Neurosciences and the Human Person: New Perspectives on Human Activities
Pontifical Academy of Sciences, *Scripta Varia* 121, Vatican City 2013
www.pas.va/content/dam/accademia/pdf/sv121/sv121-frithu.pdf

Are there innate mechanisms that make us social beings?

Uta Frith[1]

Introduction

We humans consider pride ourselves in being an ultra-social species with a strong desire to learn from each other and to cooperate with each other (Boyd & Richerson, 2009; Tomasello & Vaish, 2012). Folk psychology lets us believe that we have conscious control over our most treasured social abilities, such as empathy, fairness and morality, and that we pass them on to the next generation through teaching normative rules.

However, this folk belief does not fit with the compelling demonstration that we are not conscious of most of our cognitive and social abilities, and that these are already in place in early childhood well before normative rules are taught (Tomasello, 2008). Furthermore, social abilities of an apparently high level of complexity, such as altruistic helping, mind-reading, and reputation management, can be observed in non-human animals including fish, insects and birds (Frith & Frith, 2012). In the present paper, I will reflect on where our social abilities come from, how they are organised and what happens when they go wrong.

At present we struggle to understand the putative presence of conscious and unconscious systems in the mind. Daniel Kahneman (2011) elaborated on these systems, economically dubbing them I and II, or fast and slow. According to Kahneman, the unconscious system I is fast and powerful and rules our mental lives far more than we realise. However, it does submit occasionally to reasoning as presented by system II. The conscious system II is slow, weak and also more error prone than we realise. It provides justifications and rationalisations for behaviour that is actually caused by unconscious processes. I will argue that the fast system is based on *instincts*, and that this is where we should look for innate social mechanisms.

The problem with instincts

The term instinct has had an image problem, and understandably so. There are two main obstacles to using this unfashionable term. First, instincts

[1] UCL Institute of Cognitive Neuroscience and Interacting Minds Centre University of Aarhus

conceived as a sequence of rigidly hard-wired behaviour are incompatible with contemporary ideas of how the brain works. The very word hard-wired contradicts the idea of brain plasticity. Second, instincts are often defined as excluding learning, when learning is in many ways the essence of what the brain does. We learn all the time and adapt our brain to new conditions.

I propose to use the term 'start-up kit' to suggest that there is both a given predisposition and learning. Learning is needed to realise and to tune up any mental ability, however innate. I will continue to use the term innate, even though, like instinct, this also has had a bad press. Some difficult obstacles have to be overcome. For instance, it is inimical to emphasise inequality that is inherent in the notion of genetically based abilities and disabilities and smacks of genetic determinism. However, genetically based individual differences are undeniable, while genetic effects are probabilistic and act in interaction with environmental and social factors. Strong environmental effects can sometimes trump genetic effects. For example, genetically caused disorders are amenable to intervention, and gene therapy has been shown to be successful (Sheridan, 2011).

It is generally agreed that the brain is a prediction machine (Knill & Pouget, 2004). The prediction of the social behaviour of other agents is of as much value and importance as the prediction of events in the physical world. There is much less controversy in the assumption that the brain comes equipped with innate start-up kits for predicting the physical world. To mention just two examples: the ability to detect of numerical magnitude depends on an intact intra-parietal sulcus (Piazza *et al.*, 2007), and the ability to navigate in space depends on an intact hippocampal formation (Burgess & O'Keefe, 2011). It is encouraging that specific neural mechanisms have been identified for several cognitive processes, and it is highly likely that this is true also for the social mind/brain. But how can we know there are circumscribed social mechanisms, and how would we test the hypothesis that they rest on innate start-up kits? It is research on the cognitive basis of autism, which gives me strong reasons for this belief.

What do 'start-up kits' do?

The argument can be made that powerful learning mechanisms are sufficient to explain social behaviour. Social learning is likely to be the main driver for human adaptation over generations (Boyd, Richerson & Henrich, 2011; Heyes, 2012). If learning is so powerful, why require innate mechanisms at all? I believe we need them to explain social learning within a single brain. For one thing, life is too short to learn everything that can be learned, if there are no predispositions to set priorities. For another thing,

not everything can be learned equally easily and this is hard to explain if learning mechanisms don't have prior dispositions. For instance, fear of snakes can be learned quickly, but not fear of flowers (Cook & Mineka, 1989). Learning proceeds within remarkably strict limits. These are often referred to as constraints. However, I believe it is time to think of innate predispositions not as constraining learning, but as enabling learning.

It is also important to get rid of the idea that innate mechanisms must be present at birth and look for them only in young infants. Indeed there are some time bombs, which are set to detonate at different stages of life, e.g. sexual maturity and child bearing. Evolution has resulted in adaptive mechanisms, manifest in the way the brain is organised, and in the way learning gets off to a quick start when it is needed. This can happen at different ages.

For the purposes of this paper I assume that start-up kits are genetically programmed predispositions for specific computational processes, located in different circuits of the social mind/brain. They enable fast track learning in vital domains. This learning can be seen as a necessary calibration of the mechanism for given environmental conditions, with recalibration when circumstances change sufficiently. The idea is that there is a smoothly working mechanism that responds automatically to the right stimuli, like lock and key.

Autism as a crucible

Autism Spectrum Disorder (ASD) is a heterogeneous collection of disorders caused by a multitude of genetic and epigenetic causes (Geschwind, 2009). In view of this heterogeneity the fact that a clinical diagnostic category of autism exists is remarkable. One way to explain this coexistence of multiplicity at one level and unity at another level, is that all distal causes converge in a bottleneck. This bottleneck is the social brain, which develops atypically (Abrahams & Geschwind, 2010; Pelphrey *et al.*, 2011). At another level, the causal paths diverge again to give rise to impairments in a variety of social behaviours. These behaviours are also affected by other distal factors, such as education and learning, which alleviate or aggravate the condition. It is easy to picture this in a three-level framework (Morton & Frith, 1998) as shown in Figure 1. I believe that it is by investigating the nature of the mechanisms at the cognitive level that the anatomy of the social mind/brain can be laid bare.

I will focus here on a particular cognitive function, Theory of Mind, also referred to as ToM or mentalising (for a recent review see Apperley, 2012). I would place '*mentalising failure*' in the big oval space in the middle of Figure 1, with smaller ovals representing other hypothetical cognitive dysfunctions. The mentalising hypothesis claims to apply to all individuals on the autistic

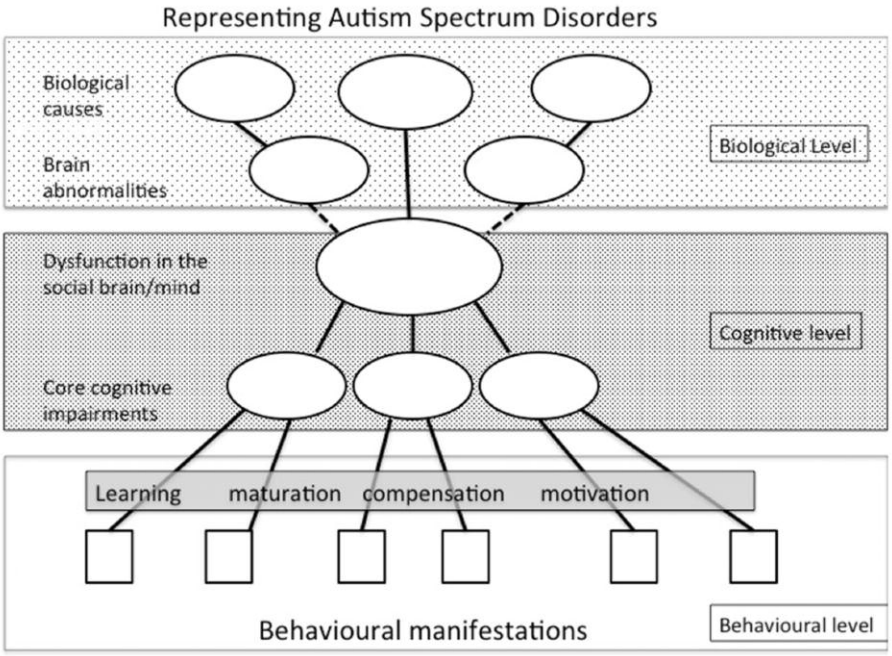

Figure 1. Three-level framework.

spectrum, regardless of what additional cognitive problems they might have. Empirical studies over the last 20 years have given some credence to this claim. They have shown that autistic individuals show atypical brain activity in the mentalising system of the social brain. They have also shown that their characteristic social impairments can be parsimoniously explained via mentalising failure (for a recent review see Frith, 2012). However, there may be additional dysfunctions that act as proximate causes of the social impairments seen in autism (for a recent review see Chevallier *et al.*, 2012).

Gotts *et al.* (2012) derived a summary picture of the social brain as revealed by autism. This is shown in Figure 2 (**see p. xxx**). The brain regions indicated in red are active when volunteers perform various social tasks in the scanner. For instance, they are active when a task involves mentalising, but not active, when it does not involve mentalising. The difference between these two conditions is what appears as red blobs in typical brain images. Over and over again the blobs have been found in these regions: medial prefrontal cortex, superior temporal sulcus, temporoparietal junction, temporal poles, basal temporal regions, inferior frontal gyrus, amygdala, pre-

cuneus and anterior insula. Together the regions form a network that we can call the social brain. Of course, to understand mental mechanisms it is not sufficient to know their location in the brain. However, the evolutionary history of the brain regions pinpointed can give clues to the origins of the mechanisms and how they became incorporated into genetic programmes.

A metaphor for the social mind/brain

I am going to use the visual prop of a house as a structure representing the human social mind/brain. I will act as an estate agent to tempt you to invest in this substantial property. I know you are not going to move in, it is after all a metaphor, but you may enjoy viewing it. I will guide you through its five floors and tell you about its ancient history. Two firms inhabit the property, Kahneman I and II. The bulk of the house belongs to the firm of Kahneman I; only the upper floor and the attic belong to the firm of Kahneman II.

The most important thing to know about the house is that it is a radio station. The daily business of the inhabitants is broadcasting and receiving news. When you view the house you can experience how utterly enthralled the inhabitants are by everything that is going on in other houses. You can also see that decisions have to be made constantly to switch on or off various listening and broadcasting devices. As a scrupulous estate agent, I will not conceal that some of the switches and information processing devices are liable to dysfunction. In extreme cases, parts of the firms may close down. This is not entirely surprising given the highly complex nature of the information they trade.

Here is a quick look at the floor plan followed by more detailed viewing.

Basement – Recognising other agents
Ground floor – Being part of a group
Bel-étage – Taking account of others' mental states
Top Floor – Intelligence services: explicit Theory of Mind
Attic Tower – The Self

Basement – Recognising other agents

The house has sound foundations. The cavernous basement contains ancient structures. This is tried and trusted machinery that never fails. It is shared with virtually all other animals, particularly mammals and is and deeply unconscious. This machinery allows us to recognise social agents, kin and non-kin, and moreover, individual others. Survival depends on ultra-fast learning on which agent to approach and which to avoid. These

mechanisms are in excellent order and suggest an innate predisposition to prioritise other agents over any other stimuli in the environment. This may be because social stimuli are intrinsically rewarding (Delgado, 2007; Krach *et al.*, 2010). The reward yields dividends in affiliation and love, but also in gaining status and winning competitions.

To help the visitor, there is a video in a roomy vault, which explains how mammals bond with their young to suckle and nurture them. This may have provided a powerful push towards a social capacity in the human brain that we refer to as love. We can glance into a lab producing potions made from hormones such as oxytocin, vasopressin and testosterone. The effects of the potions go from nurturing the young to sexual interaction and pair bonding to fierce parochialism and competition (Carter *et al.*, 2008; McCall &Singer, 2012).

Another video shows fighting deer with huge antlers. Here you can see a mechanism in action that can evaluate individual differences in prowess. This may also be relevant to recognition of dominance hierarchies. We go back to the first video and see that already at the stage of suckling, there is differentiation between siblings, so that some get a better place at the milk producing glands than others. It is not only rewarding to be near the top of the hierarchy, but it can be fatal to be at the bottom (e.g. Petticrew & Davey-Smith, 2012).

A slide show gives an overview over the ancient evolutionary roots of the machinery and we can see that this machinery may primarily serve selfish aims. We see flashing up the phrase 'nature red in tooth and claw' and realise that this applies to social lives in particular. The mechanisms come as integral parts of the house, but they improve by learning. Learning is not left to chance, but is guided by strong rewards. Rewards are strongly linked to affiliation, sexual reproduction, nurturing, having status and winning over rivals (e.g. Zink & Meyer-Lindenberg, 2012). It would seem difficult to subdue these robust mechanisms and the strong emotions they give rise to, and yet it is possible to do so. They can be silenced temporarily using conscious inhibition. This type of control, using cultural rules and moral guidelines is the job of the firm Kahneman II, at the top floor of the house.

Ground floor – Being part of a group

It is party time. If you are a gregarious extravert, you will love this floor. The mechanisms here are of particular value to animals whose survival depends on living in groups. These mechanisms here too are robust and well honed over millennia of evolution. They rarely fail, and if you are an introvert, you can always find an off switch. Still, you should hesitate to use this switch. The mechanisms can make life very rewarding for you. They foster

alignment to other members of your group. The result is a We-mode of social perception (Tuomela, 2007). Turn on the We-mode and you share experiences with others. In the basement, the I-mode is more often in use. The basement takes account of others in terms of how well they serve the Self, but this egocentric mode is bad manners here on the ground floor. You are enveloped and submerged in a greater whole. There is dance music with inviting rhythms; there is choral singing and in one room the Ode to Joy can be heard in the background: "Seid umschlungen Millionen…"

A video of a football game is on show, with the camera turned mainly on the crowd who chant and move as one. Another video shows herds of animals in the African Savannah. It all makes sense when you think about the ground-floor mechanisms as herd instinct (Raafat, Chater & Frith, 2009). Another video shows you that the alignment in motion and emotion also supports empathy (Singer & Lamm, 2009). But there are more complex forms of empathy, (Blair, 2005), and Kahneman II is in charge of them on the top floor.

You cannot be alone on this floor for long. On every video being demonstrated you see other people. The videos impress on you that human beings, like other social animals, have a strong tendency for compliance and conformity (Cialdini & Goldstein, 2004). They want to be where other human beings are. Doing as others do is rewarding and makes you feel benevolent towards those less fortunate than you (van Baaren *et al.*, 2004). In one room you are invited to play a game on a slot machine set up in a virtual web with other players. You soon find that a good way of winning is by copying others (Rendell *et al.*, 2010). Social learning pays. One neat mechanism works in such a way that eye gaze follows eye gaze. This makes social learning very simple: you automatically follow the gaze of another person.

There are some dark rooms, but I will show them to you. Here we can glimpse oppositional behaviour, and also aggression and violence. We see that an individual who behaves out of line with the group and sticks out, risks ostracism. Even the bare hint of ostracism makes the mechanism that promotes copying behaviour work harder (Over & Carpenter, 2008). While it is rewarding to conform, it is also painful not to conform (Haun, van Leeuwen & Edelson, 2012). A video shows that strong tendencies to do what others do result in social network formation (Christakis & Fowler 2012). They can transmit good feelings as well as bad; and they are responsible for all sorts of ideas spreading quickly. The analogy with infectious diseases is not far-fetched. Still, control systems can stop the spread.

More rooms with more video shows can be visited, as this is a very large floor. We can take part in a quick experiment. You are asked to imagine

yourself in an emergency. You have a split second to decide which other people you would turn to. The odds are that these others are similar to you. This automatic tendency is also the basis of social prejudice. We can do another experiment to show how instantly we can classify others into an in-group or outgroup. It takes only the slightest prompts (for a review see Shkurko, 2012). A horror movie about Zombies shows how frighteningly easy it is for human beings of all ages to turn off prosocial group oriented behaviour, and become aggressive and nasty to members of an out-group (Bernhard, Fischbacher & Fehr, 2006).

There is so much to see on this floor, so many emotions to feel, that while hurrying through we can only get a glance at what regulates group behaviour. In human societies there is an abundance of culturally evolved rules and regulations, some enshrined in law. Here Kahneman II is in charge. But the ground floor has its own built-in rules. These are the same that play out in automatic fashion throughout the animal kingdom. For instance, a tit-for-tat strategy is common across many social species. We recognise the almost uncontrollable feeling of revenge, but this mechanism counteracts the invasion by free riders (Raihani & Bshary, 2011). Troubleshooting like this is something that the basically altruistic ground floor often has to do. How come it is so altruistic? One fundamental drive in many social species is an egalitarian drive, whose evolutionary roots have been modelled (Gavrilets, 2012). Hence, fairness is rewarding (Tabibnia & Lieberman, 2007) and inequity is aversive. Aversion to inequity can induce individuals to punish selfish culprits even if it is at a cost to themselves (Raihani & McAuliffe, 2012). A brief slide presentation of neighbouring properties shows us that not all houses are the same. There are a minority of individuals who value their own selfish interest much more than equity (Haruno & Frith, 2010; Morishima *et al.*, 2012). Group oriented individuals too can become free riders, particularly if they see other free riders doing it and getting away with it too.

It is difficult to see inside the hub of the ground floor machinery: all sensory modalities have a role to play, the movement, sound, smell and touch of others, but visual cues are particularly important. Many believe that the real engine here is the mirror neuron system. Located in the inferior frontal gyrus and the intraparietal regions, this serves as a neural basis for the links between perception and action (Kilner, Friston & Frith, 2007). Innate mechanisms might set up the capacity to have a mirror neuron system but processes such as association learning are also likely to be involved (Heyes, 2011). Perhaps the most amazing thing about the hub is that the same mechanism applies to glue together perception and action within one and the same person as across persons. This truly is a We-mode engine.

You might ask whether the ground floor mechanisms are exclusively in the service of the We-mode or whether they can they also work in the I-mode. I think they can. For instance, fairness tends to be monitored within in-groups, and thus implicitly compares individuals (Fehr, Bernhard & Rockenbach, 2008). Altruism can be seen as the easy thing to do in the We-mode, or the effortful thing to do in the I-mode. A general house rule is: save energy and always do the easy thing.

The understanding of group hierarchies is another crucial aspect of group living: you need to know your place (Zink *et al.*, 2008). We have already met examples of mechanisms that underlie dominance and status in the basement. This was when they involved an egocentric perspective. However, status can also be a group concern. It seems that rapid switching between I-mode and We-mode is part of our social nature and there are clear benefits to be had from switching. There is direct reciprocity and the I-mode is sufficient here: you scratch my back and I scratch yours. But there is also indirect reciprocity: you scratch my back, and somebody else will scratch your back, or perhaps your friend's back (Nowak & Sigmund, 1998). This demands a We-mode.

I have to point out that things sometimes do go wrong on the ground floor. In the case of autism, suggestions have been made for a number of the mechanisms located here to be faulty. Processes underlying eye gaze following have a reasonably strong claim of being critically abnormal in autism, and there are problems in spontaneously copying others and in experiencing emotional contagion. Autistic people stick out, and this leads to ostracism, for instance in the form of bullying or neglect. Anecdotal observations suggest that many autistic people actually like to be where other people are; many are keen, even desperate, to have friends. They join clubs and maintain web-based support groups. They learn from observing other people; they imitate them, if sometimes clumsily. They pick up on social stereotypes and can make in-group – out-group distinctions (Hirschfeld, Bartmess, White & Frith, 2007). For instance they like to distinguish themselves from NTs, or neurotypical people and are proud to be different. There is then a strong possibility that the mechanisms on this floor of the house are dissociable and that autism provides a way to dissociate them.

Bel-étage – Taking account of other's mental states

This floor is a big selling point of the house. You can walk through some spacious public rooms, set up ready for visitors such as you. Here is the place of a rather more delicate mechanism, not as ancient as the ones on the lower floor and more exclusive and also more fragile. In fact this machinery is rare

in the animal kingdom. It makes humans social in a rather special way. It senses and processes information at ranges that are invisible to sensory modalities: It automatically takes into account your point of view, your desires, your intentions and beliefs, in short, your mental states that you might have thought were entirely private. In fact they give invaluable clues as what you are going to do next and that is what the information processing machines on this floor are interested in. They can read your wishes before you even realise you have them and can extrapolate from your beliefs what events will surprise you. This ability is known as *implicit mentalising*. It is a We-mode thing.

This automatic mind reading mechanism is well worth having. It saves time, effort and errors, because it spares you from having to make complex conscious inferences just to know that someone who has not seen an event will not know about it. You see a demo where you watch a little boy who believes that an empty box contains sweets: you can instantly anticipate that he will be disappointed when looking in the empty box.

A slide show explains that Apperly & Butterfill (2009) argue for two systems to track beliefs and other mental states, one implicit, and one explicit. Implicit mentalising is present in human babies (Onishi & Baillargeon, 2005; Kovacs, Teglas & Endress, 2010) and in some non-human animals (Emery & Clayton, 2009; Bugnyar, 2011). The explicit model is on the top floor, belonging to Kahneman II. The idea of two models chimes in with the finding that in autism implicit mentalising remains faulty even in very able adults who have acquired excellent skill in explicit mentalising (Senju *et al.*, 2009; Frith, 2012).

If you like interesting conversations, this is the floor to be. Mentalising supports rhetoric and ostensive communication, the hallmark of human social communication. By this I mean, that there is not simply an exchange of information between agents, but agents try and influence each other's minds. Inevitably there is a dark side, namely the aim to mislead the other individual, perhaps by concealing information, perhaps by implanting false beliefs. A video shows that clever birds can excel at this (Clayton, Dally & Emery, 2007). Let us glance also at some shining exhibits of beneficial processes that thrive on implicit mentalising: spontaneous helping (Buttelmann, Carpenter & Tomasello, 2009), automatic perspective taking (Samson *et al.*, 2010), and natural pedagogy (Csibra & Gergely, 2009). Spontaneous helping behaviour has been shown in human toddlers, because they recognise when others need help. It is relatively difficult to demonstrate in chimpanzees and other animals (Rekers, Haun & Tomasello, 2011).

What are the stimuli that trigger the implicit mentalising system? They are delivered through efficient dumb waiters from the ground floor and base-

ment. These are cues that suggest agency and are abundantly emitted by other agents. This includes eye gaze and facial expressions. Delays in mentalising performance by deaf and blind children suggest that trigger stimuli are multimodal and do not rely on one sense alone. On this floor the stimuli coming from other agents are sorted into intentional and accidental ones. Only intentional stimuli demand attention. Ostensive stimuli, such as calling your name or intently looking at you, instantly ignite brain regions that are part of the brain's mentalising system (Kampe, Frith & Frith, 2003).

How is the information processed that is signalled by the igniting stimuli? A model of an illuminated brain shows a highly connected network flashing when stimuli are delivered from the dumb waiters. We see lighting up medial prefrontal cortex, superior temporal sulcus at the temporal-parietal junction, parts of the basal temporal lobes, and amygdala regions (Frith & Frith, 2006). Studies have shown that in autism the connectivity in this network is much weaker (Castelli *et al.*, 2002; Kana *et al.*, 2009). There are some helpers provided by the lower floors. For instance, emotional contagion and the categorisation of another agent as a member of the in-group, facilitates attribution of mental states. This could be simply because the mental states are shared in this case, and it is not necessary to compute a different point of view.

While looking at the model you can see some sophisticated equations being displayed (Hampton, Bossaerts & O'Doherty, 2008; Yoshida, Dolan & Friston, 2008; Behrens, Hunt & Rushworth, 2009)). Maybe it is these equations that guide neuronal processes when they automatically compute causal explanations where causes are mental states, not physical states.

Top floor – Intelligence Services: explicit Theory of Mind

At last we come to the realm of Kahneman II. 'No children under 4 admitted' it says on the door. This floor is furnished in minimalist style, flooded with light and far removed from the boisterous bonhomie of the ground floor. Rational analytic thought pervades the atmosphere. Here we can find some tasteful plaques that document stages in the triumph of cultural evolution over biological evolution, awards for religion, for philosophy, for science, and many more. The firm of Kahneman II lives in relative isolation from the rest of the house but does intervene on occasion, although it is often too late and too little.

The main exhibit in one room is explicit mentalising. We see the props for classic false belief tasks, starting with Wimmer & Perner's (1983) Maxi paradigm. All test explicit mentalising ability: Maxi didn't know his mother had moved the chocolate and so he looks for it in a place where it no longer

is. This accomplishment is expected around the fifth year of life. The understanding of second order false beliefs, such as white lies and double bluff, does not lag far behind, but refinements are observed at later ages and adolescence represents a stage where changes in brain regions that support explicit mentalising show some radical changes (Dumontheil, Apperly & Blakemore 2010). Thus, the contribution of the medial prefrontal cortex in the network wanes in favour of the contribution of parts of the superior temporal cortex. Metacognition in the service of social communication is a work in progress. It can be argued that meta-cognition in the sense of 'thinking about thinking' is a logical consequence of explicit mentalising (Carruthers, 2009), giving rise to such notions as 'I think that I think and I believe that my thoughts are different from other people's thoughts' (Leslie, 1987).

Language is one of the most important tools for our self-awareness and acts as Chief Executive Officer in the firm of Kahneman II. It is indispensible for the explicit form of mentalising and drives its development. This is shown by the studies of mentalising ability over two generations of the Nicaraguan deaf community (Pyers & Senghas, 2009). Here, the first generation lacked any signs for mental states, and failed badly on standard false belief tasks. By the second generation, sign language had become established, and now everyone succeeded on the tasks.

Conscious thought and self-awareness depend on mechanisms that are not automatic, but require an effort to use, and that is why it is a slow system. When the wheels are turning, remarkably interesting processes take place. Education and learning has a vital role to keep them running. You can put on headphones to hear lectures on social etiquette; on how to make friends and influence people; on what it means to be responsible for your actions and what is involved in making a moral choice. The most impressive space is a gallery displaying architectural models of schools, universities, churches, law courts and museums. These models illustrate the enormous reach of Kahneman II over time to set up institutions that regulated group living and improve cooperation.

There are some secret rooms, labelled Security. Here a spymaster and a spin-doctor analyse the signals received from other houses for their truth value. They can also send strategically placed signals, for example, insincere flattery. A rogue's gallery shows cartoons telling of trickery, deception, persuasion and outright manipulation. There is an elaborate monitoring system dedicated to the business of reputation management (Tennie, Frith & Frith, 2010). Security services care for the maintenance of the house and decide on repairs and embellishments. Diplomacy rules here and this means that the system is flexible enough to forgive breeches of trust and solve social

dilemmas (Van Lange & Joireman, 2008). This system can override the rigid tit-for-tat mechanism operated on the ground floor.

I speculate that the flexible and even devious social processes in Kahneman II cannot be traced to innate social mechanisms. I find it unlikely that there are innate predispositions to sue rules politeness or to present diplomas to reward distinguished service. Instead these behaviours emerge at various stages of civilisation and culture. They can be lost, however useful they are in providing the oil to grease the machinery of complex social relationships in evolved societies.

There is another reason that I think no predispositions are needed to acquire explicit mentalising and all that follows from it. We know that in autism explicit mentalising can be acquired without any sign of implicit mentalising being present (Senju *et al.*, 2009). This separately acquired mentalising works best when off-line, for instance, using written communication with the cost of using a slow system. Of course, not all autistic individuals can become so accomplished, and this would be due to additional cognitive problems, which restrict basic information processing capacity. Individuals with impaired intellectual ability and without language struggle to acquire explicit mentalising and other tools that can be used for social manipulation. If there are cases where deficits in Kahneman II processes can be observed, I would not look for innate social mechanisms. Instead I would look for causes in lack of education, lack of cultural knowledge, and lack of general intellectual resources.

Attic – The Self

By a spiral staircase we can enter the attic and come to a tower. The surrounding walls are entirely made of mirrors and windows. The tower is the abode of the Self and my metaphor for the Self is a transparent balloon. It floats in the tower and is in constant need of being inflated. This is done by feeding its vanity (Sharot, Korn & Dolan, 2011). Obviously, the Self is selfish, but it can disguise selfishness by ostentatiously worn altruism. The Self is snobbish and identifies with conscious processes. But it has some inkling that there are many floors in the house. Its vocation is to rise above them. It can also ignore them, for instance when condemning prejudice even though an automatic in-group – out-group separation is still happening on the ground floor. The Self strives to be a free agent and it can occasionally impose its will and exert control over the rest of the house (Filevich, Kühn, & Haggard, 2012). However, the real power of control resides in many other switches on the lower floors that continuously turn on and off the complex machinery that sits there.

A sketch of the house is presented in Figure 3 (**see p. xxx**). It is meant to serve as a mnemonic for the different mechanisms that may be resident in our social brain.

Social mechanisms in autism

Now that we have seen around the house, it is possible to think again about autism. In the following table I list social abilities that have been investigated in autism. I have ordered the studies in terms of their likelihood of being impaired in all or most individuals of the autism spectrum. My belief is that some of them will turn out to qualify as innate mechanisms, but this will depend on whether evidence can be obtained, first, from their clearly identifiable neural basis; second, from their specific dysfunction in well defined neuropsychological conditions; third, from their evolutionary origin In most cases, such evidence has not yet been obtained.

Table 1.

Social processes likely to be faulty in ASD

1. Eye gaze processing (Pelphrey *et al*., 2005; Grice *et al*., 2005; Kylliäinen *et al*., 2012)
2. Detecting biological motion (Kaiser & Pelphrey, 2012; Naeckaerts *et al*., 2012)
3. Mimicry/Rapport (Gallese *et al*., 2012; Scambler *et al*., 2007)
4. Self vs other distinction (Happé, 2010; Lombardo *et al*., 2010)
5. Mentalizing (Castelli *et al*., 2002; Kana *et al*., 2009; Yoshida *et al*., 2012)
    a. Perspective taking level 2 (Hamilton *et al*., 2009)
    b. Recognising social emotions (Shamey-Tsoory, 2008)
    c. Moral judgement (Moran *et al*., 2011; Gleichgerrcht *et al*., 2012)
    d. Reputation management (Izuma & Adolphs, 2011; Chevallier *et al*., 2012)

Social processes not likely to be faulty in ASD

1. Detecting agents (Johnson, 2003)
2. Identifying other's goals (Hamilton, 2009; Falck-Ytter, 2010)
3. Mirror neuron function (Hamilton, 2009; Leighton *et al*., 2008)
4. Attachment (Rutgers *et al*., 2004)
5. Salience of social stimuli (Fletcher-Watson, 2008; New *et al*., 2009)
6. Perspective taking, level 1 (Zwickel *et al*., 2011)
7. Cooperation, spontaneous helping (Colombi *et al*., 2009; Liebal *et al*., 2008)
8. Attributing social stereotypes of gender and race (Hirschfeld *et al*.)
9. Managing social hierarchies? (White *et al*., 2006)
10. Ingroup-outgroup formation? (Hirschfeld *et al*.)
11. Conformity? (Bowler & Worley, 1994)
12. Fairness/Inequity aversion? (Hill, Sally & Frith, 2004)

Social processes likely to be faulty, but not specific to ASD

1. Alexithymia (Silani *et al*., 2008; Bird *et al*., 2010)
2. Empathy (Singer *et al*., Jones *et al*., 2010; Bird *et al*., 2010)
3. Face processing (McPartland *et al*., 2011; Weigelt, Koldewyn, & Kanwisher, 2012)

As the Table shows in the case of autism the metaphorical house is perfectly sound in many ways. Sure, there are hits sustained by a number of separable social mechanisms. But there are also many putative social mechanisms that seem to be working well. Implicit mentalising is clearly the worst affected. However, it is interesting to speculate whether this mechanism is decomposable into more basic components and whether these too are compromised in autism. There are some promising computational models for mentalising (Hampton, Bossaerts & O'Doherty, 2008; Yoshida, Dolan & Friston, 2008; Behrens, Hunt & Rushworth, 2009), and relevant brain circuits have been pinpointed already.

The Table includes potential faults in top-down control mechanisms. For example, a number of researchers (e.g. Spengler, Bird & Brass, 2010; Cook & Bird, 2012; Grecucci *et al.*, 2012; Wang & Hamilton, 2012) have carried out studies of mimicry and argue that the problem here resides not in the mirror mechanism, but in a lack of top-down control. Chawarska, Makari & Shic (2012) take on the case of eye gaze and provide data that supports the hypothesis that lack of social attention is context dependent. Thus, they showed that autistic toddlers spend less time gazing at a face, relative to other children, only in situations where an adult makes a bid for their attention. Still even then, they look more at the adult compared to situations when there was no such bid. These types of argument point to new theories. These do not postulate problems in innate social mechanisms, but instead they point to problems in the control of these mechanisms. This is reminiscent of contemporary research on genetic diseases where the spotlight of attention has moved towards regions in the genome that switch on and off genes, rather than the genes themselves.

In the house metaphor I deliberately placed some highly active switches, for instance, between I-mode and We-mode. Perhaps a fault in this switch might explain what causes the characteristic egocentrism in autistic social interaction. To me autism has suggested an 'absent self', which is ironic, since I suggest that there is both too much and too little self (Frith, 2003; Frith, 2008). However, I could imagine that if the switch was stuck in the We-mode, there would be too little self, and if stuck in the I-mode, too much.

Two more points can be made through the Table. First, it is time to acknowledge that a remarkable amount of sociability can be present in autistic individuals, some of it still waiting to be revealed. Second, it is time to attend to possible overlap of pathologies in different disorders. Perhaps this is how we can understand that problems, such as lack of empathy for other people's feelings, or alexithymia, the inability to identify own feelings, are not restricted to autism but point to other types of pathology of the social brain.

Concluding remarks

The metaphor of the house serves to illustrate the complexity of the social mind/brain, up to a point. However, its main function may be to emphasise the need to search more deeply for the underlying neuro-cognitive mechanisms that are the secret drivers of our complex social world. What the metaphor makes easy to see is that there are likely to be many conditions of neuro-developmental origin, which are characterised by a variety of social deficits. The autism spectrum needs to be taken apart into distinct subgroups. Some of the subgroups may be characterised by failures in one or more of the social mechanisms themselves, while others may have problems only in the top-down control of these mechanisms.

While I believe that eventually a considerable number of innate social mechanisms will be found residing in Kahneman's System I, I also believe that most mechanisms in Kahneman's System II may not have an innate basis. So far no specific failure in this domain has been identified in autism over and above impairments in intellectual ability. For this reason I believe processes that are part of System II can usefully guide compensatory learning. Thus, autistic individuals can acquire explicit mentalising. The surprising picture that emerges from the mosaic of putative social mechanisms is that they seem rather independent. That is, a very basic mechanism, such as biological motion detection may be impaired, but this does not apparently affect the ready categorisation of in-groups and out-groups, nor the ability to show spontaneous empathy. It remains to be seen what kind of interdependence there is and what kind of compensatory learning can be achieved when biologically rooted impairments restrict social life.

References

[1] Abrahams B.S., Geschwind D.H. (2010). Connecting genes to brain in the autism spectrum disorders. *Arch Neurol*. 67(4):395-9. Review.

[2] Apperly, I. (2012). What is Theory of Mind? Concepts, cognitive processes and individual differences. *Quarterly Journal of Experimental Psychology*, 65, 825-39.

[3] Apperly I.A., Butterfill S.A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychol Rev*. 116(4):953-70.

[4] Behrens T.E., Hunt L.T., Rushworth M.F. (2009). The computation of social behavior. *Science*. 324(5931):1160-4.

[5] Bernhard H., Fischbacher U., Fehr E. (2006). Parochial altruism in humans. *Nature*. 442(7105):912-5.

[6] Bird G., Press C., Richardson D.C. (2011). The role of alexithymia in reduced eye-fixation in Autism Spectrum Conditions. *41*. 41(11):1556-64.

[7] Bird, G., Catmur, C. Silani, G., Frith, C., & Frith, U. (2006). Attention does not modulate neural responses to social stimuli in autism spectrum disorders, *Neuroimage*, 31, 1614-24.

[8] Bird, G., Silani, G., Brindley, R., White, S. Frith, U., & Singer, T. (2010). Em-

pathic brain responses in insula are modulated by levels of alexithymia but not autism. *Brain*, 133(Pt 5),1515-25.

[9] Blair R.J. (2005). Responding to the emotions of others: dissociating forms of empathy through the study of typical and psychiatric populations. *Consciousness and Cognition*, 14, 698-718.

[10] Bowler D.M. & Worley K. (1994). Susceptibility to social influence in adults with Asperger's syndrome: a research note. *J Child Psychol Psychiatry*. 35(4): 689-97.

[11] Boyd R., Richerson P.J., Henrich J. (2011). The cultural niche: why social learning is essential for human adaptation. *Proc. Natl. Acad. Sci. U.S.A.* 108 Suppl 2:10918-25.

[12] Boyd R., Richerson P.J. (2009). Culture and the evolution of human cooperation. *Philos Trans R Soc Lond B Biol Sci*. 364(1533):3281-8.

[13] Bugnyar T. (2011). Knower-guesser differentiation in ravens: others' viewpoints matter. *Proc Biol Sci*. 278(1705): 634-40.

[14] Burgess, N. & O'Keefe, J. (2011). Models of place and grid cell firing and theta rhythmicity. *Curr Opin Neurobiol*. 21(5):734-44.

[15] Buttelmann D., Carpenter M., Tomasello M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition*, 112(2);337-42.

[16] Carruthers P. (2009). How we know our own minds: the relationship between mindreading and metacognition. *Behav Brain Sci*. 32(2):121-38; discussion 138-82.

[17] Carter C.S., Grippo A.J., Pournajafi-Nazarloo H., Ruscio M.G., Porges S.W. (2008). Oxytocin, vasopressin and sociality. *Prog Brain Res*. 170:331-6.

[18] Castelli, F., Frith, C.D., Happé, F., & Frith, U. (2002). Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. *Brain*, 125, 1839-1849.

[19] Chawarska, K. Makari S. & Shic F. (2012). Context modulates attention to social scenes in toddlers with autism. *J Child Psychol Psychiatry*. 53(8):903-13.

[20] Chevallier C., Kohls G., Troiani V., Brodkin E.S., Schultz R.T. (2012). The social motivation theory of autism. *Trends Cogn Sci*. 16(4):231-9.

[21] Chevallier, C., Molesworth, C. & Happé. F. (2012). Diminished social motivation negatively impacts reputation management: Autism Spectrum Disorders as a case in point. *PLoS ONE*, 7 (2012), p. e3110

[22] Christakis N.A. & Fowler J.H. (2012). Social contagion theory: examining dynamic social networks and human behavior. *Stat Med*. 2012 E-Pub.

[23] Cialdini, R.B., & Goldstein, N.J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology*, 55, 591-621.

[24] Clayton N.S., Dally J.M., Emery N.J. (2007). Social cognition by food-caching corvids. The western scrub-jay as a natural psychologist. *Philos Trans R Soc Lond B Biol Sci*. 362(1480):507-22.

[25] Colombi C., Liebal K., Tomasello M., Young G., Warneken F., Rogers S.J. (2009). Examining correlates of cooperation in autism: Imitation, joint attention, and understanding intentions. *Autism*. 13(2):143-63.

[26] Cook J.L., Bird G. (2012). Atypical social modulation of imitation in autism spectrum conditions. *J Autism Dev Disord*. 42(6):1045-51.

[27] Cook & Mineka S. (1989). Observational conditioning of fear to fear-relevant versus fear-irrelevant stimuli in rhesus monkeys. *J Abnorm Psychol*. 98(4):448-59.

[28] Csibra G., Gergely G. (2009). Natural pedagogy. *Trends Cogn Sci*. 13(4):148-53.

[29] Delgado M.R. (2007). Reward-related responses in the human striatum. *Ann NY Acad Sci*. 1104:70-88.
[30] Dumontheil I., Apperly I.A., Blakemore S.J. (2010). Online usage of theory of mind continues to develop in late adolescence. *Dev Sci*. 13(2):331-8.
[31] Emery N.J., Clayton N.S. (2009). Comparative social cognition. *Annu Rev Psychol*. 60:87-113.
[32] Falck-Ytter, T. (2010). Young children with autism spectrum disorder use predictive eye movements in action observation. *Biol Lett*. 6(3):375-8.
[33] Fehr E., Bernhard H., Rockenbach B. (2008). Egalitarianism in young children. *Nature*. 454(7208):1079-83.
[34] Filevich, E., Kühn, S. & Haggard, P. (2012). Intentional inhibition in human action: the power of 'no'. *Neurosci Biobehav Rev*. 36(4):1107-18.
[35] Fletcher-Watson S., Leekam S.R., Findlay J.M., Stanton E.C. (2008). Brief report: young adults with autism spectrum disorder show normal attention to eye-gaze information-evidence from a new change blindness paradigm. *J Autism Dev Disord*. 38(9):1785-90.
[36] Frith, C.D., Frith, U. (2006). The neural basis of mentalizing. *Neuron*, 50(4) 531-4.
[37] Frith C.D. & Frith U. (2008). Implicit and explicit processes in social cognition. *Neuron*. 2008 Nov 6;60(3):503-10.
[38] Frith, C.D. & Frith, U. (2012). Mechanisms of Social Cognition. *Annual Review of Psychology*, 63, 287-313
[39] Frith, U. (2004). Emmanuel Miller lecture: Confusions and controversies about Asperger syndrome, *Journal of Child Psychology and Psychiatry*, 45, 672-686.
[40] Frith, U. (2008). *A Very Short Introduction to Autism*. Oxford: Oxford University Press
[41] Frith, U. (2012). Why we need cognitive explanations of autism. *Quarterly Journal of Experimental Psychology* [August 21 E-pub].
[42] Frith, U. (2001). Mindblindness and the brain in Autism. *Neuron*, 32, 969-979.
[43] Gallese V, Rochat MJ, Becchio C. (2012). The mirror mechanism and its potential role in autism spectrum disorder. *Dev Med Child Neurol*. E-pub.
[44] Gavrilets S. (2012). On the evolutionary origins of the egalitarian syndrome. *Proc. Natl. Acad. Sci. U.S.A*. 109(35):14069-74.
[45] Geschwind D.H. (2009). Advances in autism. *Annu Rev Med*. 60:367-80.
[46] Gleichgerrcht E., Torralva T., Rattazzi A., Marenco V., Roca M., Manes F. (2012). Selective impairment of cognitive empathy for moral judgment in adults with high functioning autism. *Soc Cogn Affect Neurosci*. E-Pub.
[47] Grecucci A., Brambilla P., Siugzdaite R., Londero D., Fabbro F., Rumiati R.I. (2012). Emotional Resonance Deficits in Autistic Children. *J Autism Dev Disord*. E-Pub.
[48] Grice S.J., Halit H., Farroni T., Baron-Cohen S., Bolton P., Johnson M.H. (2005). Neural correlates of eye-gaze detection in young children with autism. *Cortex*. 41(3):342-53.
[49] Gotts S.J., Simmons W.K., Milbury L.A., Wallace G.L., Cox R.W., Martin A. (2012). Fractionation of social brain circuits in autism spectrum disorders. *Brain*. 135(Pt 9):2711-25.
[50] Hamilton A.F. (2009). Goals, intentions and mental states: challenges for theories of autism. *J Child Psychol Psychiatry*. 50(8):881-92.
[51] Hamilton, A.F., Brindley, R., & Frith, U. (2009). Visual perspective taking impairment in children with autistic spectrum disorder. *Cognition*, 113, 37-44.
Hampton A.N., Bossaerts P., O'Doherty J.P.

(2008). Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proc. Natl. Acad. Sci. U.S.A.* 105(18):6741-6.

Happé, F. (200) Theory of mind and the self. *Ann NY Acad Sci.* 1001:134-44.

Haruno M., Frith C.D. (2010). Activity in the amygdala elicited by unfair divisions predicts social value orientation. *Nat Neurosci.* 13(2):160-1.

Haun, D., van Leeuwen, E. & Edelson, M. (2012). Majority influence in children and other animals. *Dev Cog Neurosci.* E-pub.

Hermans E.J., van Wingen G., Bos P.A., Putman P., van Honk J. (2009). Reduced spontaneous facial mimicry in women with autistic traits. *Biol Psychol.* 80(3):348-53.

Heyes C. (2011). Automatic imitation. *Psychol Bull.* 201137(3):463-83.

Heyes C. (2012). Grist and mills: on the cultural origins of cultural learning. *Philos Trans R Soc Lond B Biol Sci.* 367(1599):2181-91.

Hill, E.L., Berthoz, S., & Frith, U. (2004). Brief Report: Cognitive processing of own emotions in individuals with autistic spectrum disorder and their relatives. *Journal of Autism and Developmental Disorders*, 34, 229-235.

Hill, E.L., Sally, D. and Frith, U. (2004). Does mentalising ability influence cooperative decision making in a social dilemma? Introspective evidence from a study of adults with autism spectrum disorder. *Journal of Consciousness Studies*, 11, 1-18.

Hirschfeld L., Bartmess E., White S. & Frith U. (2007). Can autistic children predict behavior by social stereotypes? *Current Biology*, 17(12):R451-2.

Izuma K., Matsumoto K., Camerer C.F., Adolphs R. (2011). Insensitivity to social reputation in autism. *Proc. Natl. Acad. Sci. U.S.A.*, 108. 17302-17307.

Johnson, S.C. (2003). Detecting agents. *Philos Trans R Soc Lond B Biol Sci.* 358(1431):549-59.

Kahneman, D. (2011). *Thinking, Fast and Slow.* N.Y.: Straus and Giroux.

Kaiser M.D., & Pelphrey K.A. (2012). Disrupted action perception in autism: behavioral evidence, neuroendophenotypes, and diagnostic utility. *Dev Cogn Neurosci.* 2(1):25-35.

Kampe, K., Frith, C.D., Dolan, R.J., Frith, U. (2001). Attraction and gaze – the reward value of social stimuli. *Nature*, 413, 589.

Kana R.K., Keller T.A., Cherkassky V.L., Minshew N.J., Just M.A. (2009). Atypical frontal-posterior synchronization of Theory of Mind regions in autism during mental state attribution. *Soc Neurosci.* 4(2):135-52.

Kanwisher N., McDermott J., Chun M.M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. *J Neurosci* 17:4302-11.

Kilner J.M., Friston K.J., Frith C.D. (2007). The mirror-neuron system: a Bayesian perspective. *Neuroreport.* 18(6):619-23.

Knill D.C. & Pouget A. (2004). The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* 27(12):712-9.

Knoblich G., Sebanz N. (2008). Evolving intentions for social interaction: from entrainment to joint action. *Philos Trans R Soc Lond B Biol Sci.* 363(1499):2021-31.

Kovács Á.M., Téglás E., Endress A.D. (2010). The social sense: susceptibility to others' beliefs in human infants and adults. *Science*, 330(6012):1830-4.

Krach, S., Paulus, F.M., Bodden, M. & Kircher, T. (2010). The rewarding nature of social interactions. *Front Behav Neurosci.* 4:22.

Kylliäinen A., Wallace S., Coutanche M.N., Leppänen J.M., Cusack J., Bailey A.J., Hietanen J.K. (2012). Affective-motivational brain responses to direct gaze

in children with autism spectrum disorder. *J Child Psychol Psychiatry*. 53(7):790-7.

Leighton J., Bird G., Charman T., Heyes C. (2008). Weak imitative performance is not due to a functional 'mirroring' deficit in adults with AutismSpectrum Disorders. *Neuropsychologia*. 46(4): 1041-9.

Leslie, A.M. (1987) Pretense and representation. The origin of "theory of mind". *Psychological Review*, 94, 412-426.

Liebal K., Colombi C., Rogers S.J., Warneken F., Tomasello M. (2007). Helping and cooperation in children with autism. *J Autism Dev Disord*. 38(2):224-38.

Lombardo, M.V., Chakrabarti, B., Bullmore, E.T., Sadek, S.A., Pasco, G., Wheelwright, S.J., Suckling, J., MRC AIMS Consortium, Baron-Cohen, S. (2010). Atypical neural self-representation in autism. *Brain*, 133, 611-24.

McCall C., Singer T. (2012). The animal and human neuroendocrinology of social cognition, motivation and behavior. *Nat Neurosci*. 15(5):681-8.

McPartland J.C., Webb S.J., Keehn B., Dawson G. (2011). Patterns of visual attention to faces and objects in autism spectrum disorder. *J Autism Dev Disord*. 41(2):148-57.

Minio-Paluello I., Baron-Cohen S., Avenanti A., Walsh V., Aglioti S.M. (2009). Absence of embodied empathy during pain observation in Asperger syndrome. *Biol Psychiatry*. 65(1):55-62.

Moran J.M., Young L.L., Saxe R., Lee S.M., O'Young D., Mavros P.L., Gabrieli J.D. (2011). Impaired theory of mind for moral judgment in high-functioning autism. *Proc Natl Acad Sci U S A*. 108(7):2688-92.

Morishima, Y. Schunk, D., Bruhin, A., Ruff, C.C. & Fehr, E. (2012). Linking brain structure and activation in the temporoparietal junction to explain the neurobiology of human altruism. *Neuron*, E-pub.

Morton, J. & Frith, U. (1995). Causal modeling: Structural approaches to developmental psychopathology. In D. Cicchetti, & D. Cohen (Eds.), *Developmental Psychopathology* (pp. 357-90). New York, NY: Wiley.

Nackaerts E., Wagemans J., Helsen W., Swinnen S.P., Wenderoth N., Alaerts K. (2012). Recognizing biological motion and emotions from point-light displays in autism spectrum disorders. *PLoS One*. 7(9).

New J.J., Schultz R.T., Wolf J., Niehaus J.L., Klin A., German T.C., Scholl B.J. (2009). The scope of social attention deficits in autism: prioritized orienting to people and animals in static natural scenes. *Neuropsychologia*. 48(1):51-9.

Nowak M.A., Sigmund K. (1998). The dynamics of indirect reciprocity. *J Theor Biol*. 21;194(4):561-74.

Onishi, H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 8, 308(5719), 255-8.

Over, H. & Carpenter, M. (2008). Priming third-party ostracisms increases affiliative imitation in children. *Dev Sci*. 12(3) F1-F8.

Pelphrey, K.A., Morris, J.P., & McCarthy, G. (2005). Neural basis of eye gaze processing deficits in autism. *Brain*, 128(Pt 5):1038-48.

Pelphrey K.A., Shultz S., Hudac C.M., Vander Wyk B.C. (2011). Research review: Constraining heterogeneity: the social brain and its development in autism spectrum disorder. *J Child Psychol Psychiatry*. 52(6):631-44.

Petticrew M., Davey Smith G. (2012). The monkey puzzle: a systematic review of studies of stress, social hierarchies, and heart disease in monkeys. PLoS One. 2012;7(3)

Piazza M., Pinel P., Le Bihan D., Dehaene S. (2007) A magnitude code common

to numerosities and number symbols in human intraparietal cortex. *Neuron*, 18;53(2):293-305.

Press C., Richardson D., Bird G. (2010). Intact imitation of emotional facial actions in autism spectrum conditions. *Neuropsychologia*. 48(11):3291-7. Epub 2010 Jul 16.

Pyers J.E., Senghas A. (2009). Language promotes false-belief understanding: evidence from learners of a new sign language. *Psychol Sci*. 20(7):805-12.

Raafat R.M., Chater N., Frith C. (2009). Herding in humans. *Trends Cogn Sci*. 13(10):420-8.

Raihani N.J., Bshary R. (2011). Resolving the iterated prisoner's dilemma: theory and reality. *J Evol Biol*. 24(8):1628-39.

Raihani N.J., McAuliffe K. (2012). Human punishment is motivated by inequity aversion, not a desire for reciprocity. *Biol Lett*. 8(5):802-4.

Rekers Y., Haun D.B., Tomasello M. (2011). Children, but not chimpanzees, prefer to collaborate. *Curr Biol*. 21(20):1756-8.

Rendell L., Boyd R., Cownden D., Enquist M., Eriksson K., Feldman M.W., Fogarty L., Ghirlanda S., Lillicrap T., Laland K.N. (2010). Why copy others? Insights from the social learning strategies tournament. *Science*. 328(5975):208-13.

Rutgers A.H., Bakermans-Kranenburg M.J., van Ijzendoorn M.H., van Berckelaer-Onnes I.A. (2004). Autism and attachment: a meta-analytic review. *J. Child Psychol. Psychiatry*, 45, 1123-1134

Samson D., Apperly I.A., Braithwaite J.J., Andrews B.J., Bodley Scott S.E. (2010). Seeing it their way: evidence for rapid and involuntary computation of what other people see. *J Exp Psychol Hum Percept Perform*. 36(5):1255-66.

Scambler D.J., Hepburn S., Rutherford M.D., Wehner E.A., Rogers S.J. (2007). Emotional responsivity in children with autism, children with other developmental disabilities, and children with typical development. *J Autism Dev Disord*. 37(3):553-63.

Shamay-Tsoori, S.G. (2008). Recognition of 'fortune of others' emotions in Asperger syndrome and high functioning autism. *J Autism Dev Disord*. 38(8):1451-61.

Schneider D., Lam R., Bayliss A.P., Dux P.E. (2012). Cognitive Load Disrupts Implicit Theory-of-Mind Processing. *Psychol Sci*, 23(8):842-7.

Senju, A., Southgate, V., White, S. and Frith, U. (2009). Mindblind eyes: an absence of spontaneous Theory of Mind in Asperger Syndrome, *Science*, 325(5942):883-5.

Sharot, T., Korn, C.W. & Dolan, R. (2011). How unrealistic optimism is maintained in the face of reality. *Nat Neurosci*. 14(11):1475-9.

Sheridan C. (2011). Gene therapy finds its niche. *Nat Biotechnol*. 29(2):121-8.

Shkurko (2012). Is social categorization based on relational ingroup/outgroup opposition? A meta-analysis. *Soc Cogn Affect Neurosci*. 2012 Jul 30.

Silani, G., Bird, G., Brindley, R., Singer, T., Frith, C., & Frith, U. (2008). Levels of emotional awareness and autism: an fMRI study. *Social Neuroscience*, 3, 97-112.

Singer T., Lamm C. (2009). The social neuroscience of empathy. *Ann. N.Y. Acad. Sci*., 1156:81-96.

Tabibnia, G. & Lieberman, M.D. (2007). Fairness and cooperation are rewarding. *Ann. N.Y. Acad. Sci*., 1118, 90-101.

Tennie C., Frith U., Frith C.D. (2010). Reputation management in the age of the world-wide web. *Trends Cogn Sci*, 14, 482-8.

Tomasello, M. (2008). *Origins of Human Communication*. Cambridge Mass: MIT Press.

Tomasello M, Vaish A. (2012). Origins of Human Cooperation and Morality. *Annu Rev Psychol*. Jul 12.

Tuomela, R. (2007). *The philosophy of so-*

ciality. Oxford Scholarship Online. September 2007.

van Baaren, R.B., Holland, R.W., Kawakami, K., & van Knippenberg, A. (2004). Mimicry and prosocial behavior. *Psychological Science*, 15(1), 71-74.

van Lange, P. & Joireman, J.A. (2008). How we can promote behaviour that serves all of us in the future. *Social Issues and Policy Review*, 2(1):127-57.

Wang Y. & Hamilton, A. (2012). Social top-down response modulation (STORM): a model of the control of mimicry in social interaction. *Front Hum Neurosci*. 6:153. Epub.

Warneken F., Tomasello M. (2006). Altruistic helping in human infants and young chimpanzees. *Science*. 311(5765):1301-3.

Weigelt S., Koldewyn K., Kanwisher N., (2012). Face identity recognition in autism spectrum disorders: a review of behavioral studies. *Neurosci Biobehav Rev.* 2012 Mar;36(3):1060-84.

White S., Hill E., Winston J., Frith U. (2006). An islet of social ability in Asperger Syndrome: judging social attributes from faces. *Brain Cogn*. 61(1):69-77.

Williams J.H. (2008). Self-other relations in social development and autism: multiple roles for mirror neurons and other brain bases. *Autism Res.* 1(2):73-90.

Yoshida W., Dolan R.J., Friston K.J. (2008). Game theory of mind. *PLoS Comput Biol*. 4(12)

Yoshida W., Dziobek I., Kliemann D., Heekeren H.R., Friston K.J., Dolan R.J. (2010). Cooperation and heterogeneity of the autistic mind. *J Neurosci*. 30(26):8815-8.

Zink C.F., Meyer-Lindenberg A. (2012). Human neuroimaging of oxytocin and vasopressin in social cognition. *Horm Behav*. 61(3):400-9. 3

Zink C.F., Tong Y., Chen Q., Bassett D.S., Stein J.L., Meyer-Lindenberg A. (2008). Know your place: neural processing of social hierarchy in humans. *Neuron*. 58(2):273-83.

Zwickel J., White S.J., Coniston D., Senju A., Frith U. (2011). Exploring the building blocks of social cognition: spontaneous agency perception and visual perspective taking in autism. *Soc Cogn Affect Neurosci*. 6(5):564-71.

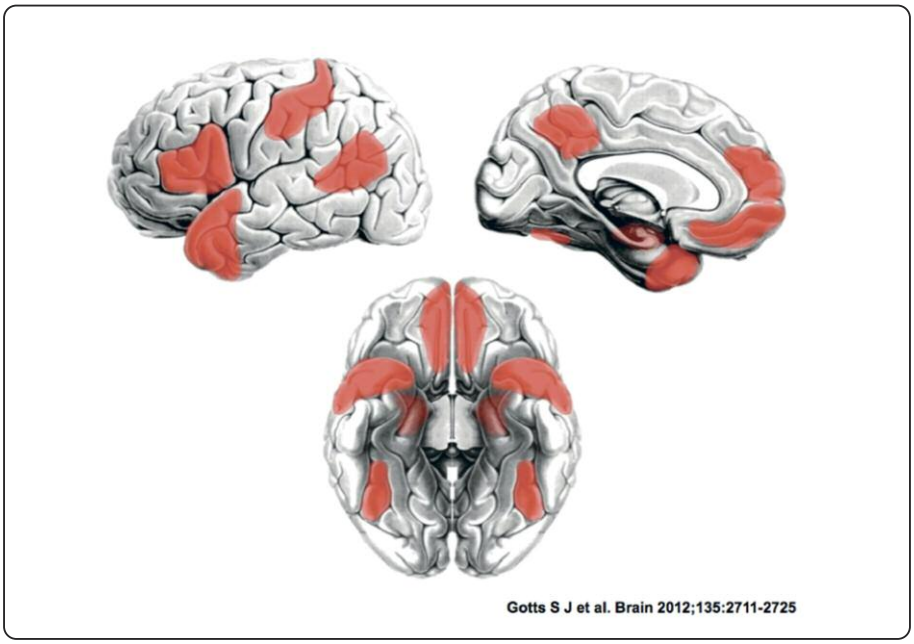

Figure 2. The social brain as revealed by autism.

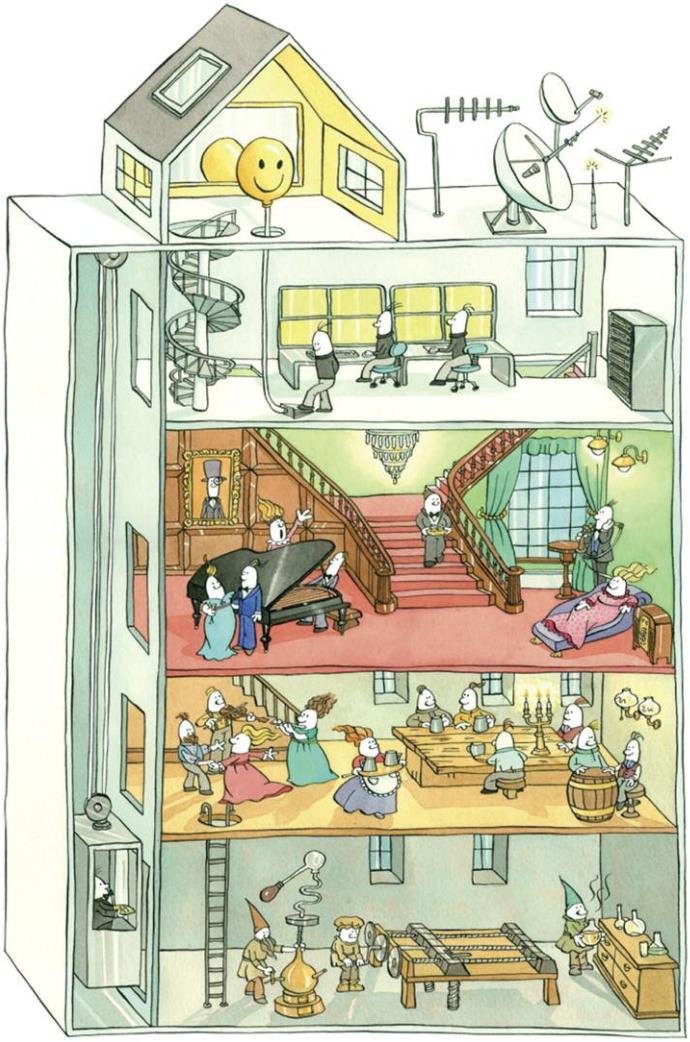

Figure 3. ©Jan McCafferty.