

Combining the old and the new

Jon Brock looks at Bayesian and predictive coding accounts of autistic cognition

Bernard Rimland's classic text, *Infantile Autism* (1964), marked the beginning of research on autism as a disorder of cognition. Rimland's hypothesis – that autistic individuals have difficulty relating new and old experiences – has recently been updated in the form of Bayesian and predictive coding accounts of the condition. These new approaches have the potential to explain a wide range of symptoms associated with autism, linking differences in cognition to their underlying neurobiology. But as with all contemporary theories of autism, the challenge will be to address the huge variability that exists within the autism spectrum and the overlap with other supposedly distinct conditions.

question

How might researchers differentiate between alternative accounts of Bayesian and predictive coding?

How does autism research move forward if there are in fact many different autisms all lumped together?

resources

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36, 181–204.

'How did Nate Silver predict the US election?' www.theguardian.com/science/grrlscientist/2012/nov/08/nate-silver-predict-us-election

Mark Rimland was not like other babies. As his father, Bernard, later recalled, 'Mark was a screaming, implacable infant who resisted being cuddled and struggled against being picked up'. The paediatricians were baffled, and it was only after his mother, Gloria, remembered reading about Leo Kanner's description of 'autism' in a psychology textbook that a diagnosis was eventually made.

Today, autism is considered to be a neurodevelopmental disorder arising from the complex interplay of genetics and environmental factors during pre- and early post-natal development. However, in 1958, when Mark was first diagnosed, it was generally regarded as the child's reaction to a lack of maternal affection – the infamous 'refrigerator mother' hypothesis. Bernard knew that this was nonsense and set out to determine the real cause of his son's difficulties. His literature search led in 1964 to the publication of a book, *Infantile Autism*, that was to become a landmark in autism research. Not only did he conclusively debunk the refrigerator mother hypothesis, Rimland also set out his own revolutionary theory. Autism, he argued, was a cognitive dysfunction. The 'diversity of symptoms and manifestations' could be traced to 'a single critical disability: The child with early infantile autism is grossly impaired in a function basic to all cognition: the ability to relate new stimuli to remembered experience... The child is thus virtually divested of the means for deriving meaning from his experience...

He cannot integrate his sensations into a comprehensible whole.'

In his subsequent work, Rimland moved away from cognitive theorising. He promoted some questionable biomedical interventions and was a strong advocate of the now discredited notion that autism is caused by traces of mercury in vaccines. However, the ideas laid out in *Infantile Autism* inspired a generation of researchers to investigate autistic cognition.

Half a century later and eight years after his death, Rimland's original ideas are back in vogue. In a 2012 paper, Liz Pellicano and David Burr proposed that autistic cognition could be understood in terms of Bayesian statistics – a mathematical framework for combining new and old information. Autism, they suggested, is characterised by a reduced influence of prior knowledge. In response to Pellicano and Burr, other researchers have argued that autistic cognition might be more usefully characterised in terms of a related theoretical framework, predictive coding and, more specifically, how the brain deals with the inevitable discrepancies between reality and expectations.

Central coherence

Rimland's ideas about autistic cognition were highly speculative, based on observation and intuition alone. However, UK psychologists Beate Hermelin and Neil O'Connor were already beginning to conduct experiments that spoke directly to his hypothesis (Hermelin & O'Connor, 1967).

In one study they gave children a series of words to repeat out loud. As expected, non-autistic children made fewer errors if the words made up meaningful sentences. 'The fish swims in the pond' was easier to recall than the random word string 'By is go tree stroke lets'. Children with autism, however, failed to show this effect, suggesting that they were repeating the individual words without considering the meaning of the

references

- Adams, R.A., Stephan, K.E., Brown, H.R. et al. (2013). The computational anatomy of psychosis. *Frontiers in Psychiatry*, 4, 47.
- Brock, J., Norbury, C., Einav, S. & Nation, K. (2008). Do individuals with autism process words in context? *Cognition*, 108, 896–904.
- De Lacy, N. & King, B.H. (2013). Revisiting the relationship between autism and schizophrenia. *Annual Review of Clinical Psychology*, 9, 555–587.
- Fletcher, P.C. & Frith, C.D. (2009). Perceiving is believing: A Bayesian approach to explaining the positive symptoms of schizophrenia. *Nature Reviews Neuroscience*, 10, 48–58.
- Frith, U. (1989). *Autism: Explaining the enigma*. Oxford: Blackwell.
- Frith, U. & Snowling, M. (1983). Reading for meaning and reading for sound in autistic and dyslexic children. *Journal of Developmental Psychology*, 1, 329–342.
- Friston, K.J., Lawson, R. & Frith, C.D. (2013). On hyperpriors and hypopriors: Comment on Pellicano and Burr. *Trends in Cognitive Sciences*, 17, 1.
- Geschwind, D.H. & Levitt, P. (2007). Autism spectrum disorders: Developmental disconnection syndromes. *Current Opinion in Neurobiology*, 17, 103–111.
- Happé, F., Ronald, A. & Plomin, R. (2006). Time to give up on a single explanation for autism. *Nature Neuroscience*, 9, 1218–1220.
- Hermelin, N. & O'Connor, N. (1970). Psychological experiments with autistic children. Oxford: Pergamon Press.
- Kahan, J. & Foltynie, T. (2013).

sentence, exactly as Rimland might have predicted.

Hermelin and O'Connor were unaware at the time of Rimland's book, but it inspired their PhD student Uta Frith, to continue this line of work. In one ingenious experiment (Frith & Snowling, 1983) she and Maggie Snowling gave children 10 sentences to read out loud. Each contained a homograph – a word such as *tear* with a dual meaning. By noting how the children spoke the homographs, Frith and Snowling could determine which meaning had been assigned – and whether it made sense in the broader context of the sentence. Children with autism performed poorly on this task. They would pronounce *tear* the same, regardless of whether the sentence suggested crying or ripping, indicating that they had failed to take the earlier sentence context into account. These and other findings led Frith to propose her 'weak central coherence' account of autism. Autistic cognition, she argued,

is characterised by a focus on detail and a failure to 'weave together' information to extract meaning (Frith, 1989).

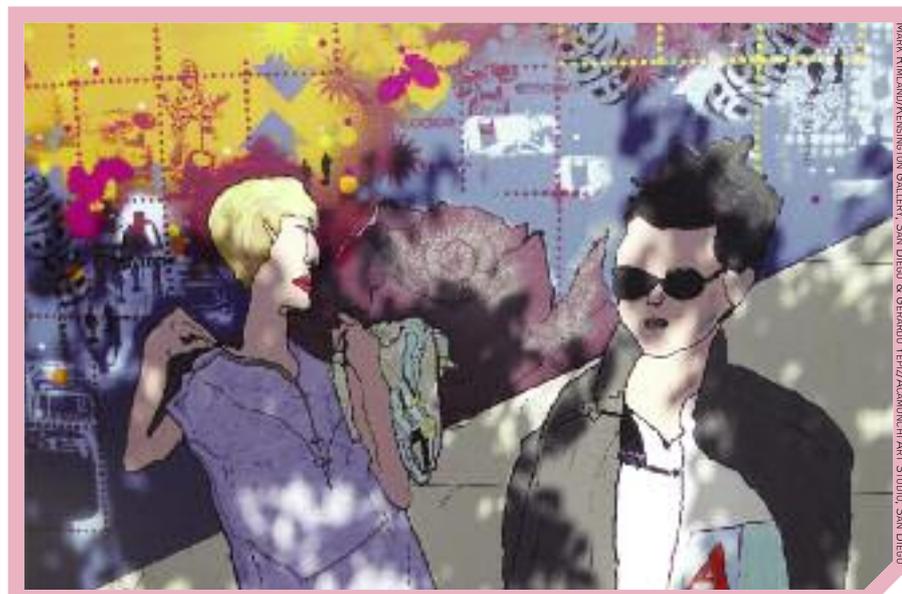
Central coherence was, in essence, a fleshing out of Rimland's earlier proposal. It provided a way of understanding some of the difficulties facing people with autism, including problems with language and social understanding where contextual nuances are often critical. Importantly, however, the account also emphasised the strengths of autistic individuals. Being able to ignore the bigger picture and focus on the details might in some circumstances be advantageous. Frith and her colleague, Amita Shah had found that children with autism performed relatively well on certain visuo-spatial tasks that required attention to detail. For example, on the embedded figures test in which they had to locate geometric shapes hidden in a larger more complex picture, autistic children outperformed non-autistic children of similar overall cognitive ability (Shah & Frith, 1983).

The central coherence account proved highly influential. Over the years, it has inspired hundreds of studies and has succeeded in focusing attention on the non-social, non-diagnostic features of autism. However, the account remains frustratingly vague. The term 'central coherence' is not one that is recognised in mainstream cognitive psychology and it's unclear what cognitive mechanism might actually be involved. It's more a description of the kind of explanation researchers are looking for than an explanation in and of itself. Pellicano and Burr's Bayesian account is an attempt to go beyond central coherence and pin it down to something more precise, testable and, ultimately, falsifiable.

The Bayesian perspective

The basic premise underlying Bayesian statistics is that information is inherently unreliable, so a better estimate of reality comes from combining new information with prior knowledge. Perhaps the best-known example of Bayesian statistics in action came from the 2012 US presidential election campaign, when the *New York Times* blogger Nate Silver was able to correctly predict the results in all 50 states. Rather than just relying on the latest opinion poll, Silver, developed a mathematical model of voting intentions. He then used Bayesian theory to incrementally adjust the model in the light of each new opinion poll. The weight put on a new poll depended on its margin of error – how much it fluctuated from poll to poll and how well it had predicted the outcome of previous elections.

The Bayesian principles underlying Silver's model have also been applied to cognitive theories of perception. In the same way that an opinion poll has a margin of error, incoming sensory information is inherently ambiguous and unreliable. For instance, the light hitting the back of the eye could come from an infinite number of different arrangements of objects in the real world. However, the



Bernard Rimland's quest to understand his son Mark – who created this artwork – led to the publication of *Infantile Autism* in 1964

Understanding DCM: Ten simple rules for the clinician. *Neuroimage*, 83, 542–549.

Lawson, R.P., Rees, G. & Friston, K.J. (2014). An aberrant precision account of autism. *Frontiers in Human Neuroscience*, 8, 302.

Mitchell, P., Mottiron, L., Soulières, I. & Ropar, D. (2010). Susceptibility to the Shepard illusion in participants with autism: Reduced top-down

influences within perception? *Autism Research*, 3, 113–119.

Pellicano, E. & Burr, D. (2012). When the world becomes 'too real': A Bayesian explanation of autistic perception. *Trends in Cognitive Sciences*, 16, 504–510.

Rimland, B. (1964). *Infantile autism: The syndrome and its implications for a neural theory of behavior*. Chicago: Appleton-Century-Crofts.

Shah, A. & Frith, U. (1983). An islet of ability in autistic children. *Journal of Child Psychology and Psychiatry*, 24, 613–620.

Teufel, C., Subramaniam, N. & Fletcher, P.C. (2013). The role of priors in Bayesian models of perception. *Frontiers in Computational Neuroscience*, 7, 25.

van Boxtel, J.J. & Lu, H. (2013). A predictive coding perspective on

autism spectrum disorders. *Frontiers in Psychology*, 4, 19.

van de Cruys, S., de-Wit, L., Evers, K. et al. (2013). Weak priors versus overfitting of predictions in autism: Reply to Pellicano and Burr [TICS, 2012]. *i-Perception* 4, 95–97.

Wass, S. (2011). Distortions and disconnections: Disrupted brain connectivity in autism. *Brain and Cognition*, 75, 18–28.

autism

problem is constrained by our experience of what objects are likely to be out there, and we are rarely conscious of the potential for alternative interpretations.

Pellicano and Burr's proposal is not that people with autism ignore prior knowledge altogether, but rather that their priors are broader – and so perception is less constrained by past experience. In terms of the US election analogy, this is akin to placing less weight on previous opinion polls and thus being extra-sensitive to the most recent opinion poll. The result, Pellicano and Burr suggest, is that the world is somehow 'more real' for people with autism. There is, however, some poetic licence here. Under most circumstances, the prior information is useful and so improves the accuracy of perception, bringing us *closer* to reality (see Teufel et al., 2013).

A notable exception is the case of visual illusions. In the Shepard illusion, for example, the dimensions of the two table tops are identical (see Figure 1). However, our experience of rectangular surfaces in the real world leads us to perceive the table on the left as being long and thin, while the one on the right appears short and fat. Prior knowledge helps us to make a judgement about the three-dimensional object being represented, but it hinders judgements about the actual two-dimensional image. Consistent with the Bayesian account, Peter Mitchell and colleagues found that people with autism were less susceptible to this effect than non-autistic people, suggesting in turn that they are less affected by what Mitchell and colleagues term 'the curse of knowledge' (Mitchell et al., 2010).

Pellicano and Burr's Bayesian account is essentially a mathematical formalisation of Mitchell and colleagues' proposal – and of Rimland's original idea. But this formalisation is important. It finally grounds research on autistic cognition in mainstream cognitive psychology research. It also generates novel and testable predictions. In any situation where performance can be understood in terms of the combining of current and prior information, prior knowledge should be underweighted in individuals with autism. While the current emphasis is on visual perception, the Bayesian framework could in principle be applied to a wide range of symptoms associated with autism.

Predictive coding

Rimland's initial focus in *Infantile Autism* was on cognition, but he was conscious that differences in information processing

must ultimately arise from differences in brain function. 'In some very real way', he suggested, 'memories, thoughts, and ideas are somehow locked into separate compartments of the autistic child's brain'. This same intuition underpins the more recent 'underconnectivity' hypothesis (see Wass, 2011), according to which, the symptoms of autism arise due to a lack of communication between different parts of the brain. There is now growing evidence for atypical brain connectivity in autism, but the relationship between brain activity and cognition in autism remains opaque.

In a commentary on Pellicano and Burr's paper, Karl Friston and colleagues suggested that the Bayesian account of autism could be usefully reframed in terms of the predictive coding framework (Friston et al., 2013; see also van Boxtel & Lu, 2013). Like the Bayesian approach, this concerns the way in which old and new information are combined, but it provides a more explicit link to neural mechanisms. Predictive coding starts with the simple idea that the main purpose of the nervous system is to try to anticipate what will happen next. Put another way, the organism tries to minimise its prediction errors – the difference between what it predicted would happen and what actually transpired. To some extent, prediction errors are inevitable. A key feature of predictive coding, therefore, is that the organism encodes the precision of its predictions – that is, how confident it is in their accuracy. If precision is low, then the organism takes even large prediction errors in its stride. But if precision is high and strong predictions are violated, this is cause to sit up, take notice – and learn from the experience, allowing better predictions to be made in future.

Friston and colleagues hypothesise that precision is reduced in autism: individuals with autism make less strong predictions and so their perception is dominated by the noisy, ambiguous sensory information. However, in a further commentary, Sander van de Cruys

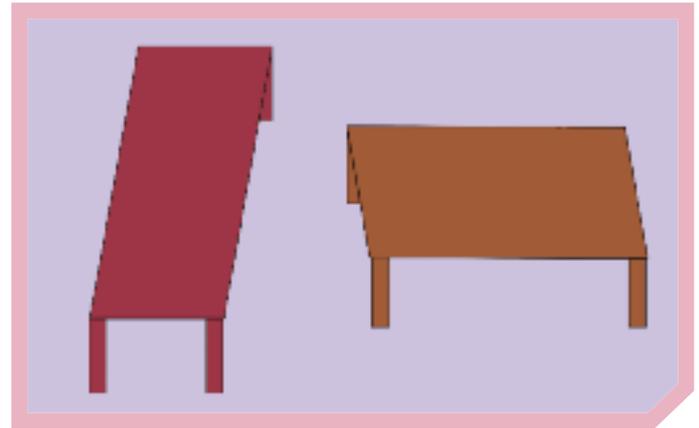


Figure 1: The Shepard illusion – knowledge of three-dimensional objects makes the two table surfaces appear to have different proportions, despite having the same two-dimensional shape

and colleagues suggest that autism could in fact be characterised by *over*-precise prediction errors (van de Cruys et al., 2013). Somewhat counterintuitively, this could have very similar effects to reduced precision because even small deviations from predictions are treated as significant prediction errors.

Within the predictive coding framework the brain is conceived as a hierarchy, with information cascading up and down the different levels. Brain Region A will make predictions about activity in Brain Region B, and the prediction error will then be sent back to Brain Region A. Region B will interact with Region C in a similar way, and so on down the chain. Differences in predictive coding thus correspond

"Autism is complicated and messy... There are no clear boundaries"

directly to the way in which brain regions at different levels of the hierarchy interact. In principle at least, it should be possible to test

hypotheses about atypical connectivity by applying dynamic causal modelling to neuroimaging data from individuals with autism. This involves generating different models of how brain regions interact during a particular task and then determining which model comes closest to predicting the neuroimaging data recorded (see Kahan & Foltynie, 2013).

Predictive coding also offers insights into how changes in brain connectivity and cognition might arise from differences at the synapse (Lawson et al., 2014). The precision of predictions is thought to involve NMDA receptors, which act like an amplifier in a hi-fi system, determining the size of a neuron's response to prediction errors. NMDA receptors are part of the glutamate neurotransmitter system, which in turn interacts with other neurotransmitters, including serotonin, dopamine and

GABA, as well as the hormone oxytocin, all of which have all been implicated in autism. There may, therefore, be a number of different but converging routes to atypical predictive coding and connectivity.

The schizophrenia connection

While still in their early days, the Bayesian and predictive coding accounts of autism have considerable potential for understanding autistic cognition, its underlying neurobiology and, ultimately, its genetics. But like all theories of autism, they face a number of challenges. Prominent amongst these is to specify exactly how autism differs from other conditions. In particular, very similar accounts have been proposed to explain schizophrenia (see Adams et al., 2013).

Historically, autism was considered to be an early-onset form of schizophrenia, and there are certainly overlaps between the two conditions (see De Lacy & King, 2013). Both are associated with impaired social reasoning and executive skills, as well as a reduced sensitivity to context (although the term 'central coherence' is not used in the schizophrenia literature). Both have been attributed to reduced or atypical brain connectivity, and certain genetic variations appear to convey risk for both conditions. There are, however, some clear differences. As Rimland remarked in *Infantile Autism*, 'the writer is reminded of the story of the two men who were indistinguishable in appearance except that the tall thin one had a red beard and only one leg' (p.68).

In particular, schizophrenia is associated with hallucinations and delusions in schizophrenia, which are not generally present in autism. Paul Fletcher and Chris Frith (2009) have argued that these 'positive' symptoms can be understood as the failure of predictive coding. Ordinarily, our internally generated thoughts and actions have highly predictable consequences – because we have made them ourselves. However, if prediction breaks down, then our own actions may feel like we didn't initiate them. Likewise, our inner voice may seem like it is someone else's. The question then is, if people with autism also have issues with predictive coding, why don't they too experience hallucinations and delusions?

One possibility is that autism and schizophrenia are both disorders of predictive coding, but the exact nature of the impairment is different. However, it is also important to consider the different developmental trajectories. In schizophrenia, atypical predictive coding

is occurring in the context of relatively well-developed adult internal model of the world. In autism the same atypicality present from very early in life (and possibly before birth) would entail that the individual's model of the world is atypical from the start.

Studies directly comparing autism and schizophrenia may prove particularly useful in refining accounts of both disorders. However, there is an underlying assumption here that autism is a distinct condition that a person either does or does not have; and that it makes sense, therefore, to think of a theory of *autism* – an explanation that applies to all or at least the majority of people with autism. This view, however, is one that is increasingly under question.

One autism or many?

As an appendix to *Infantile Autism*, Rimland provided an 80-item symptom checklist to assist in the diagnosis of autism. Total scores over a certain threshold indicated that the child was autistic. Although many different combinations of symptoms could lead to an autism diagnosis, Rimland's assumption was that the variation reflected different manifestations of the same underlying condition. Modern diagnostic checklists work in much the same way (and carry the same assumption), but the concept of autism is far broader than in the 1960s, and the heterogeneity within autism is accordingly even greater. Two individuals may both meet diagnostic criteria for an autism spectrum disorder but present with very few symptoms in common. Recognition of this heterogeneity problem is growing. In an influential 2007 paper, Daniel Geschwind and Pat Levitt argued that we should think of 'the autisms' (plural) as a collection of disorders that share some superficial similarities. The central challenge for autism researchers, then, is to tease apart the various autisms in order to identify the most appropriate interventions and support required by any individual. One possibility, therefore, is that the Bayesian/predictive coding account only applies to a subset of the many autisms. In other words, there will be some autistic people who show atypical priors or precision across a wide range of situations. Other individuals will be similar to non-autistic people in this regard – and would, therefore, require an alternative explanation for their symptoms.

A different and arguably more radical viewpoint is that we should stop thinking about syndromes like autism altogether

and instead focus research efforts on specific symptoms (Happé et al., 2006). Rather than considering atypical priors or precision to be a whole-brain phenomenon affecting all of cognition equally, it might make more sense to think of atypicalities in relation to specific brain systems and cognitive processes. Our own work on language comprehension tends to support this view (Brock et al., 2008). A strong prediction of the Bayesian/predictive coding account is that children with autism should make less use of contextual information when identifying spoken words. By testing autistic children with a wide range of language skills, as well non-autistic children with and without language impairment, we were able to show that context sensitivity was in fact related to a child's degree of language impairment irrespective of whether they had an autism diagnosis.

What next?

The issues that face the Bayesian and predictive coding accounts of autism are by no means unique. Autism is complicated and messy. Almost every aspect of cognition can be affected, but the pattern of strengths and difficulties varies hugely across the autistic population. Even so-called 'core' symptoms manifest differently in different individuals and at different stages of development. There are no clear boundaries. Autism overlaps with a wide range of other supposedly distinct conditions and appears itself to have multiple different causes.

Despite his remarkable perceptiveness, it is now apparent that Rimland wildly underestimated the complexity of the challenge he faced. Indeed, it seems unlikely that there will ever be a straightforward answer to the autism puzzle. Seen in this light, the Bayesian and predictive coding accounts of autism are clearly over-simplistic, but they do at least provide a framework for testing hypotheses and guiding research. As Rimland's successors, we may at last be edging closer to an understanding of how and why autistic individuals think differently to those of us without autism.



Jon Brock
ARC Centre of Excellence in
Cognition and its Disorders,
Macquarie University,
Australia
jon.brock@mq.edu.au