From 'Theory of Mind' in 1985 to 'Empathizing-Systemizing' in 2012[1].

Simon Baron-Cohen, Cambridge University, Autism Research Centre

It is now 27 years since the publication of the first article testing whether children with autism have a theory of mind (in 1985). The original study by my colleagues Alan Leslie and Uta Frith and I found that some 80% of children with autism failed this test that most typical children of 4 years old could pass. It concluded that this could not be attributed to general developmental delay or learning difficulties, because in a control group of children with Down Syndrome, some 80% of them passed. The results has been replicated dozens of times, suggesting this is a robust finding. This initial paper is considered important for 10 reasons.

First, until that point, there was no simple cognitive account of why these children have difficulties in forming social relationships and with the subtleties of communication (pragmatics), as well as why such children were relatively disinterested in fiction (such as pretend play). The identification of this 'cognitive deficit' had the power to make sense of these 3 cardinal signs of autism.

Second, philosophers such as Daniel Dennett had in 1980 argued that a 'theory of mind' (defined as the capacity to attribute mental states to others in order to make sense of and predict the other person's behaviour) is *essential* for social interaction and communication. With a theory of mind, other people's actions are easy to understand. Without a theory of mind, other people's actions seem confusing, uninterpretable, even terrifying, and would be expected to cause the child to withdraw from and avoid the social world. This triggered interest in this study not just from psychologists and psychiatrists, but also philosophers, demonstrating the value of applying philosophy to cognitive and clinical science.

Third, the study showed – at a glance – that a theory of mind is not a function of general intelligence but is *independent* of it. This was demonstrated by the inclusion of and results from the group with Down Syndrome. This pointed towards the possibility of a theory of mind being a separate, 'modular' aspect of human cognition. Again, this triggered the interest of philosophers such as Jerry Fodor who had argued in 1983 that the human mind might have some modular properties.

Fourth, the idea that the mind might have a module that was so basic to human social survival and functioning that without it, autism would result, raised the possibility that a theory of mind itself has a genetic basis and that a theory of

---

[1] This article is adapted from Baron-Cohen, S, (2011) Cognizione ed empatia nell'autismo. Dalla teoria della mente a quella del "cevello maschile estremo". In Agliuzzio, V, (ed) Bricks. Erikson; and Baron-Cohen (in press) Empathy deficits in autism and psychopaths: mirror opposites? In M. Banaji & S. A. Gelman (Eds.), Navigating the social world: What infants, children, and other species can teach us. Oxford University Press.

mind might be the end product of millions of years of human evolution. The fact that autism had in 1977 been understood to be a genetic condition, through Michael Rutter and Susan Folstein's first study of twins with autism, bolstered this possibility. This also triggered the interest of evolutionary psychologists, given that Leda Cosmides and John Tooby had in 1992 created the field of evolutionary psychology.

Fifth, the idea that a theory of mind had evolved prompted primatologists to question whether this ability was unique to humans, or whether our primate relatives such as chimpanzees had a theory of mind. The theory of mind study in autism was published soon after the first study of chimpanzees by David Premack and Guy Woodruff in 1978, bringing developmental psychologists into direct interaction with primatologists. Although the first chimpanzee study concluded that chimpanzees had elements of a theory of mind (they might be able to attribute goals and intentions to others), philosophers such as Dan Dennett provided commentaries on that study, prompting the design of the 'false belief' litmus test of a theory of mind. Subsequent studies by primatologists such as Daniel Povinelli and Mike Tomasello concluded that chimpanzees lack a theory of mind. This pointed to the idea that a theory of mind is one of the defining qualities of *Homo Sapiens*.

Sixth, the idea that a theory of mind might be modular in the mind led inexorably to the prediction that one might be able to 'image' a theory of mind using the new tools of functional brain scanning. A study by our group in 1994 using Single Photon Emission Computed Tomography (SPECT) implicated the orbito-frontal cortex as involved in a theory of mind. A second study in 1995 by Chris Frith and colleagues using Positron Emission Tomography (PET) implicated the left ventromedial prefrontal cortex as involved in a theory of mind. A third study in 1999 by our group using functional magnetic resonance imaging (fMRI) implicated the amygdala as another brain region involved in theory of mind. These 3 studies fitted nicely with a theory in 1990 by Leslie Brothers that a special circuit in the human brain was dedicated to processing information about the social world, which she called the 'social brain'. In all of these 3 studies, people with autism showed under-activity in each of these brain regions whilst they were taking engaged in a theory of mind task. This provided a neural basis to the presence of theory of mind in the typical population and to its impairment in autism.

Seventh, the idea that a theory of mind might be impaired in autism led our group to look at educational implications since the implication was that whilst a typical child might not need any special teaching in order to acquire and develop a theory of mind, a child with autism might need highly tailored special educational techniques in order to do so. This led to books such as *Teaching Children with Autism to Mindread* (Wiley, 1997) and to special educational software such as the DVD *Mindreading* (www.jkp.com/mindreading) in 2003, and to clinical trials demonstrating that such teaching did lead to children with autism learning to pass tests of theory of mind.

Eighth, the idea that by 4 years old a typical child might already have a theory of mind in place and that a child with autism might not prompted developmental psychologists to look at the developmental *precursors* of theory of mind. In particular, it led to the suggestion that the infancy skills of 'joint attention' and 'pretend play' at the end of the first year of life might be fundamental building blocks in the acquisition of theory of mind. This built on the research by child psychologists Jerome Bruner in 1975 and Alan Leslie in 1987 showing that by 9-14 months old joint attention (such as gaze monitoring and protodeclarative pointing using the index finger) and pretend play develop *spontaneously* in typical human toddlers, and for Leslie to argue that a lack of a theory of mind in autism might also thus explain why these other behaviours (joint attention, pretend play) might be absent in autism.

Ninth, the idea that precursors to a theory of mind such as pretend play and joint attention might be delayed or absent in autism led our group to develop an early-identification screening test called the Checklist for Autism in Toddlers (the CHAT) for paediatricians, Health Visitors, General Practitioners/Family Physicians. This 5 minute behavioural test was intended to be used in frontline clinical practice, to determine if a child might need a full clinical assessment for autism. Thus, out of the basic science came an application for the health service and clinical practice. A study of infant siblings in families where there was already a child with autism showed that the CHAT could accurately predict which infants at 18 months old would go on to develop autism. A large population study of 18 month olds also showed that a child who failed on the CHAT had an 80% likelihood of going on to develop autism.

The CHAT was developed when classic autism was known about but when the related condition of Asperger Syndrome (AS) was not, and it was later found that the CHAT failed to predict AS. As such, it could not be used as a screening test for the whole autistic spectrum and this stimulated research by our group to revise the CHAT (now called the Q-CHAT) to detect such milder, more subtle cases such as AS. Such research is ongoing, but the CHAT showed the potential clinical applications of the original theory of mind research to public health.

Lastly, tenth, all of the above work was brought together in a slim monograph entitled *Mindblindness* (MIT Press, 1995) as a new theory of autism and as a new way to conceptualize this neurodevelopmental condition. Where previously autism was a puzzling condition that had elicited fanciful but non-evidence based theories of autism (such as these children being the product of parental neglect, by psychoanalyst Bruno Bettelheim), this provided a highly specific genetic-neurological-cognitive theory of the disability. Where previously autism had been poetically described as 'feral/wolf children' (as if they had developed outside of human culture) or as if 'living in a glass bubble' (mystically seen as unreachable) or even as if from another planet (and thus mystified by the humans they find on this planet), there was now a theory with plenty of evidence that made autism explicable in terms of non-mystical factors.

These 10 reasons provide some – but not all – of the reasons why the original study of autism and theory of mind was of interest to a wide range of sciences, disciplines, clinicians, and educators.

****

Like any theory, the initial formulation of the mindblindness theory needed some revision. 5 are summarized next.

First, the results of the original study indicated this was not an all or none phenomenon, since 20% of children with autism passed the theory of mind test. This was interpreted as a sign of *specific developmental delay*, rather than a total and permanent absence of a theory of a theory of mind. The finding by Francesca Happe that those children with autism who do pass this test tend to be at least 11 years old suggested the size of this delay was huge.

Second, the original study gave the impression that a theory of mind was categorical (either you have this skill/pass this test or do not), later research showed that when chronological and mental age-appropriate, 'advanced' theory of mind tests were used (such as the 'Reading the Mind in the Eyes' test, or the Faux Pas Test, from our group, or the Strange Stories test, or the Animated Triangles test from Uta Frith's group), a theory of mind deficit could be seen even in adults with autism and even in those with Asperger Syndrome, making it a core (i.e., universal) cognitive feature of autism spectrum conditions.

Third, for me, although theory of mind seemed modular, it did not exist in isolation but could be seen as working (in the typical person) as the cognitive component of *empathy*, the other component being the affective component. The cognitive component is defined as the drive to recognize another person's thoughts and feelings, and the affective component is defined as the drive to respond with an appropriate emotion to another person's thoughts and feelings. This led us to develop instruments such as the Empathy Quotient (EQ) showing that people with autism spectrum conditions have below average empathy. In other words, we needed a broader construct (empathy) to understand autism, which acknowledged the emotional responsivity dimension, not just the cognitive skill.

Fourth, for me, the sole focus on autism as a deficit (in empathy) was inadequate because it ignored the other side of autism: the areas of strength. Whilst autism and AS are disabilities, Uta Frith had argued convincingly that autism and AS involved cognitive *assets* as well as deficits and that autism should be viewed as a different cognitive style. She argued (perhaps paradoxically) that even the assets reflected another cognitive deficit that they termed 'weak central coherence' (or the inability to integrate featural or detailed information into global or gestalt represention). Frith had elegantly shown years earlier that children with autism had *excellent attention to detail* on measures like the Embedded Figures Test but saw this as a failure to use context or achieve 'strong central coherence', to see the bigger picture.

The weak central coherence (WCC) theory had value in highlighting how children with autism may suffer from information overload in being forever lost in the detail but unable to see the bigger picture. However, the WCC theory still painted autism as a disability, which to my mind overlooked that some people with autism (especially those with AS) *could* achieve mastery of a total system (such as their computer), a phenomenon that the WCC theory could not easily explain. For me, the assets in autism are not part of another disability but are signs of intact or even precocious development of a second psychological process, which I called 'systemizing'. This was the fifth major theoretical revision.

Systemizing is defined as the drive to analyse or build a system. Any system. This might be a mechanical system (such as a computer), a natural system (like the weather), an abstract system (like mathematics), or a collectible system (such as a library catalogue). Systems are defined as anything that follows rules or laws. Rules or laws are defined as unchanging, repeating, timeless patterns. I argued that individuals on the autistic spectrum (whether with classic autism or AS) could be characterized by below average empathy (E) alongside intact or even above average systemizing (S). This is a two-factor theory.

In my recent book *Zero Degrees of Empathy* (Penguin UK, 2011) I argue that hyper-systemizers such as people with autism have to 'step out of time' to identify such timeless, repeating patterns. Whilst the empathy deficit could account for the social features of autism, the intact or hyper-systemizing could account for the non-social features of autism. A surprising advantage of this E-S theory was that it could make sense of an otherwise puzzling set of 'symptoms' in autism, namely the strongly repetitive behaviour, narrow interests (obsessions) and the resistance to change/need for sameness that had been noted from the first description of autism by Leo Kanner in 1943.

This is because when you systemize you have to do what professional scientists are trained to do: focus on one detail/one variable at a time, whilst holding the rest of the system constant (or unchanging) to see what happens when you change one component (variable) in the system at a time. Without doing this an understanding of the system will be slow, potentially shallow, and incomplete. Children with autism seemed to need the world to remain unchanging (for example, putting all electrical switches in the house into their original positions, so that when one switch is changed, one can see the cause-effect rules that govern the system (which switch controls which light coming on).

Intact or even hyper-systemizing can lead not to deficits but to talents or precocious development in understanding 'how things work' and helps make sense of why a child who appears 'mindblind' (unaware of others' thoughts and feelings) could nevertheless figure out how to operate the family video recorder or computer, or who collects a class of object or information 'obsessively', or takes an object apart to disassemble it into its component parts and then reassemble it (one part at a time) to confirm the rules that govern the system. And it conceptualized the repetitive behaviour and excellent attention to detail seen in autism as being *in the service of* systemizing. The WCC theory saw these

behaviours as the signs of a disability. The E-S theory saw these behaviours as sign of cognitive strengths.

Earlier views of autism had seen the repetitive behaviour as 'purposeless' whilst the E-S theory now saw this as highly purposeful. Earlier educational practice had advocated that one should *discourage* the child's obsessions and repetitive behaviour whereas the E-S theory now saw these as a sign of a very different cognitive style, and that preventing a child from systemizing would be akin to preventing a child with a musical talent from spending time listening to music, or a preventing a child with congenital blindness from using their sense of touch. It led to the proposal that educators should *encourage* the child with autism to pursue their repetitive behaviour and narrow interests since these would lead to the child mastering one aspect of the environment at a time, however small. It could be a system like the names of every dinosaur, or the patterns in a railway timetable, or to predict which day of the week any calendar date might fall, so-called *savant* abilities.

Finally, as a sixth major revision, the E-S theory also had one other intriguing aspect: namely, that there are sex differences on average in each of the two domains. Females on average score higher on tests of empathy, and males on average score higher on tests of systemizing. This meant, psychometrically, that autism could represent an extreme of the typical male brain. This was explored in my book *The Essential Difference'* (Penguin, 2003). The E-S theory has given rise to the 'extreme male brain' theory of autism that is now being tested at the neural level as well as in terms of its possible link to levels of prenatal testosterone and sex-linked genes.

****

Empathy has two separable components: a cognitive component (the ability to recognize someone else's thoughts, intentions, and feelings) and an affective component (the drive to respond to someone else's thoughts and feelings with an appropriate emotion). Two different neurodevelopmental conditions involve empathy deficits: autism and psychopathy. So why do they not result in a similar outcome? If they both share low empathy, why do people with autism tend to avoid other people, struggle with relationships, and show a commitment to honesty and truth, whilst psychopaths often hurt other people, and manipulate and deceive others?

The answer to this riddle may lie in the two 'fractions' of empathy. The idea is that whilst people with autism have well-established difficulties in 'theory of mind' (the cognitive component of empathy) alongside intact affective empathy, psychopaths have intact theory of mind but impaired affective empathy. Put otherwise, people with autism have trouble keeping track of others' intentions, beliefs, knowledge, desires, and emotions, but still get upset when they hear of someone's suffering, whilst psychopaths find it easy to mindread others and do so to their own advantage, but don't *care* about others' thoughts and feelings.

So if in psychopaths cognitive empathy is intact whilst affective empathy is impaired, and if in autism the profile is precisely the opposite way around, could this explain the differences we observe in their behaviour? Certainly it is a parsimonious explanation for why 'zero degrees of empathy' can result in cruelty in psychopaths on the one hand, and social withdrawal and confusion in people with autism. People with autism often show high levels of morality – to the point of becoming whistle-blowers when they perceive others as breaking the rules – whilst psychopaths show high levels of amorality.

For some people it will come as a surprise – given the 'mindblindness' theory of autism – that many people with Asperger Syndrome show high levels of care for others. They look after their ageing parents, their pets (some even take in dozens of lost or injured animals), and many are devoted parents to their own children and show care towards the sick in their community. In addition, they give to or work for charities that provide care to those less fortunate than themselves. So the 'care' principle is often well developed in autism and Asperger Syndrome. Equally, many people with Asperger Syndrome become passionate lobbyists for social change towards greater social justice, campaigning and protesting and marshalling their political arguments for concepts such as equality, fairness and justice, feeling outrage for those who are unjustly interned, for example, and compassion for their plight.

And the same is true of their feelings of loyalty. They are often described as the most loyal of employees, recognizing that betrayal is immoral and the importance of sticking with your team, whether as a football supporter or as a member of a group. People with Asperger Syndrome also show keen attention to social hierarchy, not just their own position within it but a close scrutiny of those at the top, wanting their leaders to prove they deserve our respect by behaving consistently, honestly, and ethically.

For people with autism, their intact morality has a second cast iron platform: their strong systemizing. As mentioned earlier, people with autism are strong systemizers, typically becoming highly focused ('obsessed') with particular systems. Perhaps because of their difficulties with cognitive empathy, they rely even more on systemizing to understand the social world, wanting people to be consistent and to follow rules. These can include rules of morality. This drive to understand the system in all its exquisite detail, and in a black-and-white, binary fashion, is not just an advantage when it comes to figuring out gadgets, building Lego structures, and piecing together train timetables or the names of every dinosaur (just some of the 'obsessions' that develop in autism) but can also lead to a strong moral code[2].

[2] Some people with Asperger Syndrome sometimes argue they are not strong systemizers because they were awful at math: but the systemizing theory does not propose that people with autism or Asperger Syndrome should be good at understand *all* systems, since it is the nature of systemizing that one latches onto just one system at a time to understand it deeply. For some people it may be math, but for others it may be calenders, horse-riding, map-collecting, names of birds, or any other system.

****

We have come a long way from the first experimental demonstration of theory of mind deficits in autism, but there is still a long journey ahead to unravel the complexity of the autistic mind and brain.