Implicit and Explicit Processes in Social Cognition

Chris D. Frith[1,3,*] and Uta Frith[2,3]
[1]Wellcome Trust Centre for Neuroimaging
[2]Institute of Cognitive Neuroscience
University College London, London, WC1N 3AR UK
[3]Center of Functionally Integrative Neuroscience, Aarhus University Hospital, DK-8000 Aarhus C, Denmark
*Correspondence: c.frith@ucl.ac.uk
DOI 10.1016/j.neuron.2008.10.032

In this review we consider research on social cognition in which implicit processes can be compared and contrasted with explicit, conscious processes. In each case, their function is distinct, sometimes complementary and sometimes oppositional. We argue that implicit processes in social interaction are automatic and are often opposed to conscious strategies. While we are aware of explicit processes in social interaction, we cannot always use them to override implicit processes. Many studies show that implicit processes facilitate the sharing of knowledge, feelings, and actions, and hence, perhaps surprisingly, serve altruism rather than selfishness. On the other hand, higher-level conscious processes are as likely to be selfish as prosocial.

A Brief History of Social Cognitive Neuroscience

Less than 20 years ago, the idea of identifying human brain systems involved in social interaction would have met with incredulity and derision. Such social matters were the domain of the humanities rather than biology (Frith, 2007). Yet humans are among the most social of all primates and success in social interactions is one of the major forces driving our evolution (Humphrey, 1976). Now there are two specialist journals and innumerable laboratories dedicated to social cognitive neuroscience. Such research programs aim to uncover the physiology underlying the cognitive processes engaged during social interactions. How did the extraordinary flourishing of this topic come about?

Studies of complex social behavior in monkeys (e.g., Cheney and Seyfarth, 1990) have provided one major impetus for the development of social neuroscience. One of the first people to talk explicitly about the social brain was Leslie Brothers (1990), and her evidence for identifying the components came from the study of nonhuman animals. One influential and still unanswered question, originally posed by Premack and Woodruff (1978), was whether the chimpanzee had a ''Theory of Mind.'' This question and possible methods for answering it galvanized research in child development (e.g., Leslie, 1987; Wimmer and Perner, 1983). Baron-Cohen et al. (1985) showed that children with autism were unable to attribute false beliefs and predict someone else's behavior on the basis of their mental states at the appropriate age. From this observation of a circumscribed deficit in social cognition, it was a small step to use the recently available techniques of neuroimaging to investigate what brain processes might underlie the normally pervasive attribution of mental states (e.g., Fletcher et al., 1995).

Animal research in another field has also powered ideas on the brain basis of social cognition. This is the groundbreaking discovery of mirror neurons by Rizzolatti and his colleagues, which was rapidly taken up in human experiments (Rizzolatti and Craighero, 2004). The physiological basis of such fundamental social processes as imitation, emotional resonance, and empathy could now be studied with precision, and it brought rigor to a field that was threatened by a poverty of good experimental data.

Another crucial field that has contributed ideas to the brain basis of social interaction is experimental economics. Here also there has been a critical role for animal studies. Economic models of decision making can be applied directly to the study of animal learning and can guide the analysis of human brain imaging studies (Glimcher and Rustichini, 2004). Economic models allow concepts such as trust and altruism to be quantified and make links between evolutionary psychology and moral behavior (Axelrod and Hamilton, 1981).

Social neuroscience also owes a debt to social psychology. Social psychologists during the 20th century made many remarkable discoveries, showing for instance that an individual's behavior could be shaped and changed, even criminalized, by systematic manipulation of social interactions, as in the famous experiments by Asch, Milgram, and Zimbardo. However, it was not until the technological breakthrough of brain imaging methodologies that social psychology could transform itself into social cognitive neuroscience. At the first conference on this topic, Ochsner and Lieberman (2001) celebrated the emergence of this new field. They wished ''to infuse social psychology with brain science methodology in the hope of deciphering how the brain controls such cognitive processes as memory and attention, which then influence social behaviors such as stereotyping, emotions, attitudes, and self-control.''

An Introduction to Themes in Social Cognitive Neuroscience

Today, research in social neuroscience includes many themes. These can be classified into three principal domains. First, participants are asked to read the dispositions and emotions of other people, often via facial expressions. Second, participants are asked to read the intentions, desires, knowledge, and beliefs of other people (called theory of mind or mentalizing). Third, pairs of participants engage in real-time interactions, typically by playing economic games. The first two classes are largely concerned with the effect of social situations on individuals. Only studies in the third domain directly investigate the mutual give and take between two or more individuals involved in social interactions.

However, the processes revealed in the first two domains of study play an important role in these social interactions. In this essay we will concentrate on the distinction between implicit and explicit processes.

Implicit and Explicit Processes in Social Cognition

One idea that has emerged in all these domains is that there are two levels of social cognition. At the lower level there are fast, relatively inflexible routines that are largely automatic and implicit and may occur without awareness. At the higher level there are slow, flexible routines that are explicit and require the expenditure of mental effort. We shall illustrate this point with two examples.

Prejudice

People respond automatically to the presentation of a fearful face with a fear response. They show increased autonomic activity and increased neural activity in the amygdala (Ohman and Mineka, 2001). This happens even when the fearful face is masked in such a way that the subject is unaware that it was presented (Morris et al., 1999; Whalen et al., 1998). Implicit race prejudice is revealed when people respond to the presentation of a black face with a similar fear response. The amplitude of this response in the amygdala correlates with the degree of prejudice as measured by the Implicit Association Test, but not with explicit (conscious) measures of race prejudice (Phelps et al., 2000). In a subsequent experiment (Cunningham et al., 2004), the faces of black Americans were presented, either very briefly (30 ms) or for a longer period (535 ms). The amygdala activation associated with implicit (unconscious) race prejudice was much reduced when the faces were presented for the longer period. Furthermore, the magnitude of activity in prefrontal cortex predicted how much the amygdala activity would be reduced for the long presentations. This result suggests that activation in dorsolateral prefrontal cortex (DLPFC) and anterior cingulate cortex (ACC) is associated with deliberate attempts to control undesirable prejudicial responses to black faces. There are two forms of race prejudice: implicit and explicit. Within individuals these implicit and explicit forms are relatively independent of each of other.

Mentalizing

Mentalizing refers to the ability to read the desires, intentions, and beliefs of other people and is also referred to as having a theory of mind. The acid test of this ability is the false belief task (Wimmer and Perner, 1983). For example, we know that Maxi will look for his chocolate in the cupboard because he doesn't know that his mother has moved it to the fridge. His behavior is determined by his false belief about the location of the chocolate. Children can pass this test at around the age of 5 years. At this age they have explicit knowledge about Maxi's false belief. They can tell you why Maxi looks in the wrong place. However, an implicit understanding of false beliefs can be found in younger children. At around the age of 3 years many children will look at the correct location (where Maxi believes the chocolate to be), while telling the experimenter that Maxi will look in the other location (Clements and Perner, 1994). Furthermore, infants of between 12 and 15 months old are surprised (as defined by a longer time spent looking) when an actor's behavior is not determined by that actor's false belief (Onishi and Baillargeon, 2005; Surian et al., 2007). Here again there are two forms of mentalizing, implicit and explicit, that may be relatively independent. For example, it has been proposed (Frith, 2004) that people at the high-functioning end of the autistic spectrum lack implicit mentalizing, but have acquired an explicit form through teaching and experience.

The Role of Instructions

As yet we know very little about the relationship between these implicit and explicit processes in social cognition. Do explicit processes build upon the preexisting implicit processes? Can explicit knowledge directly affect implicit processes? Our guess is that they are largely independent. Some evidence for this independence comes from studies of the very basic physiological process of fear conditioning. If beings experience a visual stimulus, such as blue square, repeatedly followed by a shock, people (and other animals) will rapidly show a fear response to the blue square. Such learning can also occur in a social context. Observing someone else receiving the shock after the blue square will condition the observer to give a fear response to the blue square. In both these cases the learning could be driven by implicit processes: the implicit process by which shock elicits fear and the implicit process by which the sight of a fearful face elicits fear. Fear conditioning can also be achieved by explicit instructions, i.e., telling the subject that the blue square will be followed by a shock. However, this explicit route does not affect the implicit system. If the conditioned stimulus (the blue square) is presented subliminally a fear response still occurs after direct experience of a shock and after observation. It does not occur after explicit instructions (Olsson and Phelps, 2004).

The Interfering Effects of Implicit Social Processes

In addition to our lack of awareness of it, something that is always difficult to prove, another sign of an automatic, implicit process is that we have no top-down control over it. The process runs whether we want it to or not. Many processes involved in social cognition are automatic in this sense. Here are three examples from different domains.

Gaze Following

People (and many other animals) follow the gaze of others (Goossens et al., 2008). They will look at the same place where someone else is looking. As a result gaze direction can be used as a cue in covert attention tasks. If a face looks to the left just before a target appears on the left, the response to that target will be quicker. If the face had looked to the right, the response to the target would have been slower (Driver et al., 1999). This is the cue validity effect. This effect occurs even when the gaze direction cue is presented subliminally (Sato et al., 2007). Bayliss and Tipper have shown that this is an automatic response that occurs even when this interferes with performance. In these experiments the gaze cue was given by individual faces. Some faces looked in the same direction as the target, while others consistently looked in the opposite direction. However, this difference did not affect gaze following. Subjects continued to follow the gaze of the faces that looked in the wrong direction even though this slowed down their response to the target (Bayliss and Tipper, 2006).

Imitating Actions

When we observe other people in motion, activity can be seen in brain regions concerned with the production of action.

Presumably this observation relates to the mirror neurons identified in the premotor and parietal cortices of the monkey (Rizzolatti and Craighero, 2004). This overt imitation of others can spill over into overt imitations of which we seem to be unaware. For example, when people converse they tend to imitate each others' movements and gestures (the chameleon effect: Chartrand and Bargh, 1999). This automatic imitation can interfere with the actions of the observer. If we watch someone moving their arm from side to side while we are moving our arm up and down, our movements become more variable (Kilner et al., 2003). However, this interference only happens if we believe we are watching the movements of another human rather than a robot (Stanley et al., 2007).

Tracking the Knowledge of Others

In order to solve the theory of mind tasks we discussed earlier, we have to appreciate that people do not know about things that they cannot see (e.g., the food hidden behind the rock) and also that what they can see will depend on their point of view. So, for example, I know that my living room has four pictures in it: two on the east wall and two on the west wall. However, someone entering the room for the first time will only see the two on the west wall and will not know that there are two on the wall behind her. So I know that there are four pictures in the room, and I also know that my friend can only see two of them. Translating this scenario into an experimental setting, Samson et al. (2007) have shown that the mere presence of an ignorant person in the room interferes with subjects' ability to say how many pictures are in the room. They are slower and may make errors. It seems that we cannot help taking into account other peoples' knowledge.

Tracking the Goals and Intentions of Others

We have already discussed the covert attention paradigm in which invalid cues can slow down responses to targets. A related choice reaction time paradigm makes use of congruent and incongruent imperative signals. In an experiment by Sebanz and colleagues, the imperative stimulus was a pointing finger with a ring on it. If the ring was red the left button was to be pressed, while if the ring was green the right button was to be pressed. The direction in which the finger pointed (left or right) was to be ignored. In this case, a red ring with the finger pointing right is an example of an incongruous signal and elicits a slower response. This interfering effect of incongruity can be eliminated if the paradigm is turned into a go no-go task in which the subject only has to respond to the red ring. With this paradigm there is no interference when the finger points in the "wrong" direction. The critical observation is what happens when a second subject is brought in to perform the other half of the task. One subject presses the left button for the red ring. The other presses the right button for the green ring. With this paradigm the interference comes back, but only if the second subject is actually performing the task. Their mere presence is not enough (Sebanz et al., 2003). This result shows that we automatically represent the task our companion is doing even though this interferes with our own performance.

What Is the Function of These Low-Level Automatic Processes?

We have emphasized the interfering effects of these processes as evidence that they are automatic. The corollary of this demonstration is that, in most circumstances, they are beneficial. The benefit that they provide is to increase the efficiency and success of group behavior at the expense of the individual.

The Automatic Following of Gaze Shifts: Joint Attention

By following the gaze of others we discover what it is that they are attending to. In this way we achieve joint attention; that is, we both attend to the same thing and the foreground and background aspects of our attention become aligned (Tomasello and Carpenter, 2007). The effects of observing eye gaze are not simply about altering the direction of spatial attention. Joint attention allows us to learn about the world from others. Infants learn to approach or avoid novel objects by observing the response of the mother during joint attention (social referencing, Feinman et al., 1992). This effect also happens in adults. Objects looked at by other people are more liked than objects that do not receive attention (Bayliss et al., 2006). Observation of eye gaze shifts elicits activity in the posterior superior temporal sulcus (pSTS) (Pelphrey et al., 2003). This activity seems to reflect inferences about the intention of the actor since it is greater when the gaze shift is in an unexpected direction (see also Saxe et al., 2004). Furthermore, this activity is specific to social cues that direct attention including both pointing and eye gaze (Materna et al., 2008b). In contrast, nonsocial cues that orient attention elicit activity in intraparietal sulcus (IPS) (Materna et al., 2008a).

Sharing Action

When performing a task with someone else, we need to coordinate our movements and ensure that we have the same goal. This alignment in action is helped by covert and automatic representations of the actions and intentions of others. Indeed, the covert imitation manifest in the chameleon effect biases people toward prosocial behavior. If our movements are covertly imitated we will feel more friendly toward the person who is imitating us and will even feel more well-disposed toward people in general, as shown by an increased willingness to give money to charity (van Baaren et al., 2004). This effect seems only to work when we are unaware of the imitation (Lakin and Chartrand, 2003). If we become aware of the imitation it seems creepy and manipulative rather than pleasing.

However, simply imitating the actions of another is often not an aid to joint action. For example, when carrying a heavy object together, one person may walk forward while the other person walks backward (Sebanz et al., 2006). The alignment of goals is usually more important than the alignment of actions and actions may have to be complementary rather than imitative. In the study by Reed et al. (2006), two subjects had to move a lever to control the position of a pointer on a screen. The two levers were joined by a rigid rod so that the movements of the two subjects had to be precisely coordinated. Their only means of communication was the force they could sense in the rod. The striking result of this study was that the time taken to move the pointer to a target was shorter when the two subjects acted together rather than when either acted alone. This was achieved by the rapid acquisition of complementary strategies whereby one subject controlled the acceleration and the other the deceleration phase of the rapid ballistic movement.

Shared action can be observed in infants as young as 18 months (Warneken and Tomasello, 2006). Such infants will spontaneously help strange adults to achieve their goals. This helping involves complementary action rather than simple imitation, indicating an understanding of these goals and the motivation to help.

The observation of movements in humans typically elicits activity in premotor and inferior parietal cortices, which probably contain neurons equivalent to the mirror neurons found in monkeys (Dinstein et al., 2007; Rizzolatti and Craighero, 2004). However, mirror neurons are characterized as being linked to the execution of a specific action and the observation of that same action. This leads to the question of what happens in the mirror system when complementary rather than identical actions are required in a joint action task. Newman-Norlund et al. (2007) devised a task in which subjects interacted while making either the same action or complementary actions. Activity in inferior frontal and inferior parietal cortices was greater during preparation for complementary action. This suggests that these regions support a more sophisticated system in which the link between action execution and action observation is modulated by the needs of the joint action task that is being performed.

Sharing Knowledge

The automatic alignments between people that we have been discussing provide the common ground (Clark, 1996) that is important for successful interactions, including verbal interactions. During a dialog speakers become aligned at many levels from the phonological to syntax and semantics (Pickering and Garrod, 2004). This alignment improves comprehension. We suggest that there is an additional alignment relating to knowledge. To achieve successful group action, we need to know what other people know and, in particular, what other people don't know. We also need to know about common knowledge, which is the knowledge that everybody knows that everybody knows (Geanakoplos, 1994). One of the Gricean maxims for good communication is that you should not tell people things that they already know (Grice, 1989). We suggest that, in addition, there is an automatic drive to tell people things that they don't know. At 3 years infants will adjust their answers to take into account the ignorance of their questioner (Perner and Leekam, 1986). As early as 12 months, infants will point to inform others about events these others do not know about (Liszkowski et al., 2007).

To achieve common knowledge we should go further and tell people what they don't know in public so that everybody else now knows that they know. This is one of the many important features of gossip.

Underpinning this drive to tell people things they don't know is the automatic tendency to represent what other people know when it is different from what we know. As we have seen this can interfere when we are asked to say what we know in the presence of someone with different knowledge (Samson et al., 2007). Perhaps it is this tendency and the related drive to share knowledge that explains the difficulty that even adults have with hiding their privileged knowledge (Keysar et al., 2003).

The Group versus the Individual

According to Sober and Wilson (1998) when individuals compete within a group it is the uncooperative (selfish) individuals who do better. However, when groups compete it is the groups with more cooperative individuals that do better. So how do groups maintain a high proportion of unselfish individuals? The automatic processes we have considered may provide the impetus for people to behave in a cooperative and prosocial manner. However, as we have seen, these processes are largely automatic and probably occur without awareness. This observation goes against the idea that man's primitive instincts are selfish and that higher rational thought is needed to overcome this selfishness.

Bentham (1789), for example, was following a long tradition when he said that "nature has placed mankind under the guidance of two sovereign masters, pain and pleasure." To be driven solely by pleasure and pain would seem to be fundamentally selfish. However, as Adam Smith (1759) pointed out, man's greatest pleasure comes from the approbation of others and his greatest pain from the disapproval of others. The formation and preservation of our reputation is a fundamental drive. As a result the mere presence of others makes us become more prosocial.

Audience Effects

Models of cooperation have recently incorporated the possibility of "image scoring" and reputation formation as promoters of cooperation (Nowak and Sigmund, 1998a, 1998b; Wedekind and Milinski, 2000). An example of this effect can be seen when people play the dictator game. In this game a proposer is given a sum of money, a proportion of which he can offer to a second player. A purely selfish person would offer very small amounts of money since the second player has no come-back. However, the offers made in the dictator game are still substantial, at around 20% of the total (Forsythe et al., 1994). Why should the proposer give anything away in this game? That generosity in the dictator game depends, in part, on the wish for a good reputation in the eyes of others is supported by the observation that smaller offers are made when the proposer has complete anonymity (Hoffman et al., 1996). Larger offers are made when proposers believe they have an audience. Such audience effects can be observed in children at least as young as 4, for whom the feeling of pride depends strongly upon an audience (Seidner et al., 1988).

The use of generous "gifts" to enhance reputation has a long history and can be found in all cultures (Mauss, 1924). Such use extends from traditional potlatch rituals where leaders strengthen group relations and acquire honor by giving away large amounts of goods to modern phenomena such as open source software and Wikipedia, to which many people freely donate their time and expertise (Zeitlyn, 2003).

In order to keep an explicit track of our reputation, we need to be able to represent what other people think about us. This is sophisticated form of mentalizing. We are doing more than representing other peoples' mental states, we are representing their representation of our mental states. We have speculated (Amodio and Frith, 2006; see also Saxe, 2006) that anterior rostral medial prefrontal cortex (arMPFC) might have a special role in such representations. The need to mentalize, whether to predict what our gaming partner is going to do next or to manipulate our reputation in his eyes, does not arise if we are playing against a computer (Rilling et al., 2004) or a person who is simply following a predetermined sequence of instructions (Singer et al., 2004). When subjects believe they are playing against such partners, significantly less activity is observed in brain regions concerned with mentalizing, including MPFC.

The presence of an audience also affects imitation. Bavelas et al. (1986) studied imitation when subjects observed the victim of an apparently painful injury. When we see someone in pain we tend to mimic their pain by displaying a facial expression of pain. This mimicry was significantly enhanced when the observer and victim were in eye contact. It seems that, when we believe we are engaged in a social interaction, we experience, or at least express, more empathy. Is part of this effect also related to reputation management? Are we trying to make sure that others know that we have empathy? The implication is that being empathic is a socially desirable trait.

At least some of these audience effects do not depend upon explicit processes. Bateson et al. (2006) studied the behavior of people using an honesty box to pay for hot drinks. Over a series of weeks either a photograph of a pair of eyes or a photograph of flowers were put on display in the coffee room. There was a significant increase in the amount of money put in the honesty box on weeks in which eyes rather than flowers were displayed. Since the eyes in this experiment were clearly not real and no actual audience was involved, the effect must reflect an implicit process that was resistant to the subjects' explicit knowledge of the situation.

Selfish Emotion and Social Reason?

A traditional Cartesian view has been that automatic processes, such as emotion, are primitive and selfish, meaning that prosocial, altruistic behavior requires these processes to be overridden by reason in the form of explicit high-level control processes. There is increasing evidence that this view is wrong (Damasio, 1994). A good example comes from studies of the ultimatum game. In the ultimatum game, as in the dictator game that we have already discussed, one player (the proposer) is given a sum of money and then must choose how much to offer the other player (the responder). In contrast to the dictator game, the responder may accept the offer or refuse it. A refusal means that both parties get nothing. A rational responder should accept any offer however small, since some money is better than none. However, this is not what actually happens. In the ultimatum game the majority of responders will refuse offers of less than a third of the total (see Camerer and Thaler, 1995 for a review). Responders consider that small offers are "unfair," and this emotional response is reflected in the brain activity elicited with low offers that is observed in the anterior insula (Sanfey et al., 2003). Furthermore, the greater the activity in the anterior insula, the more likely the offer is to be rejected. The justification for interpreting this in terms of emotion is that activity in the anterior insula has often been associated with subjective feeling states, such as the subjective unpleasantness of painful stimulation (Craig, 2002) or the feeling of disgust (Wicker et al., 2003): it reflects how pleasant or unpleasant we feel a situation to be.

Altruistic Punishment

The implicit assumption is that this emotional response makes the decision less rational. However, it is also less selfish. When we consider the group rather than the individual, turning down unfair offers can be seen as a good decision. When the responder turns down an unfair offer in the ultimatum game, he is effectively punishing the proposer who will not get any money as a result of the refusal. This is an example of altruistic punishment since the responder foregoes monetary gain in order to punish the proposer. We suggest that altruistic punishment is another example of prosocial behavior.

Altruistic punishment has been shown to have a vital role in maintaining cooperation in groups (Bowles and Gintis, 2002, 2004; Boyd et al., 2003; Fehr and Gachter, 2002; Gintis, 2000). In common good games involving several players, the group benefits from the investments of individual players. Each time an individual invests, the group as a whole gains, while the individual investor loses a little. As long as everyone invests, then everyone also gains. But in such situations free riders will appear. These players accept the benefit from the investments of others while withholding their own money. The free riding individuals gain at the expense of the group. Once free riders have appeared in the group cooperation breaks down. If, however, altruistic punishment is possible, then free riding is reduced and cooperation flourishes (Fehr and Gachter, 2002). Here altruistic punishment is applied even though it brings a material loss to the individual player. But while the individual may lose each time he applies punishment, he benefits in the long run from the increased cooperation occurring in the group. A loss to the individual is converted into a gain for the group.

As with the other examples of prosocial behavior that we have already discussed, our propensity to apply altruistic punishment in the ultimatum game depends upon whether or not we are interacting with another person. When the offer in the ultimatum game is based on the spin of a roulette wheel, much lower offers are accepted (Blount, 1995). The same result was observed by Rilling et al. (2004) when subjects believed the offer was being made by a computer.

Are Higher-Level Control Processes Prosocial?

This result emphasizes the fact that not all automatic, "primitive" processes are selfish. Some are prosocial. So are explicit, high-level control processes needed to bias our behavior toward these automatic prosocial processes and away from the selfish ones? This is clearly not the case. As we have seen the presence of others is sufficient to bias us automatically toward prosocial processes. But, of course, this doesn't mean that explicit high-level control processes have no role.

In fact the evidence is that high-level control processes are just as likely to be selfish as prosocial and can be used to override automatic prosocial behavior. This is illustrated in an experiment by Valdesolo and DeSteno (2008). This experiment examined the processes underlying the moral hypocrisy through which subjects judge their own unfair behavior to be less objectionable than the same behavior in others. The key finding was that this biased judgment was eliminated when subjects had to perform a competing cognitive task. These findings suggest that the hypocrisy stems from explicit volitionally guided judgments used to override automatic, more prosocial attitudes to fairness. Indeed, it is through becoming explicitly aware of these low-level prosocial processes that we can subvert them. Thus, rather than share knowledge we may keep it to ourselves to manipulate and deceive others. Furthermore, we use our high-level cognition to provide explicit and "rational" justifications as to why in our case "greed is good."

Are Automatic Prosocial Processes Necessarily Good?

Sober and Wilson (1998) suggest that altruistic behavior emerges from group competition, since groups with a larger proportion of prosocial individuals will do better. One implication of this idea is that the same process that causes us to help members of our own group may cause us to obstruct members of other groups. Does this mean that prosocial behavior such as the sharing of feelings, actions, and knowledge with others does not occur with members of other groups? A suggestion that this might be the case comes from the study of antisocial punishment. We already mentioned the finding that the altruistic punishment of free riders (i.e., low contributors to the common good) can increase their contributions (Fehr and Gachter, 2002). However, it is observed that, rather than increasing their contributions, low contributors will sometimes start punishing the high contributors (Bochet et al., 2006). This is referred to as perverse or antisocial punishment. A cross-cultural study of altruistic and antisocial punishment (Herrmann et al., 2008) found that altruistic punishment dominated in democratic countries such as UK, Denmark, and USA. In contrast, antisocial punishment dominated in less democratic countries such as Saudi Arabia, Russia, and Belarus. Gintis (2008) suggests that this effect may occur because, in these less democratic and more traditional societies, many players may feel themselves part of the small group of family and friends rather than the larger group of unrelated strangers from the same country. This is an example of prosocial processes acting against people from a different group.

At the beginning of this essay we discussed the implicit race prejudice that often occurs in response to the presentation of black faces (Phelps et al., 2000). Prejudice is a fundamental process by which the brain enables us to decide what to do on the basis of inadequate evidence. If there is not enough evidence available in the current stimuli, then we must rely on prior knowledge and expectation. This will always be the case when we meet someone new, who we have never seen before. Prejudice is not intrinsically bad, but, inevitably, it will not do justice to this person. However, when prejudice leads us to believe that this is not a member of our group, then automatic prosocial processes may come into play that act against people who are not in our group. Here our high-level cognitive processes are needed to suppress undesirable prosocial processes.

Automatic implicit processes, and higher-level strategically guided social actions, show an intricate interplay that future studies of social cognition may unravel. If they can do this, we may find ways to prevent some of the deep conflicts between selfishness and altruism that mark human history.

REFERENCES

Amodio, D.M., and Frith, C.D. (2006). Meeting of minds: the medial fronal cortex and social cognition. Nat. Rev. Neurosci. *7*, 268–277.

Axelrod, R., and Hamilton, W.D. (1981). The evolution of cooperation. Science *211*, 1390–1396.

Baron-Cohen, S., Leslie, A.M., and Frith, U. (1985). Does the autistic child have a "theory of mind"? Cognition *21*, 37–46.

Bateson, M., Nettle, D., and Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. Biol. Lett. *2*, 412–414.

Bavelas, J.B., Black, A., Lemery, C.R., and Mullett, J. (1986). I Show How You Feel - Motor Mimicry as a Communicative Act. J. Pers. Soc. Psychol. *50*, 322–329.

Bayliss, A.P., and Tipper, S.P. (2006). Predictive gaze cues and personality judgments: Should eye trust you? Psychol. Sci. *17*, 514–520.

Bayliss, A.P., Paul, M.A., Cannon, P.R., and Tipper, S.P. (2006). Gaze cuing and affective judgments of objects: I like what you look at. Psychon. Bull. Rev. *13*, 1061–1066.

Bentham, J. (1789). An Introduction to the Principle of Morals and Legislations (Oxford, UK: Blackwell, 1948).

Blount, S. (1995). When social outcomes aren't fair: the effect of causal attributions on preferences. Organ. Behav. Hum. Decis. Process. *63*, 131–144.

Bochet, O., Page, T., and Putterman, L. (2006). Communication and punishment in voluntary contribution experiments. J. Econ. Behav. Organ. *60*, 11–26.

Bowles, S., and Gintis, H. (2002). Homo reciprocans. Nature *415*, 125–128.

Bowles, S., and Gintis, H. (2004). The evolution of strong reciprocity: cooperation in heterogeneous populations. Theor. Popul. Biol. *65*, 17–28.

Boyd, R., Gintis, H., Bowles, S., and Richerson, P.J. (2003). The evolution of altruistic punishment. Proc. Natl. Acad. Sci. USA *100*, 3531–3535.

Brothers, L. (1990). The social brain: a project for integrating primate behaviour and neurophysiology in a new domain. Concepts Neurosci. *1*, 27–51.

Camerer, C., and Thaler, R.H. (1995). Anomalies: Ultimatums, Dictators and Manners. J. Econ. Perspect. *9*, 209–219.

Chartrand, T.L., and Bargh, J.A. (1999). The chameleon effect: the perception-behavior link and social interaction. J. Pers. Soc. Psychol. *76*, 893–910.

Cheney, D.L., and Seyfarth, R.M. (1990). How monkeys see the world (Chicago: University of Chicago Press).

Clark, H.H. (1996). Using Language (Cambridge: Cambridge University Press).

Clements, W.A., and Perner, J. (1994). Implicit understanding of false beliefs. Cogn. Dev. *9*, 377–395.

Craig, A.D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. Nat. Rev. Neurosci. *3*, 655–666.

Cunningham, W.A., Johnson, M.K., Raye, C.L., Chris Gatenby, J., Gore, J.C., and Banaji, M.R. (2004). Separable neural components in the processing of black and white faces. Psychol. Sci. *15*, 806–813.

Damasio, A.R. (1994). Descartes' error: emotion, reason, and the human brain (New York: G.P.Putnam's Sons).

Dinstein, I., Hasson, U., Rubin, N., and Heeger, D.J. (2007). Brain areas selective for both observed and executed movements. J. Neurophysiol. *98*, 1415–1427.

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E., and Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orienting. Vis. Cogn. *6*, 509–540.

Fehr, E., and Gachter, S. (2002). Altruistic punishment in humans. Nature *415*, 137–140.

Feinman, S., Roberts, D., Hsieh, K.F., Sawyer, D., and Swanson, K. (1992). A critical review of social referencing in infancy. In Social Referencing and the Social Construction of Reality in Infancy, S. Feinman, ed. (New York: Plenum Press).

Fletcher, P.C., Happe, F., Frith, U., Baker, S.C., Dolan, R.J., Frackowiak, R.S., and Frith, C.D. (1995). Other minds in the brain: a functional imaging study of "theory of mind" in story comprehension. Cognition *57*, 109–128.

Forsythe, R., Horowitz, J.L., Savin, N.E., and Sefton, M. (1994). Fairness in Simple Bargaining Experiments. Games Econ. Behav. *6*, 347–369.

Frith, C.D. (2007). Making up the Mind: How the Brain Creates our Mental World (Oxford: Blackwell).

Frith, U. (2004). Emanuel Miller lecture: confusions and controversies about Asperger syndrome. J. Child Psychol. Psychiatry *45*, 672–686.

Geanakoplos, J. (1994). Common Knowledge. In Handbook of Game Theory, R. Aumann and S. Hart, eds. (Amsterdam: Elsevier Science B.V), pp. 1438–1496.

Gintis, H. (2000). Strong reciprocity and human sociality. J. Theor. Biol. *206*, 169–179.

Gintis, H. (2008). Behavior. Punishment and cooperation. Science *319*, 1345–1346.

Glimcher, P.W., and Rustichini, A. (2004). Neuroeconomics: the consilience of brain and decision. Science *306*, 447–452.

Goossens, B.M.A., Dekleva, M., Reader, S.M., Sterck, E.H.M., and Bolhuis, J.J. (2008). Gaze following in monkeys is modulated by observed facial expressions. Anim. Behav. *75*, 1673–1681.

Grice, H.P. (1989). Studies in the Way of Words (Cambridge, MA: Harvard University Press).

Herrmann, B., Thoni, C., and Gachter, S. (2008). Antisocial punishment across societies. Science *319*, 1362–1367.

Hoffman, E., McCabe, K., and Smith, V.L. (1996). Social distance and other-regarding behavior in dictator games. Am. Econ. Rev. *86*, 653–660.

Humphrey, N.K. (1976). The social function of intellect. In Growing Points in Ethology, P.P.G. Bateson and R.A. Hinde, eds. (Cambridge: Cambridge University Press), pp. 303–317.

Keysar, B., Lin, S.H., and Barr, D.J. (2003). Limits on theory of mind use in adults. Cognition *89*, 25–41.

Kilner, J.M., Paulignan, Y., and Blakemore, S.J. (2003). An interference effect of observed biological movement on action. Curr. Biol. *13*, 522–525.

Lakin, J.L., and Chartrand, T.L. (2003). Using nonconscious behavioral mimicry to create affiliation and rapport. Psychol. Sci. *14*, 334–339.

Leslie, A.M. (1987). Pretense and representation: The origins of "theory of mind." Psychol. Rev. *94*, 412–426.

Liszkowski, U., Carpenter, M., and Tomasello, M. (2007). Pointing out new news, old news, and absent referents at 12 months of age. Dev. Sci. *10*, F1–F7.

Materna, S., Dicke, P.W., and Thier, P. (2008a). Dissociable roles of the superior temporal sulcus and the intraparietal sulcus in joint attention: a functional magnetic resonance imaging study. J. Cogn. Neurosci. *20*, 108–119.

Materna, S., Dicke, P.W., and Thier, P. (2008b). The posterior superior temporal sulcus is involved in social communication not specific for the eyes. Neuropsychologia *46*, 2759–2765.

Mauss, M. (1924). Essai sur le don. Forme et raison de l'échange dans les sociétés archaiques (London: Routledge). First published by Presses Universitaires de France in "Sociologie et Anthropologie."English translation by W.D. Halls, 1990.

Morris, J.S., Ohman, A., and Dolan, R.J. (1999). A subcortical pathway to the right amygdala mediating "unseen" fear. Proc. Natl. Acad. Sci. USA *96*, 1680–1685.

Newman-Norlund, R.D., van Schie, H.T., van Zuijlen, A.M., and Bekkering, H. (2007). The mirror neuron system is more active during complementary compared with imitative action. Nat. Neurosci. *10*, 817–818.

Nowak, M.A., and Sigmund, K. (1998a). Evolution of indirect reciprocity by image scoring. Nature *393*, 573–577.

Nowak, M.A., and Sigmund, K. (1998b). The dynamics of indirect reciprocity. J. Theor. Biol. *194*, 561–574.

Ochsner, K.N., and Lieberman, M.D. (2001). The emergence of social cognitive neuroscience. Am. Psychol. *56*, 717–734.

Ohman, A., and Mineka, S. (2001). Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. Psychol. Rev. *108*, 483–522.

Olsson, A., and Phelps, E.A. (2004). Learned fear of "unseen" faces after Pavlovian, observational, and instructed fear. Psychol. Sci. *15*, 822–828.

Onishi, K.H., and Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? Science *308*, 255–258.

Pelphrey, K.A., Singerman, J.D., Allison, T., and McCarthy, G. (2003). Brain activation evoked by perception of gaze shifts: the influence of context. Neuropsychologia *41*, 156–170.

Perner, J., and Leekam, S.R. (1986). Belief And Quantity - 3-Year Olds Adaptation To Listeners Knowledge. J. Child Lang. *13*, 305–315.

Phelps, E.A., O'Connor, K.J., Cunningham, W.A., Funayama, E.S., Gatenby, J.C., Gore, J.C., and Banaji, M.R. (2000). Performance on indirect measures of race evaluation predicts amygdala activation. J. Cogn. Neurosci. *12*, 729–738.

Pickering, M.J., and Garrod, S. (2004). Toward a mechanistic psychology of dialogue. Behav. Brain Sci. *27*, 169–190.

Premack, D., and Woodruff, G. (1978). Does the chimpanzee have a theory of mind? Behav. Brain Sci. *4*, 515–526.

Reed, K., Peshkin, M., Hartmann, M.J., Grabowecky, M., Patton, J., and Vishton, P.M. (2006). Haptically linked dyads: are two motor-control systems better than one? Psychol. Sci. *17*, 365–366.

Rilling, J.K., Sanfey, A.G., Aronson, J.A., Nystrom, L.E., and Cohen, J.D. (2004). The neural correlates of theory of mind within interpersonal interactions. Neuroimage *22*, 1694–1703.

Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. Annu. Rev. Neurosci. *27*, 169–192.

Samson, D., Apperly, I.A., Braithwaite, J.J., and Andrews, B.J. (2007). Seeing it your way: Cognitive processes underlying visual perspective taking. In Proceedings of EuroCogSci ′07: The European Cognitive Science Conference, S. Vosniadou, D. Kayser, and A. Protopapas, eds. (Lawrence Erlbaum Associates), pp. 909–909.

Sanfey, A.G., Rilling, J.K., Aronson, J.A., Nystrom, L.E., and Cohen, J.D. (2003). The neural basis of economic decision-making in the Ultimatum Game. Science *300*, 1755–1758.

Sato, W., Okada, T., and Toichi, M. (2007). Attentional shift by gaze is triggered without awareness. Exp. Brain Res. *183*, 87–94.

Saxe, R. (2006). Uniquely human social cognition. Curr. Opin. Neurobiol. *16*, 235–239.

Saxe, R., Xiao, D.K., Kovacs, G., Perrett, D.I., and Kanwisher, N. (2004). A region of right posterior superior temporal sulcus responds to observed intentional actions. Neuropsychologia *42*, 1435–1446.

Sebanz, N., Knoblich, G., and Prinz, W. (2003). Representing others' actions: just like one's own? Cognition *88*, B11–B21.

Sebanz, N., Bekkering, H., and Knoblich, G. (2006). Joint action: bodies and minds moving together. Trends Cogn. Sci. *10*, 70–76.

Seidner, L.B., Stipek, D.J., and Feshbach, N.D. (1988). A developmental analysis of elementary school-aged children's concepts of pride and embarrassment. Child Dev. *59*, 367–377.

Singer, T., Kiebel, S.J., Winston, J.S., Dolan, R.J., and Frith, C.D. (2004). Brain responses to the acquired moral status of faces. Neuron *41*, 653–662.

Smith, A. (1759). The Theory of Moral Sentiments (Indianapolis, IN: Liberty Clasics, 1982).

Sober, E., and Wilson, D.S. (1998). Unto Others (Cambridge, MA: Harvard University Press).

Stanley, J., Gowen, E., and Miall, R.C. (2007). Effects of agency on movement interference during observation of a moving dot stimulus. J. Exp. Psychol. Hum. Percept. Perform. *33*, 915–926.

Surian, L., Caldi, S., and Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. Psychol. Sci. *18*, 580–586.

Tomasello, M., and Carpenter, M. (2007). Shared intentionality. Dev. Sci. *10*, 121–125.

Valdesolo, P., and DeSteno, D. (2008). The Duality of Virtue: Deconstructing the Moral Hypocrite. J. Exp. Soc. Psychol. *44*, 1334–1338.

van Baaren, R.B., Holland, R.W., Kawakami, K., and van Knippenberg, A. (2004). Mimicry and prosocial behavior. Psychol. Sci. *15*, 71–74.

Warneken, F., and Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. Science *311*, 1301–1303.

Wedekind, C., and Milinski, M. (2000). Cooperation through image scoring in humans. Science *288*, 850–852.

Whalen, P.J., Rauch, S.L., Etcoff, N.L., McInerney, S.C., Lee, M.B., and Jenike, M.A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. J. Neurosci. *18*, 411–418.

Wicker, B., Keysers, C., Plailly, J., Royet, J.P., Gallese, V., and Rizzolatti, G. (2003). Both of us disgusted in My insula: the common neural basis of seeing and feeling disgust. Neuron *40*, 655–664.

Wimmer, H., and Perner, J. (1983). Beliefs About Beliefs - Representation and Constraining Function of Wrong Beliefs in Young Childrens Understanding of Deception. Cognition *13*, 103–128.

Zeitlyn, D. (2003). Gift economies in the development of open source software: anthropological reflections. Res. Policy *32*, 1287–1291.