

# Understanding Autistic Individuals

Inês Hipólito, Daniel Hutto and Nick Chown

‘I will teach you differences!’

*King Lear*, Act I, Scene 4

## Introduction

What autistic people tend to think about, the way they think about things, and the ways they interact with others is atypical when compared to the population at large. The cognitive diversity of autistic people, and its many variations, is well-documented and much discussed.

This chapter has the potential to contribute to the neurodiversity movement by providing philosophically motivated reasons for thinking differently about the cognitive styles of autistic individuals. In particular, it challenges the prevalent mindreading characterisation of everyday social cognition that promotes the view that autism is an underlying condition that is best explained in terms of deficiencies in inferential capacities to form and test hypotheses.

In a recent opinion piece on the neurodiversity movement published in *Scientific American* on 30 April 2019, Baron-Cohen (2019) reminds of what he describes as the ‘huge heterogeneity’ among those people who fall within the autistic spectrum. Some autistic people have no functional language and severe developmental delay (both of which I would view as disorders), others have milder learning difficulties, while yet others have average or excellent language skills and average or even high IQ. What all individuals on the autism spectrum share in common are social communication difficulties (both are disabilities), difficulties adjusting to unexpected change (another disability), a love of repetition or ‘need for sameness,’ unusually narrow interests, and sensory hyper- and hypo-sensitivities (all examples of difference). Autism can also be associated with cognitive strengths and even talents, notably in attention to and memory for detail, and a strong drive to detect patterns (all

of these are differences). How these are manifested is likely to be strongly influenced by language and IQ. Undeniably, there is enormous variability in the full spectrum of cognitive styles exhibited by autistic individuals. Every autistic individual has their own distinct cognitive style, just as every non-autistic individual does.

Beyond merely acknowledging and carefully cataloguing the heterogeneous cognitive diversity of autistic individuals, those attracted to the medical model have made persistent attempts to classify these cognitive styles, treating them as a part of a constellation of symptoms that are expressive of an underlying condition – a condition that is typically denoted by the labels ‘autism’ or ‘autistic spectrum disorder’.

Over the decades there have been many attempts to understand the true character and ultimate basis of the totality of symptoms generated by the supposed underlying autistic condition from which autistic people allegedly suffer. Yet the current received view is that the hunt for a single condition that explains and accounts for the full set of autistic symptoms in a unified manner is a snark hunt. It is now widely accepted that ‘no single aetiology can account for all differences in presentation’ (Ure et al., 2018, p. 1068). Naturally, the idea that autism is comprised of a cluster of underlying conditions – not a single condition – lends itself naturally to classifying the heterogeneity of autistic individuals in terms of various species and subtypes of autism.

The aim of empirical research into these assumed underlying conditions that make up autism is directed at identifying specific biological markers for distinct aspects of autistic disorder – aspects which, by the lights of those who buy into the medical model wholesale, are understood to be neurodevelopmental in nature.

In line with these developments, some researchers have set their explanatory sights more modestly. They zoom in to focus only on what underpins the atypical patterns of social cognition exhibited by autistic individuals – namely, their distinct style interacting with others and their limitations in fluidly understanding what motivates actions. For them, understanding what explains the social cognitive aspects of autism alone would still constitute a major advance. This would surely be so if, as Baron-Cohen (2019) maintains, social communicative difficulties are found across the autistic spectrum and ‘aspects of social cognition reflect areas of disability in autism, and are often the reason for why they seek and receive a diagnosis’.

The Diagnostic and Statistical Manual of Mental Disorders (DSM-5) (APA, 2013) concurs, taking the atypical styles of social cognition of autistic individuals to be among the most diagnostically important criteria and defining features of the condition. The DSM-5 tells us that Autism Spectrum Disorder is characterised by persistent

deficits in social-emotional reciprocity, ranging, for example, from abnormal social approach and failure of normal back-and-forth conversation ... to failure to initiate or respond to social interactions; Deficits in nonverbal communicative behaviours used for social interaction; Deficits in developing, maintaining, and understand relationships. (American Psychiatric Association, 2013, p. 50)

The goal of accounting for the distinctive patterns of social cognition of autistic individuals has launched a thousand explanatory ships, all of which set forth to discover what underpins at least the social cognitive aspects of the autistic mind. This chapter raises doubts about a specific class of explanations of the social cognitive styles of autistic individuals – Theory Theory, or ‘TT’, explanations.<sup>1</sup> TT comes in many forms – but what is common to all versions is the proposal to explain the basis of our social cognition by appeal to machinery of mind that makes use of theories of some sort that are understood to involve contentful inferential processes. TT is a worthwhile target since it is the dominant and most popular framework for answering the Explanatory Challenge of what, supposedly, underpins everyday social cognition and what goes systematically wrong with social cognition for autistic people.

Social cognition can be neutrally understood as denoting ‘our ability to understand and interact with others’ (Spaulding, 2010, p. 120). As such, it is important to remind ourselves that everyday social cognition can be depicted in various ways. Appropriately characterising the nature of social cognition poses a special kind of philosophical challenge. Let us call this the Characterisation Challenge.

Ultimately, we will argue that the various explanatory proposals of TT only look promising so long as a certain kind of answer is given to the Characterisation Challenge. That is to say, TT proposals look promising if we accept the dominant characterisation of social cognition – the standard mindreading story – which holds that ‘in order to understand and successfully interact with other agents, neurotypical adult humans attribute mental states to other agents in order to explain and predict their behavior’ (Spaulding, 2018, p. 7).

Depicting social cognition in such spectatorial terms licenses the received and longstanding view in the field that, ‘many people with autism spectrum conditions have a specific impairment in

---

<sup>1</sup> In this chapter we focus exclusively on TT proposals: however, our conclusions generalise. They apply with equal force to any Simulation Theory, or ST, proposal which assumes that neutrally based mindreading involves some kind of inference. ST proposals hold that inferences about other minds are achieved by co-opting or reusing planning or practical reasoning systems – rather than forming theories about goal planning, as is the wont of TT proposals (see, e.g. Gordon, forthcoming). However, on our analysis, the difference between ST and TT about the character of the central processes involved in modelling other minds – while technically interesting – is unimportant with respect to the larger concerns we raise in this chapter. This is because both classes of theory, TT and ST, characterise everyday social cognition in mindreading terms. Our objections apply to any explanation that takes the mindreading characterisation of everyday social cognition for granted. Hence, our analysis applies with equal strength to any mentalising proposal – whether ST or T – which assumes that it is the job of the brain to infer the goals or reasons (or otherwise ‘the causes’) that lie behind another agent’s observed behaviour.

mindreading' (Heyes, 2018, p. 149). As we shall reveal, it is no accident that the credibility of TT explanations of the alleged social cognitive deficits of autistic individuals depend on thinking of social cognitive styles of autistic individuals in terms of impoverished mindreading – as impaired attempts to get at the contents of other minds, that occur whenever autistic people attempt to engage in everyday social cognition.

This chapter argues that the fate of *any* proposed TT answer to the Explanatory Challenge stands or falls with the appropriateness of giving a Mindreading answer to the Characterisation Challenge. The first two sections provide details of the current state of the art with respect to TT proposals about how to understand and explain autistic social cognition. Section 2 focuses on old school modularist TT proposals, noting their theoretical and explanatory limitations. Section 3 examines new school predictive processing TT proposals, highlighting what has made them appear more theoretically and explanatory promising to many researchers.

Section 4 then provides a diagnosis of why we should reject any kind of TT proposal about the supposed social cognitive deficits of autistic individuals. Our concerns are about any TT proposal that takes seriously the core assumption that the primary and pervasive way that we engage and connect with others is by means of theorising. We raise objections to TT proposals as a class by providing reasons for thinking that a Mindreading answer to the Characterisation Challenge obscures the true nature of everyday social cognition. We conclude that the crucial step of characterising the social cognitive styles and tendencies of autistic individuals as some kind of Mindreading deficit is a mistake. The paper closes by encouraging the adoption of alternative, non-Mindreading ways of understanding the social cognitive styles of autistic individuals. We contend that an enactivist alternative can offer an antidote to thinking of the diverse social cognitive styles of autistic individuals in terms of underlying sub-personal cognitive deficits rather than in terms of the cognitive differences of whole persons.

## **2.1. Old School, Mental Module TT**

A longstanding, high profile TT hypothesis about what best explains the distinctive social cognitive patterns of autistic individuals holds that these stem from those individuals having a faulty or poorly functioning Theory of Mind or, ToM, module.

In general Theory Theory, or TT, proposals about social cognition assume that when we understand minds in daily life, we use the same sorts of tools that we use to understand other non-mental phenomena. That is to say, we use the same sort of tools we use everywhere in the sciences – namely, theories that aim to tell us about the unobservable, hidden causal structure of the world.

A ToM is a very particular kind of theory; it is assumed to have a distinctive sort of content. A ToM is made up of mental-state concepts that feature in theoretical postulates that comprise the core general principles of a theory of everyday psychology. The content of the ToM that normally developing humans use, so the story goes, is what enables most of us to navigate our everyday social world fluidly and with ease. We succeed in understanding others if we manage to infer their mental states correctly by applying a ToM, thereby bringing the laws of everyday psychology to bear on particular cases.

As such, ToM variants of TT hold that, for most of us, the heavy lifting in everyday social cognition is done by our acquaintance with and use of laws of governing everyday social cognition. Modularists take this idea a step further. They hold that ToM laws are housed in a special cognitive mechanism – a ToM Module, or ToMM. A ToMM is an architecturally distinct mental module – one that is solely dedicated to the special work of enabling us to predict and explain the actions of others by accurately attributing mental states contents to them. Believers in classic ToMMs assume that ‘the mind contains a single mental faculty charged with attributing mental states (whether to oneself or to others)’ (Carruthers 2011, p. 1).

Those who posit ToMMs assume that however such modules are acquired, they are the means by which everyday social cognition is normally conducted by our species. Nativist accounts of ToM assume that it is a biological device that comes built-in as standard for all normally developing members of our species (Fodor 1983). Others hold that ToMMs are acquired during ontogeny (Karmiloff-Smith, Klima, Bellugi, Grant, & Baron-Cohen, 1995). Some, such as Scholl and Leslie (1999) even propose that ‘normal children seem to develop the same ToM at roughly the same time’ (p. 138).

In their heyday, modularist theories of mind aided and abetted the idea that impaired mindreading abilities, rooted in damaged or atypically functioning neurocognitive machinery of a ToMM, were responsible for the social cognitive patterns displayed by autistic individuals.

We see these ideas brought together in the work of Baron-Cohen. According to its original formulation, Baron-Cohen’s (1995) mindblindness hypothesis proposed that ‘children and adults with the biological condition of autism suffer, to varying degrees, from mindblindness’ (1995, p. 5; Brewer, Young, & Barnett, 2017; Gillean, Xie, & Strelchuk, 2017; Livingston and Colvert 2019). By Baron-Cohen’s (2000) lights, difficulty in social cognition, understood as a mindreading impairment, is the ‘core and possibly universal abnormality of autistic individuals’ (Baron-Cohen, 2000, p. 3).

Summarising work in this vein, Brüne (2005) reports a range of findings that suggest to many that we find the fingerprint of an ‘impaired ToM in a variety of neuropsychiatric disorders’ (Brüne, 2005, p. 21). Concomitantly, modularists hypothesised that these various patterns of autistic

mindblindness are caused and explained by problems with the ToMMs of people in these populations. Thus, a standard proposal in the field, even today, is that ‘the functional or structural disruption of the neural mechanisms underlying ToM may give rise to various types of psychopathology’ (Brüne, 2005, p. 21).

The faulty ToMM proposal contends that autistic individuals are prevented from attributing contents to other minds accurately or, in the most extreme cases, doing so at all. What makes the faulty ToMM hypothesis about the social cognitive tendencies of autistic individuals attractive to many is that

to see a person with autism, we are told, is to see what happens to a human being when the ability to mentalize ... is switched off ... On the surface, this is neatly specific ... The ‘theory of mind’ explanation seems to fit the facts. (Belmonte, 2009, p. 121)

As traditionally understood, a defining feature of modules is that they are informational encapsulated, both vertically and horizontally (Matthews, 2019; Quilty-Dunn, 2019; Raftopoulos, 2019). Information contained in each module is vertically encapsulated from other modules and horizontally encapsulated from the information available in the cognitive system.

Modules have limited cognitive interests and concerns. In getting their epistemic work done, they operate on a strictly need-to-know basis, and they – apparently – don’t need to know much. It is assumed that modules work better and faster by restricting their concerns to specialised dealings with only certain topics. It is because of their informationally encapsulated domain-specificity that they are not informed and updated by all the contentful knowledge that might possibly be communicated to them. The limited communicative repertoire of modules is the peculiar characteristic that secures their status as mental modules.

Accordingly, each type of mental module is assumed to be restricted in the subject matter of its concern. Modules are domain-specific in a sense that only a circumscribed class of inputs will activate them. It is this feature of modules that makes them dissociable such that they can be selectively impaired, damaged, or disabled without effecting the operation of other systems.

Putting all of this together, ToMM theory holds that the main work of predicting and explaining the behaviour of others by assigning mental states is done in isolation from the operation of other cognitive systems. The essential character of ToMM is that it provides specialised theoretical knowledge of its particular domain and it can function by and large independently of other modules and central cognitive processes.

Since ToMMs are dissociable components, both from other modular devices and from central cognition, impairment of a ToMM will not directly impair the functioning of other cognitive mechanisms. In conclusion, the awkward and failed social interactions of autistic individuals are thus put down to the alleged fact that they lack a ToMM or are unable to wield their ToMs well in practice so as to accurately represent the mental states of others.

It is not enough to have a working ToMM, to mindread successfully. Successful mindreading also requires being able to apply one's ToM sensitively in ways that address the particularities of specific cases. That requires adjusting for relevant differences between cases by making allowance for a range of variables including a great deal of background belief and knowledge about what the other person knows and thinks, how they are likely to react, what is the most likely way someone would react in such cases, and so on and on. In short, believers in ToMMs are obliged to explain how the core ToM we allegedly use is applied sensitively in situ (for a more detailed discussion of this point see Hutto, 2008, Chap. 1).

It is wholly unclear how having a general ToM machinery that works in isolation from relevant background knowledge could possibly enable us to cope with ad hoc details and idiosyncratic attitudes we need to cope with in each new situation. Without supplement, ToMMs would be at a loss precisely when it comes to explaining how we deal with details; it is uncomfortably quiet on how we fluently come to understand particular people and in particular circumstances. Yet, as Maibom (2009) observes 'folk psychological knowledge is knowledge of the (empirical) world only if it is combined with knowledge of how to apply it' (p. 361).

An adequate explanation of how ToMMs could underpin everyday social cognition is required if a ToMM story is to be believed about how social agents grasp these social idiosyncratic details so as to understand and come to make sense of one another.

How could a ToMM operate successfully in isolation from the background knowledge and belief that appears to be needed to inform and direct its use in specific cases? Pivotaly, the supporting knowledge needed for applying a ToM sensitively cannot be supplied by a ToM and, indeed, given the contextual nature of the supporting knowledge, it isn't possible to specify it in advance at all. This reminds us that the business of socially engaging with and coming to understand others is deeply context sensitive. There are simply no algorithms with the right properties that would allow us to anticipate the relevant possibilities.

## 2.2. New school, Bayesian Brain TT

It is neither clear how, nor if, defenders of the classic ToMM can respond adequately to the serious theoretical and explanatory concerns raised in the previous section. In this light, a tempting way to go might be to seek to change theoretical horses by jumping on the Bayesian brain bandwagon.

The Bayesian Brain Hypothesis, or BBH, contends that cognition is, through and through, concerned with making and improving on its predictions about the causal structure of the world. Tirelessly and proactively, our brains are forever trying to look ahead in order to ensure that we have an adequate practical grip on the world in the here and now. On this view, our brains do not sit back and receive information from the world, form truth evaluable representations of it, and only then work out and implement action plans. Instead, the BBH holds that the basic work of brains is to get ahead of the curve by making the best possible predictions, in advance, about what the world is likely to throw at us. This is all part of the bigger job of cognition which, in all its varieties, is to try to get a sense of what is going on behind the sensory scenes by advancing, testing, and refining inferences to the best explanation on multiple spatial and temporal scales. Through this continual and dynamic process, so the BBH claims, we get a better and better fix on the true causal structure of the world.

Although the BBH made its name for its attempts to better account for the nature of perception–action cycles there has been a move to apply it to understand a much wider range of cognitive phenomena, including social cognition – especially when the latter is construed in Mindreading terms (see, e.g. Pezzulo, 2017).

The BBH is a full-blooded type of TT. Like its ToMM cousin it seeks to give an account of the Mindreading processes that it assumes lie at the heart of everyday social cognition. As noted above, the idea is that – at its core – cognitive activity is always about making inferences concerning the hidden causes of sensory phenomena. Advocates of the BBH hold that the processes used in Mindreading are the same basic kind used elsewhere in every variety of cognition – including acts of basic perception. The only difference in the case of social cognition is the target of the activity.

Thus, just as non-social objects in our environment are causes of our visual input, the mental states of other people are a part of the physical structure of the world that produces the stream of sensory impressions that our brains receive. In this view, mentalising occurs implicitly and shares a fundamental similarity with the representation of non-social objects: each is a natural result of the brain's endeavour to best explain its sensory input (Palmer, Seth, & Hohwy, 2015, p. 377).

There are no simple one-to-one links between sensory experience and its possible causes, which can be many and various. Things are even more tricky in the cases of the mental states that are presumed to lie behind and cause behaviour. Unlike the causes of sensory simulation that lie at



the shallow end of the perceptual pool, mental states are much more hidden, usually staying at the deep end.

By engaging with the world to test hypotheses, over and over again, cognisers actualise and maximise their learning potential, gradually improving their accuracy in representing the causal structure of the world. As Hohwy (2018) puts it, ‘the ability to minimise the prediction error over time depends on building better and better representations of the causes of sensory inputs. This is encapsulated in the very notion of model revision in Bayesian inference’ (p. 134).

An efficient system is one that ‘knows’ how to determine what is relevant within a context. Social contexts are much more complex than simple sensory feature detection. Being appropriately sensitive to varying contexts, according to the BBH, is a matter of being able to flexibly adjust the degree of attention given to particular sensory inputs (Hohwy & Palmer, 2014; Van de Cruys, Perrykkad, & Hohwy, 2014). This process, known as precision weighting, is effectively the capacity for determining the relevance of sensory inputs, differentiating between noise and signal. Clark (2016) describes it in terms of the system’s ability to ‘to treat more or less of the incoming sensory information as ‘news’, and more generally in the ability to flexibly to modify [sic] the balance between top-down and bottom-up information at various stages of processing ...’ (p. 226).

To be effective, the precision-weighting of inferences has to be context-sensitive (Ward, 2018; Van de Cruys et al., 2019). For well-adapted systems, learning and experience is the means by which they come to determine the relevance of evidence in the form of sensory inputs by asking whether these contain content that contradicts and should thus revise what is known or expected. Over time, the system becomes better and better at making these adjustments through a bootstrapping process, learning ‘from [changing] regularities in the sensory input’ (Van de Cruys et al., 2019, p. 165). This is a form of ‘experience-dependent learning that accompanies evidence accumulation’ (Friston, 2018; see also Bzdok and Ioannidis, 2019). The outcome of such learning is that the system gains the ability to attend to what is relevant and ignore what is not.

As discussed in the previous section, being able to cope with context-sensitivity is particularly important in social situations considering that they ‘always vary in their sensory details’ (Van de Cruys et al., 2019, p. 165). Crucially, it should now be clear why the Bayesian Brain TT proposals about what underlies social cognition look more promising than their Mental Module TT rivals. This is because according to the BBH the brain is always attempting to calculate which state of the world is ‘*most likely* to be causing the sensory input that our brain receives, given prior beliefs about these causes that are furnished by previous experiences, development, and evolution’ (Palmer et al., 2015, p. 377).

The BBH differs from a ToMM precisely in not being encapsulated. The BBH assumes that there is open channel communication – back and forth – that allows inferences to be updated, such that the whole predictive effort is informed by and updated at all levels. The brain’s hypotheses and inferences about expected causes of the behaviour of others are thus:

situated as part of a causal hierarchy and share reciprocal interactions with higher and lower levels of representation. Thus, mental state inferences are statistically constrained by representations of longer-term expectations – perhaps regarding, for example, the kind of mental states that people tend to have in a given context, or the sense of your friend’s mood that has been reflected in a variety of her behaviours since she showed up to the restaurant, or even culturally defined social contexts and norms. (Palmer et al., 2015, p. 378)

Not only does the BBH overcome the problem of how to account for the context-sensitive use of a ToM in situ, it has been claimed that the approach holds out the promise of illuminating ‘a variety of pathologies and disturbances, ranging from schizophrenia and autism to “functional motor syndromes” (Clark, 2016, p. 3).

Palmer et al. (2015) offer a specific BBH proposal about what explains the signature features of the social cognitive profiles of autistic individuals. Crucially, these researchers conjecture that the autistic social cognitive profiles are a result of the autistic brain’s failure to generate relevant inferences and thus to update or revise expectations about other minds. Autism, on this theory, is the result of a deficient precision estimation system.

Autistic individuals lack the flexibility to determine what is relevant because their brains treat too many sensory inputs as signals. This would explain autistic hypersensitivity or sensory overload (Clark, 2016; Pellicano & Burr, 2012) and autistic hypo-responsivity, which is the absence of context-sensitive responses (Van de Cruys et al., 2019). Accordingly, Clark (2016) suggests, this ‘would result in a constant barrage of information demanding further processing and might plausibly engender severe emotional costs’ (Clark, 2016, p. 225).

Developing this proposal, Clark (2016) suggest that ‘autistic subjects can construct and deploy strong priors but may have difficulties applying them’ (p. 226). Yet impaired precision-weighting capacities may:

lead to overspecific, overfitted internal models that will less efficiently explain away sensory inputs of, for example, social situations that always vary in their sensory details. The sparse, generalizable hidden causes that explain inputs best are not formed (or properly applied). (Van de Cruys et al., 2019, p. 165)

On this view, incapable of sifting out what is and is not relevant, the autistic brain may be unable to form or update its generative model to get an accurate representation of what causes and lies behind the behaviour of others. If the brains of autistic individuals consistently fail to advance appropriate inferences, then they will not have the opportunity to improve the accuracy of their representations of the mental causes that allegedly drive the behaviour of others.

The animating idea behind the BBH is that the work of brains is analogous to what scientists do when making inferences in an effort to best explain phenomena. Brains advance hypotheses, which are developed, refined and improved as those theories are tested against what the world has to offer. As Hohwy (2019) presents it, the BBH's core commitment is that 'perceptual inference is a process that arrives at revised models of the world, which accurately represent the world' (p. 166).

In an earlier publication, Hohwy draws attention to the longstanding analogy, drawn by both Helmholtz and Gregory, 'between perception and scientific hypothesis testing' (Gregory, 1980; Helmholtz, 1860/1962; Hohwy, 2013, p. 77). Clark (2016) makes a direct comparison between what scientific experimenters and brains do (p. 95). Likewise, Yon et al. (2019) note that BBH talk of perceptual inference 'likens perceptual processing to the scientific process' (p. 6).

How seriously should we take the brain–scientist analogy? In what respects should we take it seriously? Sometimes, the analogy is described as 'merely a heuristic description' (Hohwy, 2019, p. 166). Yet, to let go of a realistic reading of the idea that the brain poses and updates inferences raises deep questions about the status of the explanations offered by the BBH. As Hohwy (2018) observes:

If the inferential aspect is not kept in focus, then it would appear to be a coincidence, or somehow an optional aspect of perceptual and cognitive processes that conform to what Bayes's rule dictate. Put, differently, anyone who subscribes to the notion of predictive processing must also accept the inferential aspect. If it is thrown out, then the 'prediction error minimization' part becomes a meaningless, unconstrained notion. (p. 132)

In line with the claim that the BBH embeds a non-negotiable commitment to the brain trading in inferences, Hohwy (2018) tells us that the BBH operates with 'a concrete sense of "inference" where Bayes's rule is used to update internal models of the causes of the input in the light of new evidence' (p. 131).

Hohwy also tells us that the Bayesian way in which the brain updates its models and theories is unlike what goes on in explicit deductive reasoning in key respects. Hence, the BBH does not entail that the brain actually gets its work done by an 'overly intellectual application of theorems of probability theory' (Hohwy, 2018, p. 132). As such, the way the BBH construes the brain's

inferences differs ‘from the use of the term “inference” to describe a higher-order, cognitive and consciously effortful process’ (Palmer et al., 2015, p. 379).

This leads us to the view that the brain’s inferences are implicit and unconscious, unlike those inferences of scientists that are explicit and conscious. The brain’s inferences are swifter and abductive in character and thus unlike inferential operations of the sort found in deductive proofs. Yet, for all that, the brain’s inferences are like the inferences of scientists in being contentful and aiming to get an accurate depiction of the true causal structure of reality. That is what all inferences have in common. That is why ‘mentalising slots into predictive processing as constituting the same kind of unconscious inference that the brain is already engaged into represent [sic] its environment’ (Palmer et al., 2015, p. 378; see Hohwy & Palmer 2014; Kilner, Friston, & Frith, 2007).

What we can conclude from this is that the BBH, like all versions of ‘TT’, is committed to the assumption that we always and everywhere understand others by advancing and improving inferences about the hidden causes of their observable behaviour. This ‘TT’ picture of what underpins social cognition – and what explains typical and atypical varieties – trades upon and gets its life from the assumption that everyday acts of social cognition are rightly characterised in terms of Mindreading. It is that assumption that suggests that we must always adopt a spectatorial stance toward others, even if, unbeknownst to us. In the concluding section, we expose how the spectatorial assumption mischaracterises everyday social cognition and why our exposé should cast doubt on the BBH’s attempted explanations of the social cognitive profiles of autistic individuals.

### **2.3. Characterizing social cognition correctly: diversity not deficit**

‘TT’ explanations of everyday social cognition – both of the general population as well as that of autistic individuals – are only attractive to those who adhere to a Mindreading picture of such cognition. The Mindreading picture is bound up with a host of metaphors. Under its sway, philosophers are wont to claim that individuals have no direct access to other people’s minds; that mental states are the out-of-sight, hidden causes that drive behaviour; that in trying to understand what drives another’s behaviour we need to posit hypothetical entities in our efforts to accurately get at hidden causes, and so on and on.

The Mindreading characterisation of social cognition gets its life from the spectatorial assumption that holds that our situation with respect to others is fundamentally that of a scientific spectator to target phenomena (Hutto, 2004). That assumption is fostered by thinking that the

primary point and pervasive purpose of everyday social cognition is to bridge an assumed epistemic gap that exists between us and others for the purposes of accurately depicting the mental states that move them.

Despite the Mindreading picture's status as the received view, many philosophers have argued that, on close inspection, modelling everyday social cognition on a scientific enterprise paints a distorting picture of its character (Hutto, 2004, 2008; McGeer, 2007, Ratcliffe, 2007). Positively characterised, our everyday social cognition is bound up with engaging with the attitudes and emotions of others, understanding their projects and commitments, trusting or not trusting the accounts that give us why they do what they do. In these practices we are not taking up a scientific stance towards others.

The point of this reminder is not, pace Carruthers, to say that sub-personal theorising cannot be interactive because it is third-personal. The BBH demonstrates that scientific theorising can 'straddle the interaction–observation dichotomy' (Schönherr and Westra 2017, p. 5). The objection to the TT framework made by its so-called phenomenological critics is more fundamental: it is that our everyday engagements with one another are misdescribed when they are depicted as being essentially theoretical in character. We are not always and everywhere attempting to discover the underlying causes of another's behaviour. This is because we are interested in the other's reasons for acting and the best way to get at those reasons is to be told what they are without even having to ask. To understand those social exchanges aright is, to use McGeer's (2007) words, to recognise that we do not 'interact with one another as scientist to object, as observer to observed' (p. 146).

We maintain that reasons to doubt that the Mindreading picture paints a reliable portrait of everyday social cognition are also reasons to doubt that TT proposals, of whatever stripe, can provide the best explanations of social cognition. Thus, relinquishing the spectatorial assumption raises questions about the explanatory appropriateness of applying a TT gloss to characterise so-called sub-personal processes that underwrite social cognition. In other words, giving close attention to the character of our everyday social cognitive practices should make us wary of taking the brain–scientist analogy at all seriously. Rejecting the Mindreading answer to the Characterisation Challenge should make us question the credibility that scientific inference, or even something near enough, really lies at the heart of all of our social cognitive endeavours – namely, that scientific inference really is the driving force in the engine of social cognition.

Some have denied that this sort of conclusion follows. They hold that phenomenological reflections on the character of everyday social cognition do not strongly constrain theorising about its sub-personal drivers. In this vein Spaulding (2018) argues that careful introspection of what it is like to engage in social cognition should not constrain theorising about its underlying mechanisms

because ‘many of our social interactions consist in tacit or implicit mindreading, i.e. subconsciously explaining and predicting targets’ behavior on the basis of attributed mental states’ (pp. 14–15).

Apart from begging the question at issue, the trouble with this line of defence is that phenomenological critiques of the mindreading depiction of everyday social cognition are not based, pace Spaulding, on ‘careful introspection’ of our phenomenology. Rather they are based on giving careful attention to the character of our everyday practices (Hutto, 2013; Hutto & Satne, 2018). In addition, there have been other substantial critiques that raise doubts about the tenability of accounting for the source and basis of the implicit, unconscious contentful inferences upon which the Mindreading story and TT explanations rely (see, e.g. Hutto and Myin 2013, 2017). Together, this clutch of objections constitutes a pincer movement that brings the Mindreading–TT package into question from two directions, above and below.

Under pressure, Spaulding (2018) admits that when characterising the sub-personal processes that allegedly underwrite social cognition ‘one could substitute “interpretation” and “anticipation” for explanation and prediction’ (p. 15). That is certainly closer to the mark and, if we are right, that move has much better prospects of bringing attempts at providing sub-personal explanations of social cognition into line with its actual character.

In this light, Spaulding’s proposed adjustment to the TT gloss is a step in the right direction. An even bigger and better step, in our view, would be to embrace an enactive account of cognition and abandon the quest to find underlying sub-personal mechanisms that explain cognitive phenomena all together (see Hutto & Myin, 2013, 2017).

It matters which of these philosophical frameworks we adopt for thinking about cognition – the choice has practical and ethical significance. Consider that, on the one hand, it is easy to espouse that autistic individuals deserve our full respect and support as ‘persons who try to make sense of themselves and the world’ (Procter, 2001, p. 117). Yet, on the other hand, it is equally easy to hold – at the same time – that the impaired mindreading of autistic individuals bars them from making adequate choices in the social domain and, as such, ‘caring for them may require making these choices for them’ (Procter, 2001, p. 114). This way of reasoning can lead to bad outcomes, as in Melanie Yergeau’s case. She reports a harrowing story in which she was forcibly detained by therapists-cum-faculty in which, in her words: ‘I found myself deeper within a narrative of neurological determinism ... Regardless of what I said, it was my autism saying it’ (Yergeau, 2013). Enactivist approaches to mind and cognition give us tools for resisting rather than encouraging the idea that who ‘we’ are reduces to something inside us, that who ‘we’ are is the product of something inside us – the intelligent, sub-personal activity of our brains.

The foregoing analysis is not designed to deny or obscure the fact that certain autistic and cognitively typical individuals, given their particular cognitive capacities and profiles, find some kinds of social cognitive tasks difficult, perhaps even impossibly so. It does, however, serve to remind us that the cognitive challenges and achievements in question are challenges and achievement of individuals – of persons – and that these are not best explained by focusing solely or primarily on sub-personal parts of people.

Had there been more space, it would have been illuminating to provide more detail about enactivist alternatives to the mainstream cognitivist approaches that take the Mindreading characterisation of social cognition as given. In lieu of doing so in more depth, we must be satisfied with dwelling on this lesson for now: once we stop thinking of the main action of social cognition as happening in the heads of individuals and put it back in the space of interactions themselves it becomes clear that successful social engagement is a joint responsibility.<sup>2</sup> It is best conceived of as a shared endeavour in which adjustments need to be made by all parties involved to ensure successful outcomes. Success in these tasks, given their true point and purpose, is not something that can be achieved privately and separately in the heads or brains of the individuals involved, however remarkable those individuals may be at inferring each other's mental states.

Enactivists conceive of cognition in terms of dynamic, 'out-of-the-head', world-involving activities. Emphasising these aspects, they make much of the metaphor that cognitive engagements are a matter of 'laying down a path in walking'. That metaphor helps to illuminate that in all cases, when it comes to completing cognitive tasks there are other possible ways of getting to the same place. Getting there by other ways might require going slower, or taking a different path than the beaten one, or they might require the cooperation and assistance in meeting one part of the way.

The pivotal point is that we have reason to surrender the idea, enshrined in the intellectual individualism embraced by mainstream cognitivism, that successful social cognition depends on and aims at a gap-bridging epistemic achievement. Concomitantly, we have reason to avoid the idea that such gap-bridging can only be achieved by the theorising that goes on in the brains of individuals.<sup>3</sup> If we manage to resist these prevalent pictures, we open the door for thinking of the success of social cognitive tasks in ways that are not purely epistemic and to recognising that the success of such engagements depends on and is the mutual responsibility of all of the individuals, both autistic and cognitively typical, involved in social encounters.

---

<sup>2</sup> The double empathy hypothesis (Milton, 2012) states that cognitively typical people have just as much difficulty empathising with autistic people as vice versa because difficulties will inevitably arise when different cognitive styles are in communication. This hypothesis is based on the view that autism involves autistic cognition, not impaired neurotypical cognition.

<sup>3</sup> We do not discount the importance of brain-based aspects of cognition. The monotropism theory (Murray, Lesser & Lawson, 2005) is regarded by many autistic scholars, including the third author, as capable of describing the features of autism, including differences in sensory perception and sociality.