; Greene *et al.*, 2004; Ridderinkhof *et al.*, 2004) and there is evidence that it is important for behavioural adjustments (Kerns, 2004; Magno *et al.*, 2006). With age, participants increasingly recruited the anterior cingulate during unfair when compared with cooperative interactions. This may reflect the conflict between expectations of cooperation and experienced social feedback and shows a possible explanation of how feedback shapes future decisions towards more optimal levels of trust (Chang and Sanfey, 2011).

Finally, we found age-related changes in brain areas that were not part of our hypotheses, but that previously have been associated with social cognitive processes. During cooperative interactions, age was also associated with activation in the precentral and bilateral middle frontal gyri. These areas have been described as parts of the 'mirror system' (Van Overwalle and Baetens, 2009), a network which is thought to enable humans to understand the goals of observed (physical) actions of others in an intuitive way by internal simulation. Some evidence indicates that this brain network also engages in more

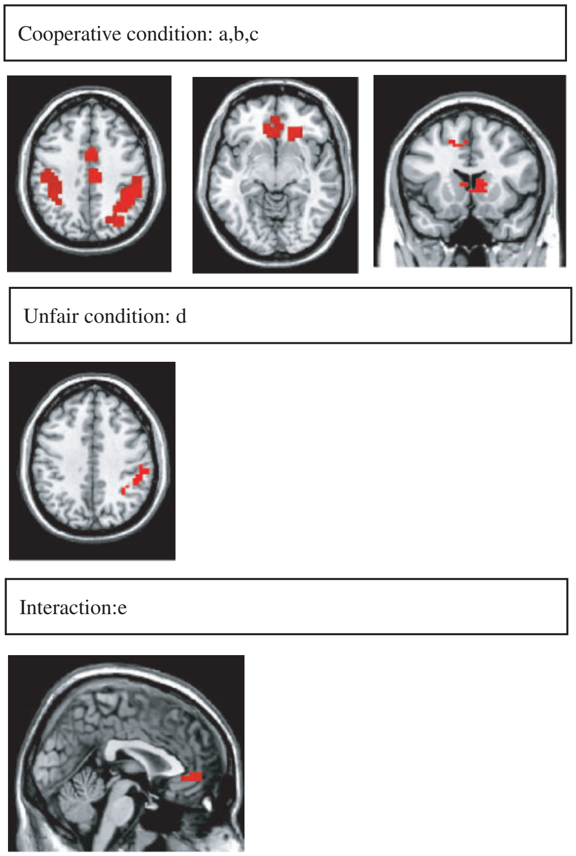

Fig. 2 Brain activation during cooperative interactions. (**a**) Increasing signal with age in the Medial frontal gyrus, precentral gyrus and TPJ. (**b**) Decreasing signal with age in the orbitofrontal cortex. (**c**) Decreasing signal with age in the caudate nucleus. Brain activation during unfair interactions (**d**) TPJ. (**e**) Interaction between age and condition (cooperative minus deceptive) in the anterior cingulate. Thresholds for the images were chosen so that there was less than one false-positive per whole-brain (main effect cooperative condition $\alpha = 0.004$, main effect unfair condition $\alpha = 0.005$, interaction $\alpha = 0.003$).

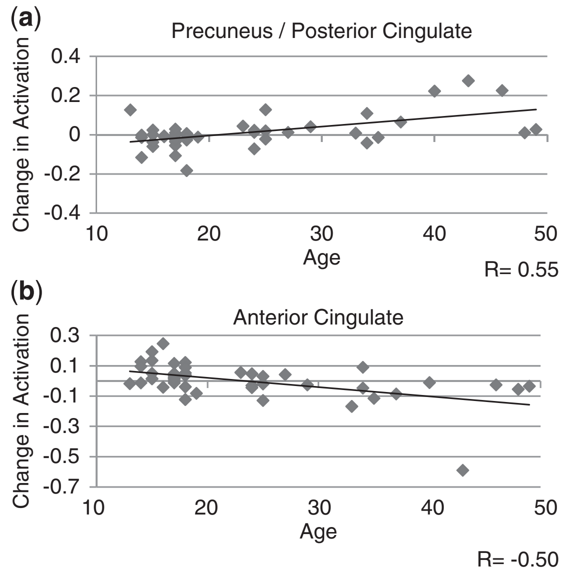

Fig. 3 Interaction in the (**a**) precuneus/posterior cingulate, (**b**) anterior cingulate. *Y*-axis = signal change in brain activation, cooperative minus deceptive investment, *x*-axis = age in years.

abstract forms of social cognition, such as mentalizing (Vogeley, 2001; Wolf *et al*., 2010) and empathy (de Greck *et al*., 2012). Previous trust game research found this area to be activated during decisions to trust (Delgado *et al*., 2005) and age-related increases of activation have been found during social perception (Beadle *et al*., 2012). During unfair interactions, the middle cingulate gyrus was increasingly active with age. This area has been suggested to function as a relay node between negative emotions and motor action and has been reported to be strongly active during decisions to trust in earlier trust game research (King-Casas *et al*., 2005). With age, activation in the dorsomedial prefrontal cortex decreased during cooperative interactions. Earlier research showed this area to be involved in the management of uncertainty in decision making, whereby more uncertainty was associated with increased activation (Volz *et al*., 2005).

Limitations

The current results have to be interpreted in the light of some limitations. First, it might be possible that the differences in trust that we see with increasing age could be due to age differences in risk aversion.While a higher risk aversion may influence the degree of trust, there is extensive literature, which shows that age is typically positively associated with increased risk aversion (Steinberg and Sheffield Morris, 2001; Steinberg, 2004; Deakin *et al*., 2004; Burnett *et al*., 2010; Paulsen *et al*., 2011, 2012). This implies that teenagers should invest more money if risk aversion would be an important motive in the trust game. Furthermore, the concern over whether behaviour in the trust game actually measures trust or risk attitudes has frequently been raised. Eckel and Wilson (2004) provide a comprehensive analysis of the way behaviour in two-person sequential trust games correlates with a variety of behavioural and survey-based risk measures (Eckel and Wilson, 2004). They did not find evidence that any of their risk measures predicts the decision to trust. Also, Houser *et al.* (2006) found that in a risk game with a computer as counterpart, the probability of investing more was significantly higher for risk seeking subjects than for subjects in the risk averse group but risk aversion did not predict investments towards a human counterpart (Houser *et al*., 2006). Second, our manipulation check showed that six individuals of the adolescent ($n = 25$) and one of the adult group ($n = 20$) did have doubts that they were playing with a real human. The fact that fewer adolescents believed that they were playing a human may have reduced their mentalizing operations. Yet, the current behavioural results show increases in trust with age that are in line with those of other research and do not support such concerns (Sutter and Kocher, 2007). Third, it is important to note that age-related effects may also be caused by differences in neurovasculature (Harris *et al*., 2011). Yet, research by Kang *et al.* (2003) showed that in voxelwise group comparison of images of visual and motor cortex regions, only minimal differences were found between children of 7 and 8 years and adults. The small differences in time courses and locations of activation foci between child and adult brains do support the feasibility of direct statistical comparison of these groups within a common space. Fourth, it should be noted that other cognitive factors that may offer an alternative explanation for age-related changes in trust game performance to increases in mentalizing, such as executive functioning or IQ, have not been included as confounders in this study. While evidence from other research of Van den Bos *et al.* (2011b) supports that non-social cognitive factors such as intelligence have a modest influence on social decision making, it would be valuable if future studies would include additional measures of theory of mind and other cognitive functions to validate the interpretation in terms of mentalizing. Fifth, subjects knew that they were taking part in a study with participants between 13 and 18 and 19 and 49, respectively. It is therefore likely that the current effects pertain to trusting behaviour regarding individuals within these age ranges. Trusting behaviour towards other age groups may differ from the current results. It would therefore be interesting for future research to systematically examine age effects with respect to the age of the trustee. Sixth, in the current paradigm, we did not directly control for aspects of monetary reward in the trust game. However, as a means to control that age-related differences in brain activation are not merely due to differences in the investments that were made, we did analyse the association between the invested amounts and brain activation (Supplementary Table S2). Age-related increases and decreases in brain activation in certain areas did not overlap with the foci that were associated with higher and lower investments. The associations between brain activation and age do not seem to be caused by differences in investments. Finally, fMRI allows for the investigation of the role of certain brain regions in certain cognitive functions and caution is required when cognitive processes are inferred from activation in specific brain regions (Poldrack, 2006). Within the brain, it is unlikely that a particular region is activated solely by one cognitive process. The current interpretations are in line with the growing literature about the social brain systems and should be regarded as a guide for future inquiries rather than direct explanations of certain findings.

Conclusion

Our findings render preliminary support to our working hypothesis of improved mentalizing as an underlying mechanism of age-related behavioural preferences of trust and cooperation. Initial trust increased with age. During interactions with a cooperative partner, older people continued to make higher investments. However, when playing with an unfair game partner, older people quickly reduced their levels of trust to those of younger ones. The neuroimaging data showed age to be associated with increased recruitment of brain areas that seem to be important for mentalizing. In line with a stronger preference for cooperation, age was associated with decreased activation in areas involved in reward learning in response to cooperative behaviour by the partner and increased activation in areas associated with cognitive control in response to an unfair game partner.

SUPPLEMENTARY DATA

Supplementary data are available at *SCAN* online.

FUNDING

This work was supported by a VIDI grant from the Netherlands Organization for Scientific Research to Prof. Lydia Krabbendam [MAGW 452-07-007] and a HEFCE Clinical Senior Lecturer Award and MRC to Dr Sukhi S. Shergill.

REFERENCES

Amodio, D.M., Frith, C.D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews Neuroscience*, *7*, 268–77.

Beadle, J.N., Yoon, C., Gutchess, A.H. (2012). Age-related neural differences in affiliation and isolation. *Cognitive, Affective, and Behavioral Neuroscience*, *12*, 269–79.

Berg, J., Dickhaut, J., McCabe, K. (1995). Trust, reciprocity and social history. *Games and Economic behavior*, *10*, 122–42.

Blakemore, S.-J., den Ouden, H., Choudhury, S., Frith, C. (2007). Adolescent development of the neural circuitry for thinking about intentions. *Social Cognitive and Affective Neuroscience*, *2*, 130–9.

Botvinick, M.M., Cohen, J.D., Carter, C.S. (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends in Cognitive Sciences*, *8*, 539–46.

Botvinick, M.M., Nystrom, L.E., Fissell, K., Carter, C.S., Cohen, J.D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature*, *402*, 179–81.

Brammer, M.J., Bullmore, E.T., Simmons, A., Williams, S.C.R., Grasby, P.M., Howard, R.J., Woodruff, P.W.R., Rabe-Hesketh, S. (1997). Generic brain activation mapping in fMRI: A nonparametric approach. *Magnetic Resonance Imaging*. 763–770.

Bullmore, E.T., Brammer, M., Williams, S.C., Rabe-Hesketh, S., Janot, N., David, A., Mellers, J., Howard, R., Sham, P.C. (1996). Statistical methods of estimation and inference for functional MR image analysis. *Magnetic Resonance in Medicine*, *35*, 261–277.

Bullmore, E.T., Suckling, J., Overmeyer, S., Rabe-Hesketh, S., Taylor, E., Brammer, M.J. (1999). Global, voxel, and cluster tests, by theory and permutation, for a difference between two groups of structural MR images of the brain. *IEEE Transactions on Medical Imaging*, *18*, 32–42.

Burnett, S., Bault, N., Coricelli, G., Blakemore, S.-J. (2010). Adolescents' heightened risk-seeking in a probabilistic gambling task. *Cognitive Development*, *25*, 183–96.

Bush, G., Luu, P., Posner, M.I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences*, *4*, 215–22.

Camerer, C.F. (2003). Psychology and economics. *Strategizing in the brain. Science*, *300*, 1673–5.

Chang, L.J., Sanfey, A.G. (2011). Great expectations: neural computations underlying the use of social norms in decision-making. *Social Cognitive and Affective Neuroscience*. (Epub ahead of print; doi:10.1093/scan/nsr094).

de Greck, M., Wang, G., Yang, X., Wang, X., Northoff, G., Han, S. (2012). Neural substrates underlying intentional empathy. *Social Cognitive and Affective Neuroscience*, *7*, 135–44.

Deakin, J., Aitken, M., Robbins, T., Sahakian, B.J. (2004). Risk taking during decision-making in normal volunteers changes with age. *Journal of the International Neuropsychological Society*, *10*, 590–8.

Declerck, C., Boone, C., Emonds, G. (2011). When do people cooperate? The neuroeconomics of prosocial decision making. http://hdl.handle.net/10067/897550151162165141 (Retrieved 11 April 2011).

Delgado, M.R., Frank, R.H., Phelps, E.A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nature Neuroscience*, *8*, 1611–8.

Deutsch Lezak, M., Howieson, D.B., Bigler, E.B., Tranel, D. (2012). *Neuropsychological Assessment*. NY: Oxford University Press.

Dumontheil, I., Apperly, I.A., Blakemore, S.-J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Developmental Science*, *13*, 331–8.

Eckel, C.C., Wilson, R.K. (2004). Is trust a risky decision? *Journal of Economic Behavior & Organization*, *55*, 447–65.

Eisenberg, N., Cumberland, A., Guthrie, I.K., Murphy, B.C., Shepard, S.A. (2005). Age changes in prosocial responding and moral reasoning in adolescence and early adulthood. *Journal of Research on Adolescence*, *15*, 235–60.

Ensink, K., Mayes, L.C. (2010). The development of mentalisation in children from a theory of mind perspective. *Psychoanalytic Inquiry*, *30*, 301–34.

Fletcher, P.C. (1995). Other minds in the brain: a functional imaging study of 'theory of mind' in story comprehension. *Cognition*, *57*, 109–28.

Friman, O., Borga, P., Lundberg, P., Knutsson, H. (2003). Adaptive analysis of fMRI data. *Neuroimage*, *19*, 837–845.

Frith, C.D., Frith, U. (2008). Implicit and explicit processes in social cognition. *Neuron*, *60*, 503–10.

Gintis, H. (2000). Beyond homo economicus: evidence from experimental economics. *Journal of Economic Literature*, *38*, 120–2.

Gobbini, M.I., Koralek, A.C., Bryan, R.E., Montgomery, K.J., Haxby, J.V. (2007). Two takes on the social brain: a comparison of theory of mind tasks. *Journal of Cognitive Neuroscience*, *19*, 1803–14.

Greene, J.D., Nystrom, L.E., Engell, A.D., Darley, J.M., Cohen, J.D. (2004). The neural bases of cognitive conflict and control in moral judgement. *Neuron*, *44*, 389–400.

Harris, J.J., Reynell, C., Attwell, D. (2011). The physiology of developmental changes in BOLD functional imaging signals. *Developmental Cognitive Neuroscience*, *1*, 199–216.

Houser, D., Schunk, D., Winter, J. (2006). *Trust Games Measure Trust*. Discussion Paper No. 1350, Munich Discussion Paper No. 2006–38, Department of Economics, University of Munich, Volkswirtschaftliche Fakultät, Ludwig-Maximilians-Universität München. Online at http://epub.ub.uni-muenchen.de/1350/1/RandT_061217_DP.pdf.

Hughes, C. (2004). What are the links between theory of mind and social relations? Review, reflections and new directions for studies of typical and atypical development. *Social Development*, *13*, 590–619.

Kang, H.C., Burgund, E.D., Lugar, H.M., Petersen, S.E., Schlaggar, B.L. (2003). Comparison of functional activation foci in children and adults using a common stereotactic space. *Neuroimage*, *19*, 16–28.

Kerns, J.G. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science*, *303*, 1023–6.

King-Casas, B., Sharp, C., Lomax-Bream, L., Lohrenz, T., Fonagy, P., Montague, P.R. (2008). The rupture and repair of cooperation in borderline personality disorder. *Science*, *321*, 806–10.

King-Casas, B., Tomlin, D., Anen, C., Camerer, C.F., Quartz, S.R., Montague, P.R. (2005). Getting to know you: reputation and trust in a two-person economic exchange. *Science*, *308*, 78–83.

Krueger, F. (2008). The neural correlates of economic game playing. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*, 3859–74.

Krueger, F., McCabe, K., Moll, J., et al. (2007). Neural correlates of trust. *Proceedings of the National Academy of Sciences*, *104*, 20084–9.

Magno, E., Foxe, J.J., Molholm, S., Robertson, I.H., Garavan, H. (2006). The anterior cingulate and error avoidance. *The Journal of Neuroscience*, *36*, 4769–4773.

Mitchell, J.P., Mason, M.F., Macrae, C.N., Banaji, M.R. (2005). *Thinking about Others: The Neural Substrates of Social Cognition*. Cambridge, MA: MIT Press.

O'Doherty, J., Kringelbach, M.L., Rolls, E.T., Hornak, J., Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience*, *4*, 95–102.

Paulsen, D., Carter, R.M., Platt, M., Huettel, S.A., Brannon, E.M. (2012). Neurocognitive development of risk aversion from early childhood to adulthood. *Frontiers in Human Neuroscience*, *5*. (Epub ahead of print; doi:10.3389/fnhum.2011.00178).

Paulsen, D., Platt, M., Huettel, S.A., Brannon, E.M. (2011). Decision-making under risk in children, adolescents, and adults. *Frontiers in Psychology*, *2*, 72.

Pfeifer, J.H., Lieberman, M.D., Dapretto, M. (2007). "I Know You Are But What Am I?!": neural bases of self- and social knowledge retrieval in children and adults. *Journal of Cognitive Neuroscience*, *19*, 1323–37.

Phan, K.L., Sripada, C.S., Angstadt, M., McCabe, K. (2010). Reputation for reciprocity engages the brain reward center. *Proceedings of the National Academy of Sciences*, *107*, 13099–104.

Poldrack, R.A. (2006). Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*, *10*, 59–63.

Ridderinkhof, K.R., Ullsperger, M., Crone, E.A., Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science*, *306*, 443–7.

Rilling, J., Sanfey, A.G., Aronson, J.A., Nystrom, L.E., Cohen, J.D. (2004). The neural correlates of theory of mind within interpersonal interactions. *Neuroimage*, *22*, 1694–703.

Rolls, E.T. (2000). The orbitofrontal cortex and reward. *Cerebral Cortex*, *10*, 284–94.

Rolls, E.T., Grabenhorst, F. (2008). The orbitofrontal cortex and beyond: from affect to decision-making. *Progress in Neurobiology*, *86*, 216–44.

Ruby, P.J. (2004). How would you feel versus how do you think she would feel? A neuroimaging study of perspective-taking with social emotions. *Journal of Cognitive Neuroscience*, *16*, 988–99.

Sanfey, A. (2007). Social decision-making: insights from game theory and neuroscience. *Science*, *318*, 598–602.

Schiffer, A.-M., Schubotz, R.I. (2011). Caudate nucleus signals for breaches of expectation in a movement observation paradigm. *Frontiers in Human Neuroscience*, *5*. (Epub ahead of print; doi:10.3389/fnhum.2011.00038).

Steinberg, L. (2004). Risk taking in adolescence: what changes, and why? *Annals of the New York Academy of Sciences*, *1021*, 51–8.

Steinberg, L., Sheffield Morris, A. (2001). Adolescent development. *Annual Review of Psychology*, *52*, 83–110.

Sutter, M., Kocher, M.G. (2007). Trust and trustworthiness across different age group. *Games and Economic Behavior*, *59*, 364–82.

Thirion, B., Pinel, P., Mériaux, S., Roche, A., Dehaene, S., Poline, J.-B. (2007). Analysis of a large fMRI cohort: Statistical and methodological issues for group analyses. *Neuroimage*, *35*, 105–120.

Van den Bos, W., Van Dijk, E., Crone, E.A. (2011a). Learning whom to trust in repeated social interactions: a developmental perspective. *Group Processes Intergroup Relations*, *15*, 243–256.

Van den Bos, W., Van Dijk, E., Westenberg, H., Rombout, S.A.R.B., Crone, E.A. (2011b). Changing brains, changing perspectives: the neurocognitive development of reciprocity. *Psychological Science*, *22*, 60–70.

Van den Bos, W., Westenberg, M., Van Dijk, E., Crone, E.A. (2010). Development of trust and reciprocity in adolescence. *Cognitive Development*, *25*, 90–102.

Van Overwalle, F. (2009). Social cognition and the brain: a meta-analysis. *Human Brain Mapping*, *30*, 829–58.

Van Overwalle, F., Baetens, K. (2009). Understanding others' actions and goals by mirror mentalizing systmems. *Neuroimage*, *48*, 564–84.

Vogeley, K. (2001). Mind reading: neural mechanisms of theory of mind and self-perspective. *Neuroimage*, *14*, 170–81.

Volz, K.G., Schubotz, R.I., von Cramon, D.Y. (2005). Variants of uncertainty in decision-making and their neural correlates. *Brain Research Bulletin*, *67*, 403–12.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised*. New York: Psychological Corporation.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. San Antonio, TX: Harcourt Assessment.

Wolf, I., Dziobek, I., Heekeren, H.R. (2010). Neural correlates of social cognition in naturalistic settings: a model free analysis approach. *Neuroimage*, *49*, 894–904.