

Out of the Mouths of Autistics: Subjective Report and Its Role in Cognitive Theorizing

Victoria McGeer

The theoretical work that emerges from a study on the work of memory, learning, and other higher functions, such as consciousness, is this: if the psychological (functional) taxonomy is ill-defined, then the search for neural substrates for those functions will be correspondingly ill-defined. There are, certainly, remarkable data, found at all levels in psychology and neuroscience, but precisely how to interpret the data in terms of a theory of neurobiological capacities, representations and processes is yet to be discovered.

– Patricia Churchland, *Neurophilosophy*

My primary concern in this chapter is with subjective report, and in particular, first-person reports of abnormal sensory and/or perceptual experiences. This topic raises interesting questions at two distinct levels: First, there are philosophical questions about the nature of subjective experience, subjective awareness of experience, and subjects' capacity to articulate what they are experiencing. Secondly, there are questions about how philosophical theories of such matters interact with empirical theories of – and research into – abnormal neurocognitive conditions. To focus my discussion of these questions, I will be considering, in particular, the phenomenon of subjective report in high-functioning individuals with autism.

1 Understanding Autism: Some Methodological and Substantive Concerns

Autism presents a highly complex challenge for researchers trying to negotiate between neurological and cognitive levels of theorizing. As

with other neurodevelopmental disorders, such research involves working along two dimensions at once. As T. W. Robbins explains:

The challenge to research into childhood autism lies in relating what appears to be a set of apparently somewhat independent symptoms . . . to corresponding deficits in brain systems. This research approach can be viewed as one of 'vertical integration' in which theories in one domain are strengthened by structural congruences in another. . . .

However, autism generally presents as a set of symptoms, the precise significance of which, relative to one another, remains to be established. . . . Analogously, the main neurobiological findings in autism seem to implicate many possible foci which are probably affected to varying extents in different individuals. These are problems of 'horizontal integration' and it seems evident that a complete understanding of the disorder has to achieve an orderly account in both the 'vertical' and the 'horizontal dimensions'. (Robbins, 1997)

There is one further dimension of complexity to add to this challenge: the problem of 'temporal integration'. Autism is a neurodevelopmental disorder, unfolding in time and hence affecting individuals (and their brains) in cascading complicated ways throughout their development. As Helen Tager-Flusberg notes,

. . . neurodevelopmental disorders are more often associated with *diffuse* cortical damage, which suggests that the impact of such disorders is more widespread, affecting complex neural systems rather than simple localized areas. Furthermore, across a range of developmental syndromes, we find that not only are particular cortical systems affected but often associated atypical subcortical structures are involved as well. For example, in autism both the cerebellum and the limbic system show significant abnormalities. These findings suggest deviations in brain development that begin early in embryology and cannot be easily classified and interpreted as later acquired focal lesions. (Tager-Flusberg, 1999)

One suggestion here is that the range of symptoms found at the cognitive/behavioural level are not related to one another at all but derive simply from abnormalities in diffuse underlying neural systems. Another possibility, not excluded by this first, is that abnormalities in specific cortical areas make for a primary cognitive deficit – for instance, a hypothesized 'theory of mind' deficit (discussed later in this section) – and this primarily deficit is responsible for (i.e., functionally related to) at least a subset of the cognitive/behavioural symptoms. A third 'neuroconstructivist' possibility proposes a developmentally more complicated set of relations among the various symptoms: Early neurological abnormalities, with typically low-level (noncognitive) behavioural manifestations, affect the developing child's interactions with its environment,

leading to atypical inputs for the developing brain. These atypical inputs themselves encourage atypical development in a number of high-level cortical structures, leading to an overall profile of abnormal brain functioning. Symptoms at the cognitive/behavioural level will both speak to and reflect this pattern of atypical development, making the relations among them impossible to sort out in purely atemporal terms (for further discussion of these alternatives, see Gerrans and McGeer, 2003). As Annette-Karmiloff Smith explains:

The neuroconstructivist modification in perspective crucially influences the way in which atypical development is considered. In this approach, the deletion, reduplication or mispositioning of genes will be expected to subtly change the course of development pathways, with stronger effects on some outcomes and weaker effects on others. A totally specific disorder will, *ex hypothesi*, be extremely unlikely, thereby changing the focus of research in pathology. Rather than solely aiming to identify a damaged module at the cognitive level, researchers are encouraged to seek more subtle effects beyond the seemingly unique one, as well as to question whether successful behaviour (the presumed 'intact' part of the brain) is reached by the same processes as in normal development. This change in perspective means that atypical development should not be considered in terms of a catalogue of impaired and intact functions, in which non-affected modules are considered to develop normally, independently of the others. Such claims are based on the static, adult neurophysiological model which is inappropriate for understanding the dynamics of developmental disorders. (Karmiloff-Smith, 1998)

With these methodological considerations in mind, consider now the range of symptoms with which researchers in autism must contend. To begin with, autism is a spectrum disorder, with individuals (usually designated 'low-functioning') that are severely retarded at one end of the spectrum and those (usually called 'high-functioning') that have normal to above average IQ at the other. (A substantial proportion of high-functioning autistics are sometimes given the differentiating diagnosis of 'Asperger's syndrome', but for the purposes of this chapter, I will consider these two high-functioning groups together, using the terms 'Asperger's syndrome' and 'high-functioning' autistic interchangeably.)¹

¹ In the latest editions of the American Psychiatric Association Diagnostic and Statistical Manual (DSM-IV) and the World Health Organization. International Classification of Diseases (ICD-10), 'Asperger's Disorder' is listed as a distinct nosological entity under the category of Pervasive Development Disorders (along with 'Autistic Disorder', 'Rett's Disorder' and 'Childhood Disintegrative Disorder'). There is still substantial disagreement amongst clinicians and researchers about whether Asperger's syndrome, in particular, really constitutes a distinct disorder from autism or merely lies on the milder

The core cognitive deficits, according to which diagnosis is made, involve characteristically impaired social, communicative and imaginative skills – the last, for instance, typically identified through an absence of pretend play in early childhood. Communicative deficits are both prelinguistic and linguistic, involving, for instance, an absence of proto-declarative pointing (to call attention to things in a shared environment), an absence of normal gestural speech, and abnormal prosody. There is limited or absent awareness of conversational cues, or the give and take of normal conversation. Pronoun reversals – for example, ‘I’ for ‘you’ – are also a common feature. As for social skills, even high-functioning autistics – those with relatively good or even superior cognitive skills – show distinctive impairments in these respects. If there is curiosity about the world, it is manifestly not directed towards other people, who seem to occupy a predominantly instrumental role in the autistic’s environment. (Those who interact with autistics often report feeling like ‘objects’.) Autistics show little or no understanding of the give and take of social life, and no capacity to engage with others in anything like a normal way. They generally occupy the world as if on their own, often developing and adhering to strict routines that make little sense from our nonautistic point of view. Similarly, their activities and interests seem to us repetitive, highly restricted and often deeply odd. Memory skills amongst autistics can sometimes be extraordinary, and a small proportion also show ‘savant talents’ in music, calculating, drawing and so on.

While these cognitive features are certainly distinctive, they do not encompass the totality of autistic abnormalities. Motor stereotypies (hand flapping, spinning, rocking) are a common feature throughout the autistic range, and motor clumsiness is also sometimes apparent, particularly amongst high-functioning autistics (perhaps distinctively associated with Asperger’s syndrome (Green et al., 2002)). Extreme emotional

end of a continuum (Klin et al., 2000; Ozonoff et al., 2000; Mayes et al., 2001; Mayes and Calhoun, 2001). The distinguishing features of Asperger’s syndrome are no evidence of general language delay or of delay in reasoning skills outside the social domain. However, other symptoms characteristic of autism, including abnormalities in language use, are generally present. Moreover, individuals who do have language and/or cognitive delay – those who fall within the current categorization of ‘Autistic Disorder’ (DSM-IV) or ‘Childhood Autism’ (ICD-10) – may develop skills in both these respects that lead to their being designated ‘high-functioning’ (For further discussion, see ‘A Note on Nosology’ in Baron-Cohen et al., 2000). Given the substantial overlap of symptoms between these disorders, it seems unlikely that the points I want to make in this chapter are affected by the diagnostic distinction. However, this is a matter for further research.

reactions, moodiness, tantrums and symptoms of anxiety disorder are also frequently observed.

Of particular interest is a further diagnostic item listed in connection with both autism and Asperger's syndrome in DSM-IV – namely, a 'persistent preoccupation with parts of objects'. Perceptual abnormalities that involve a marked preference for featural over contextual processing have been interestingly explored, with some research even indicating that autistics are less susceptible than normal or Downs syndrome control subjects to low-level perceptual illusions induced by an embedding context (e.g., the Ebbinghaus [or Titchener circles] illusion, the Ponzo illusion, the Poggendorf illusion and so on) (Happé, 1996). These abnormalities may indicate a more general cognitive tendency towards what Uta Frith calls 'weak central coherence' – a relative inability to process information, whatever its form (e.g., visual, syntactic, semantic and so on), in context (Frith, 1989). How such a tendency may be related to the particular social-cognitive difficulties described here is still unclear (Frith and Happé, 1994; Happé, 2000). However, what is clear is that such perceptual abnormalities constitute a significant element of the autistic phenotype, importantly shaping how individuals understand the world.

In addition to these perceptual abnormalities, there is another prominent aspect of the disorder – at least as revealed in subjective report – that has received surprisingly scant attention from cognitive researchers, either in their theoretical accounts or in their empirical research – namely, the prevalence of sharply abnormal sensory experiences. As Uta Frith herself remarks:

One mysterious feature that is not currently given much importance may hold further clues. Some Asperger individuals give first-hand accounts of sharply uncomfortable sensory and strong emotional experiences, often including sudden panic. From autobiographical accounts, we learn that again and again the Asperger individual's interpretation of perceptions by ear, eye, or touch tends to be extremely faint or overwhelmingly strong. There can be hyper- as well as hyposensitivity. Feeling scratchy clothes, for example, is not merely uncomfortable, but agonizing. On the other hand, pain may be tolerated to a surprising degree. (Frith, 1991)

While these reports are certainly interesting, revealing aspects of the disorder that seem otherwise overlooked, it is worth asking how much attention researchers should pay to these first-hand accounts. After all, as Frith goes on to observe, these kinds of subjective reports are restricted to a fairly small subset of the wider autistic population (about 20–25%) – a subset that is significantly reduced if those with Asperger's syndrome

are excluded. That is, they are gleaned only from individuals who are linguistically able, who tend to have normal or above-average IQs, and who very often have a late-developed capacity to pass standard 'theory of mind' tests, such as the now famous false-belief task (though it must be stressed that such a capacity does not, in these cases, indicate anything like a capacity for normal social reasoning). All in all, such individuals are unusual relative to the general autistic population; and while they may have a milder form of the same underlying disorder affecting more seriously disabled autistics, how far we can generalize from these reports remains an open question.

Nevertheless, despite exercising caution in taking verbally adept individuals to speak for the experiences of autistics in general, their reports do seem to give us a rare glimpse of what it is like to be autistic, or what the world is like from an autistic point of view. As such, they provide a unique and important set of data for research in autism, data that underscores the existence of phenomena that may not be so salient from the third-person point of view: for example, the extent and impact of autistic children's sensory abnormalities in coming to learn about and negotiate the world around them. Indeed, these abnormalities and their consequences may be peculiarly invisible from the third-person point of view – at least if we rely on it exclusively – unduly constraining proposed hypotheses for explaining some aspect of autistic behaviour.

Take, for instance, the well-known phenomenon of autistics' failing to make normal eye contact when interacting with other people. (In fact, this is one of the targeted features of social interaction in which autistics are often given explicit training: 'When you first meet people, look them in the eyes.') How might we explain this characteristic behavioural abnormality?

Here's one possibility (explored in Baron-Cohen, 1996): Although autistics have no trouble understanding what the eyes are for (i.e., looking at things), and although they are able to understand that someone is looking at them or at something else, depending on the direction of that person's gaze (as judged from their interpretation of photographs or cartoons), they take no interest in what others are looking at. They have no drive towards sharing attention – towards looking at what the other is looking at, or to securing joint attention; and they don't seem to particularly care whether or not someone is looking at them. Simply put, the eyes are not a particularly interesting or information-rich stimulus for autistic individuals; they don't make eye-contact because they're not motivated to do so.

Contrast this hypothesis with a second possibility: Far from the eyes constituting a relatively impoverished stimulus for autistics, it may be that normally sustained eye contact is too arousing for them to bear. After all, if the eyes are normally such a salient feature to human beings, and if autistics are hypersensitive in certain sensory domains, then it is at least possible that the eyes fall into this domain. Consequently, autistics avoid eye contact, not because they find other people's gaze informationally uninteresting but because they can't make use of the information available, given how sensorily overwhelming it is to look directly into another's eyes.

Why would such a hypothesis occur to us? Perhaps only by paying more heed to subjective report. For example, on being asked to make eye contact with people when speaking to them, one resistant young autistic complained, 'Their eyes are on fire!' Or consider another report from Therese Jolliffe:

Looking at people's faces, particularly into their eyes, is one of the hardest things for me to do. When I do look at people, I have nearly always had to make a conscious effort to do so and then I can usually only do it for a second. . . . People do not appreciate how unbearably difficult it is for me to look at a person. It disturbs my quietness and is terribly frightening – though the fear decreases with increasing distance away from the person. (Jolliffe et al., 1992)

These reports certainly suggest that sensory disturbances have something to do with this standardly cited example of an autistic social-cognitive deficit, making this second hypothesis a live, but overlooked, research option. Hence, I want to suggest there is a blind spot in autism research. On the one hand, there is a great deal of effort and attention devoted to understanding autistic abnormalities and the connection among them, as these are or have become evident from the third-person point of view. On the other hand, there is a growing body of (anecdotal) evidence from autistic self report – mostly in the form of autobiographical writings – that sensory abnormalities are a central feature of autistic experience, and perhaps central to understanding the nature of the disorder. Yet this possibility seems to be generally ignored or at least substantially underexplored by cognitive theorists. The question is: why? My speculative answer is that subjective reports of abnormal experiences are liable to fall prey to two methodological cum philosophical objections that ordinary subjective reports do not; and for this reason, they tend to be ignored or discounted – explained away, rather than explained in terms of the phenomena they straightforwardly express.

Of course, dealing with subjective report in cognitive research is always a tricky business, and there are two questions that really ought to be posed in connection with it: (1) In any particular case or kind of case, are these reports reliable? – that is, do they give us accurate information about what individual subjects are (really) experiencing? and (2) how can we explain what the subjects are saying? Notice that if the answer to the first question is ‘no – subjects are not reliable in making these reports’, then the type of answer given to the second question is significantly altered: Instead of trying to explain subjects’ experience as they report it to be, the researcher must now account for why they *say* what they do, given that what they say doesn’t accurately represent how they are experiencing things. For example, in cases of hysterical blindness, subjects with no apparent physiological damage may *say* that they are blind. But researchers would be justified in doubting the credibility of these reports if, for instance, these subjects do significantly *worse* than chance in forced-choice visual tests, thereby showing their capacity to use visual information to support the contention of blindness (discussed in Dennett, 1991, pp. 326–327).

Clearly, then, the first ‘reliability’ question is a critical one to pose. Still, I will be arguing that there are certain background views about subjective report that may bias the answer towards false negatives when it comes to assessing their credibility in cases where what’s reported is especially strange or unusual. These are the two methodological cum philosophical concerns I referred to: The first is not as serious as the second, though it sets the stage for examining a pervasive way researchers in both philosophy and psychology tend to regard subjective report – namely, as a sort of testimony to the nature of something that’s observed from only one point of view. I call this first source of bias a Humean resistance to miraculous testimony. The second, more significant, source of bias, which I discuss in section 3, involves an adherence to a ‘neo-perceptual’ model of introspective awareness and self-report. The important and distinctive thing about this model is that it makes self-awareness dependent on a properly functioning cognitive mechanism for tracking our mental life; and, of course, the breakdown of such a mechanism is just what might be expected in pathological cases.

These are both rather general considerations in the sense that if they count at all, they count against crediting all anomalous subjective reports – subjective reports made in a variety of pathological or abnormal cases. In my final section 4, I will consider a third, seemingly additional, reason for discounting subjective report in the particular case of autism.

It stems from what many researchers consider to be a central – perhaps, *the* central – deficit in this disorder: namely, a domain-specific inability to conceptualize, attribute or reason about specifically *mental* states – whether of *self* or other (for discussion, see Baron-Cohen et al., 1993; Baron-Cohen et al., 2000). This deficit is usually characterized as an inability to develop or deploy a ‘theory of mind’ – or ‘ToM’. However, as we shall see, not all researchers who subscribe to this view think theory of mind is theorylike in structure, or learned in a way that theories ordinarily are.

I will not here review the empirical research that supports the idea of such a domain-specific disability in autism. I take it to be rather compelling. My concern instead is with how this disability has been conceptualized by certain influential researchers in the field – and, specifically, with how this conceptualization affects researchers’ attitudes towards autistic self-knowledge and self-report. In particular, my concern is this: If having such a disability can be equated with *lacking normal introspective awareness* (Frith and Happé, 1999), then it seems reasonable to discount the reliability of autistic subjective report in general. In this case, anything unusual in these reports would be explained away in terms of the global introspecting disability, rather than explained in terms of the phenomena directly reported. Or to put this another way, the ‘valuable data’ of autistic subjective report would be valuable just in the sense of providing more evidence for a specific ToM disability. So my goal, in this final section, is to examine whether the equation between a ToM disability and lack of introspective awareness makes good theoretical or empirical sense. I will close by considering some implications of this discussion for future research in autism.

2 Subjective Report and the Humean Resistance to Miraculous Testimony

If we begin with the naive question – what is subjective report? – a commonsense answer immediately suggests itself: It is a kind of testimony to the states or events that constitute individuals’ subjective condition. That is, it is a testimony to how individuals find they are experiencing the world (including their own bodies); or it is a testimony to what they believe, desire, hope or fear – a testimony to the nature and existence of their own intentional states.

Now with regard to testimony generally, we may ask a further question – namely, when is it rational to credit another person’s reports of

states and events that we do not witness ourselves? Again, a common-sense answer suggests itself: when these states and events fall within the range of normal, expectable, or at least *explicable* occurrences. But what about reports that fall quite outside this normal range? Is it rational to credit such testimony then? In the *Enquiries Concerning Human Understanding*, David Hume considers this question with respect to testimony about the occurrence of 'miracles': that is, events that violate laws of nature and thereby stand outside of any normal experience. His answer there is unequivocally 'no'. He writes:

No testimony is sufficient to establish a miracle unless the testimony be of such a kind, that its falsehood would be more miraculous, than the fact which it endeavours to establish; and even in that case, there is a mutual destruction of arguments, and the superior only gives the assurance suitable to that degree of force, which remains, after deducting the inferior. (Hume, 1975, Section X, I, 91)

To put Hume's conclusion in modern parlance: *the less probable the reported event, the less credible the report*. Here is Hume again:

Suppose that the fact which the testimony endeavours to establish, partakes of the extraordinary and the marvellous; in that case, the evidence, resulting from the testimony, admits of a diminution, greater or less, in proportion as the fact is more or less unusual. (Section X. I. 89)

Consider now the following autistic subjective reports:

I had – and always had had, as long as I could remember – a great fear of jewellery. That terror also included hairclips and metal buttons. I thought they were frightening, detestable, revolting. If I was made to touch jewellery, I felt a sharp whistling metallic noise in my ears, and my stomach turned over. Like a note falsely electrified, that sound would creep from the base of my spine upwards until it rang in my ears, tumbled down into my throat and settled like nausea into my stomach. (Gerland, 1997, pp. 54, 157)

As much as I loved to chew scratchy and gritty textures, I often found it impossible even to touch some objects. I hated stiff things, satiny things, scratchy things, things that fit me too tightly. Thinking about them, imagining them, visualizing them . . . any time my thoughts found them, goose bumps and chills and a general sense of unease would follow. I routinely stripped off everything I had on even if we were in a public place. I constantly threw my shoes away, often as we were driving in the car. I guess I thought I would get rid of the nasty things forever! . . .

I also found many noises and bright lights nearly impossible to bear. High frequencies and brassy, tin sounds clawed my nerves. . . . Bright lights, mid-day sun, reflected lights, strobe lights, flickering lights, fluorescent lights; each seemed to sear my eyes. Together, the sharp sounds and bright lights were more than enough to overload my senses. My head would feel tight, my stomach would

churn, and my pulse would run my heart ragged until I found a safety zone. (Willey, 1999)

I perceived sound and visual information directly and consciously only at the cost of its cohesion. I could interpret the part but lost the whole. I saw the nose but lost the face, saw the hand but continued to see the body but would not know what it was except piece by piece. I'd get the intonation but lose the meaning of the words or get the meaning of the words but only at the cost of tuning out the intonation, as though independent of the words. (Williams, 1999)

How far do we credit such reports? After all, they certainly sound 'extraordinary and marvellous' in a Humean sense. Or consider another example: subjective reports of so-called synaesthetic experience, where subjects purport to have experiences in one modality as a consequence of stimulation in a quite different modality. For instance, they report seeing particular graphemes as tinged always with particular colors (5s as red, for instance), or hearing particular musical tones as colored in particular ways (C#s as purplish-blue). Falling outside our own experience – that is to say, *normal* human experience (in a statistical sense), it may be tempting to apply Hume's principle: Such reports are simply incredible. In fact, as V. S. Ramachandran and E. M. Hubbard comment in a recent article on synaesthesia, 'despite a century of research, the phenomenon is still sometimes dismissed as bogus' (Ramachandran and Hubbard, 2001b, p. 4). Could this resistance be due in part to an application of Hume's principle? If so, I think it is misguided for two distinct reasons.

In the first place, Hume's reasoning depends on assessing the improbability of the reported event. But how do we go about making these assessments when it comes to reports of subjective experience? Using 'normal' experience as a guide will be reasonable only insofar as we are quite convinced that no organic abnormality could account for the experiences *as described*. But this must come after a thoroughgoing investigation of the phenomena concerned, rather than providing a priori grounds for undercutting such investigations. After all, given the complexity of the neural systems realizing our experience and given how little we still understand about these neural systems, the possibility of such untoward phenomenon is not something we can reasonably assess simply in light of normal experience.

A second, more important, reason is that applying Hume's principle in the case of subjective report means taking the situation of testimony that Hume envisaged far too literally. In effect, it is to imagine an old-fashioned picture of the mind in which subjects themselves are the only witness to

the putative events reported – only they have direct perceptual access to their own subjective condition; and cognitive scientists are ‘on the outside’, depending exclusively or primarily on what their subjects say.

In fact, this is not the position of cognitive – or consciousness – scientists. (Whether or not it’s the position of subjects themselves, I’ll return to in section 3.) As Ramachandran and Hubbard have elegantly demonstrated in a recent set of experiments on synaesthesia, scientists do have ways of directly verifying what their test subjects say about themselves. More, their methods allow them to explore the mechanisms that explain why subjects’ experiences take the shape that they do. For instance, after noting that certain of their grapheme-color synaesthetes described their visual experience in ways that suggested a fairly low-level visuo-perceptual effect, Ramachandran and Hubbard reasoned that these subjects should experience certain phenomena under experimental conditions that are generally accepted as indicative of low-level perceptual processes – phenomena such as perceptual grouping and ‘pop-out’.²

Thus, in one such experiment designed to test for synaesthetic pop-out, they asked their subjects to identify a geometric shape notionally inscribed by the 2s embedded in a matrix of 5s. The matrix was presented in black and white, as shown in Plate 1 (following p. 310). Their responses were then compared with nonsynaesthetic control subjects. As might be expected, control subjects took some time to identify a triangle. The synaesthetic subjects, by contrast, ‘were significantly better at detecting the embedded shape’, demonstrating a genuine pop-out effect (such as displayed in Plate 2, following p. 310) (Ramachandran and Hubbard, 2001b, p. 7; see also Ramachandran and Hubbard, 2001a). These and other experiments demonstrate quite conclusively that these subjects’ synaesthetic experiences were very much in keeping with how they reported them to be.

The upshot is plain. Though ordinary folk may not be in a position to verify what others say about their experience, scientists certainly are

² It’s worth noting that in order to get this far, Ramachandran and Hubbard acknowledge a need for ‘probing the introspective phenomenological reports of these subjects, even though this strategy is unpopular in conventional psychophysics. For instance, to the question: “Do you literally *see* the number 5 as red or does it merely remind you of red – the way a black-and-white half-tone photo of a banana reminds you of yellow?” our first two subjects replied with remarks like, “Well, that’s hard to answer. It’s not a memory thing. I do *see* the colour red. But I *know* it’s just black. But with the banana I can also imagine it to be a different colour. It’s hard to do that with the 5s”, all of which further suggests that synaesthesia is a sensory phenomenon’ (Ramachandran and Hubbard, 2001b, note 3, p. 7).

if they are clever enough to figure out what other discriminating behavioural manifestations there will be of the nature and quality of a person's experience beyond verbal report. Careful scientific investigation at the behavioural level can give us – in fact, already has given us – good reason to think that abnormal experiences, unlike miracles, do occur after all. Moreover, we may have such reason even in the absence of direct physical evidence, or understanding, of the underlying organic cause of various particular abnormalities. These observations point us towards three related conclusions: First, in the case of strange or unusual sensory/perceptual reports, we have no *prima facie* grounds for doubting the reliability of the subjects in question, especially when their reports are systematic and robust; their reliability or unreliability must be established by further investigation. Secondly, if subjects are somehow miscuing us about the nature of their own experience, there is no reason to think that such miscuings cannot be detected from the third-person point of view – or indeed that what they say cannot be objectively verified. Thirdly, this very fact casts some doubt on the commonsense conception of subjective report with which we began. I turn now to a fuller examination of this final point.

3 Subjective Report as Grounded in Competing Philosophical Models of Introspection

3.1 *The Neoperceptual Model*

As indicated earlier, Hume's reasoning about miraculous testimony fits particularly well with a testimonial conception of self-report according to which persons are testifying to something that they witness but that their interlocutor does not – namely, their own experience. This way of putting things has a rather perceptual flavour not to be taken literally anymore – that is, not in the sense of persons casting an inner eye over their experience. Nevertheless, many philosophers and cognitive scientists are drawn towards thinking of reflective self-awareness in quasi-perceptual terms. In fact, they endorse what I will call a 'neoperceptual' (NP) model of introspection, which replaces the 'inner eye' with a respectable subpersonal cognitive mechanism.³ The key features of the

³ The contemporary locus classicus for this view is Armstrong (1968, 1993). For critical discussion, see Shoemaker's account of the 'broad' perceptual model in Shoemaker (1996, essays 10 and 11).

neoperceptual model are these:

- i. The mind/brain is composed of what I will call *first-order* states and processes: for example, sensations, perceptions, emotions, beliefs, desires and so forth – states and processes that, insofar as they have representational content at all, represent conditions in the world (including conditions of our own bodies).
- ii. We are aware of some of these states/processes courtesy of a (subpersonal) cognitive mechanism that causally produces in us independently existing *second-order* or '*meta-representational*' beliefs: beliefs whose representational content is about these first-order (mental) states and processes. This causal process constitutes the special faculty of 'introspection' or 'introspective judgement'.
- iii. This causal process is generally reliable, at least in normal individuals under normal conditions.
- iv. Subjective reports are normally the direct linguistic expression of the *second-order beliefs* thus generated.

Notice that in this model, there are two distinct sources of error in self-report: First, persons may not be giving adequate linguistic expression to their second-order beliefs, because of linguistic incompetence, confusion, distraction, memory problems and so forth. And, second, persons may not be expressing higher-order beliefs that are properly formed by this mechanism at all, because the mechanism is broken down, malformed or just not deployed. Some philosophers – for example, Alvin Goldman – may regard this second source of error as an advantage of the model, since it suggests that abnormal reports in pathological conditions could well be explained in terms of this sort of breakdown (Goldman, 1997). However, since this kind of breakdown is always possible on the NP model, it also suggests that researchers should not simply help themselves to the assumption, stipulated in iii, that our introspective mechanisms are generally reliable. Thus, as Goldman himself worries, if researchers want to use subjective reports as evidence for what their subjects are experiencing, it seems that they must first establish the reliability of their subjects' introspective mechanisms, even in cases where the first source of error is ruled out – that is, there is no evidence of linguistic incompetence, confusion, distraction and so on (Goldman, 1997; Goldman, 2000).

As I indicated in section 2, this task of verifying and even exploring what subjects claim about their experiences is not beyond the reach of cognitive scientists. Still, the NP model introduces, by its very structure, a source

of error or unreliability in these reports that may cast inappropriate doubt on the *prima facie* credibility of subjective report in both normal and pathological conditions. To see why this doubt is inappropriate, I introduce an alternative model of self-report and reflective self-knowledge that simply rules out this putative source of error.

3.2 *The Reflective-Expressivist Model – Self-Knowledge of Intentional States*

The reflective-expressivist (RE) model builds on a feature already present in the NP model of self-knowledge and self-report. Recall that on the NP model, we report our first-order states by directly expressing our second-order beliefs about them: That is, we directly express what is believed (e.g., that we have a certain belief or experience). On pain of regress, we don't have to form any third-order beliefs in order to express the contents of our second-order beliefs – namely, that we have a certain belief or experience. We simply express these contents directly. In fact, advocates of the NP model should accept the fact that we also very often just directly express our first-order states: If I say 'that's a latte', then under the usual provisos of sincerity and linguistic competence, I am taken to directly express the content of my belief that there's a latte before me. It only seems necessary to depart from this basic expressivist structure when we're asked to engage in a certain mode of self-reflective activity – that is, when we're asked to pay attention to ourselves, as against (so it seems) the world – and say how it is with our own subjective condition.

But now let us examine this self-reflective task more closely. Suppose I spy a drink in front of me and judge (or take) it to be a latte. Phenomenologically, as all sides agree, it seems that when I'm asked whether I believe – that is, really *believe!* – that there's a latte before me, all I can do to make my answer is attend more closely to the apparent qualities of the drink itself: Does it look like a latte? Smell like a latte? Do I have background info that would call my conclusions into doubt? Does the drink still command my unswerving latte judgement? Moreover, if I perceptually 'retake' the drink in this way – that is, as a sophisticated folk-psychologist with background knowledge, however implicit, of the norms governing belief ascription – then I know that my positive reassessment of the evidence that leads me to say there's a latte before me is sufficient for my saying that I truly believe there's a latte before me. In other words, I don't seem to answer the belief question about myself by scrutinizing or tracking my own internal states at all. I simply *speak out of my* (first-order) latte-believing condition. (For this way of putting the point and a congenial defense of this type of expressivist view, see Bar-On, 2000; Bar-On and Long, 2001)

Of course, this phenomenological evidence is not conclusive; it can be accounted for in a number of ways. Even so, it makes vivid the conceptual point that knowing what I believe really only depends on two things: (1) having background knowledge of what believing in general requires – namely, a robust inclination to judge something is the case; and (2) the capacity to make and express judgements that report particular features of the world. Hence, there is no barrier in principle to embracing the comparatively minimal RE model of self-reflective activity. The key feature of this model is that such activity does not depend on a capacity for tracking one's own first-order states in addition to tracking what those first-order states are about. One's first-order states are not the direct target of cognitive or even subcognitive activity. Instead, such activity depends simply on a capacity to look at the world again, prompted by the recognition that it may not be as it seems – that is, as one initially judges it to be – and trying to rule out possible sources of error.

There are two important things to note about this minimalist approach to subjective knowledge and subjective report. First, what makes self-reflective activity distinctively *reflective* is that a person's intentionally expressed judgements – that is, judgements expressed with intention and in intentional language – depend on (a) knowing about folk-psychological concepts and the norms that govern them – that is, what counts as believing, hoping, desiring and so forth; and (b) knowing how the norms apply in particular cases (for instance, is my hesitation in reaffirming that this drink is a latte due to something detectably strange about the drink itself? – in which case I may register my doubt by hedging what I claim about myself. Or am I hesitating because I'm surprised by your question? I look again, I see it's a latte; of course, I believe it's a latte!). The second thing to note is that if a person's knowledge of these norms is limited or abnormal, the capacity to express (intentional) self-reflective judgements will be correspondingly compromised. This will be relevant to the case of autism, which I return to in section 4 below.

3.3 *Extending the Reflective-Expressivist Model – Self-Knowledge of Experiential States*

So far I've only talked about subjective reports, or better subjective expressions, of first-order intentional states. But what about subjective expressions of our own sensory/perceptual experiences, expressions of how things look, taste, smell and so on? Are there any special problems for the reflective expressivist here?

Suppose I catch a whiff of that irresistible latte smell, or better, have a close encounter with that irresistible latte taste. If I say, 'that smells (or tastes) like a latte', then on the NP model, I must be scanning or tracking (perhaps subcognitively) some internal mental state of mine – a latte smell or taste experience. On the RE model, by contrast, I just directly express how I am experiencing things; I directly express how I take things to be, according to one sense modality or another.

But now suppose I'm asked to be more reflective about my experience, to say more precisely what my experience is *really* like. How do I go about answering this question? Once again, it seems that the only way to respond is by focussing my attention on the world, just as I do when I'm answering the belief question. Only now I focus on how in this moment things in the world, including states of my own body, really smell, taste, feel and so forth. According to the reflective expressivist, then, self-reflection in the sensory case is the same as in the belief case: It does not involve any special, introspective sort of mentalistic scanning; it does involve a special – that is, attentive, focussed – redeployment or reengagement of my sensory systems towards some aspect of the world. It involves my perceptually retaking the world in a certain modality, and it is this retaking that I express in confirming self-report: 'It smells like a rich, dark, smoothly roasted latte (to me)' (cf. Harman, 1990).

Briefly, then, and schematically, there are three things to note about making experiential reports on the RE model: First, I am able to *know* that I'm having a certain olfactory experience, say, just by virtue of being able to focus on *how* things smell in the world. Secondly, I am able to communicate that I have such an experience to others simply by virtue of knowing how to say how things smell in words (e.g., it smells bitter, sweet, like roses and so on). Thirdly, and most significantly, knowing how to say in more elaborate detail how things smell, taste, feel and so forth does not require any skilled folk-psychological understanding of the norms governing the ascription of belief and other intentional states. Again, this matters for assessing autistic subjective reports, and I'll come back to it in section 4.

3.4 Comparing the Reflective-Expressivist and Neoperceptual Models

There is much more to say in defence and elaboration of the RE model of self-awareness and self-report.⁴ For present purposes, my aim is simply

⁴ For further discussion, see Bar-On (2000), Bar-On and Long (2001); McGeer (in preparation).

to articulate a coherent alternative to the NP model that is worth exploring in its own right. In my view, what makes it worth exploring are the following salient points of comparison:

- i. The RE model is more economical than the NP model. Like the NP model, it requires an account of how subjects are linguistically enabled to express the contents of their mental states without forming (higher-order) beliefs about them. The RE model only differs with the NP model over whether it is first- or second-order states that are directly expressed in self-report. However, unlike the NP model, the RE model obviates the positing of a tracking mechanism or process responsible for generating second-order representations of particular first-order states. This economy leads to a straightforward empirical prediction: The search for such a mechanism and/or process in the brain will be forlorn. Correspondingly, at the theoretical level, the very idea that such a mechanism or process could be selectively damaged, either through trauma or in the course of development, will have no place in cognitive theory.
- ii. The RE model rules out a certain kind of error in self-report – namely, error by way of mistaking or misrepresenting one's own conscious occurrent first-order states – for instance, one's own occurrent sensory experiences, as they putatively are in themselves as against how they merely *seem* to a person to be. Experiences – or better *experiencings* – are firmly located back on the subjective side of the subjective-objective divide. This has the further advantage of repudiating the strange metaphysics of 'real seeming', what Daniel Dennett has called 'the bizarre category of the objectively-subjective – the way things actively objectively seem to you even if they don't seem to seem that way to you!' (Dennett, 1991, p. 132).

But is this really such an advantage? After all, one important motivation for retaining the category of the objectively-subjective is precisely to account for certain types of error in self-report – namely, the sorts of error for which we seem to have very good reason to believe that individuals are misreporting (misexpressing) how they are *really* experiencing things. Cases of neglect seem to provide a good example. For instance, in Anton's syndrome, involving visual anosognosia (or blindness neglect), cortically blind subjects fail to acknowledge that they cannot see. Is this not a case of subjects simply mistaking what their own experiences are really like?

In partial response to this challenge, it's important to recognize that the RE model does not rule out *all* sources of error or unreliability in subjective report. Uncontroversial sources of error or unreliability include, in either pathological or normal forms: limitations of language, limitations of memory, self- and other-deception and, of course, inattentiveness or neglect to features of the world, including, of course, features of one's own body. For example, for interesting pathological reasons yet to be fully understood in neurocognitive terms, the RE model suggests that what visual anosognosics are failing to notice is not something about the quality of their own experiences due specifically to a lack of inner-directed attention; rather, what they are failing to notice is the absence of a visible world due to a peculiar loss of outer-directed attention that accompanies their loss of visual perception. In other words, it's not just that visual anosognosics fail to see the world – this is simply blindness; it's that they *overlook* the absence of a visual world. Further, by overlooking the absence of the visual world, visual anosognosics fail to notice a problem with themselves. On this model, blindness neglect, or any pathology of neglect, is not to be viewed in terms of *misperceiving* something inner – how one's sensory experiences really are in themselves (see, for example, Goldman, 1997); rather, neglect (as the terms suggests) is more like *forgetting* something out: a usually temporary and specific *amnesia* concerning how they normally interact with the world (for a similar line of argument, see Dennett, 1991).

- iii. Advocates of the RE model can also agree with neoperceptualists that insofar as ordinary subjects have opinions about this, they may be quite dramatically wrong about what accounts for their experiences being a certain way. For these are questions not about the way their experiences are (i.e., how things are, according to how they experience them); rather, these are questions about why they experience things as they do (i.e., how things work to make it the case that they experience things a certain way). Notoriously, those who reason from the first-person point of view may have pretty wild and woolly views about such things, at least from the point of view of science, if not from the point of view of folksy commonsense. For instance, the phenomenon of *change blindness* seems to indicate that we don't recreate and store somewhere in the brain a fully detailed imagistic representation of the scene before us. And yet, given the rich detail we can perceive in the world (as soon as we

saccade to it), it may seem to us as perceiving subjects that the visual scene is recreated in our experience in all the high-resolution detail that appears before us. This seemingly plausible story is not unmasked by deeper first-person reflection on our experience, but rather by scientists experimenting with our abilities to act and react perceptually under carefully controlled conditions (Dennett, 1992; Churchland et al., 1994; Rensink, 2000).

- iv. The crucial difference between the RE and NP models of self-attribution and self report is finally this: The RE model asserts, whereas the NP model denies, that within the restricted domain of individuals saying how things *seem* to them to be – that is, of their directly expressing how they are *experiencing* things – they cannot, subject to various qualifications, be reasonably doubted. The qualifications are that they are sincere and that uncontroversial sources of error are ruled out: linguistic incompetence, confusion, distraction, neglect and so on.

4 Are There Special Reasons to Doubt the Reliability of Autistic Self-Report?

So far I've argued that there are two background considerations that might lead researchers to underestimate the probable reliability of abnormal experiential or other subjective reports, thereby losing or misreading valuable data in their exploration of these phenomena.

The first is a misplaced adherence to the Humean principle: *the less probable the reported event, the less credible the report*. While this principle is not unreasonable in itself, I think it is misapplied in the case of subjective report (as against testimony to the occurrence of miracles) on two distinct counts: (a) Most importantly, we don't just have a person's verbal report to go on (there are ways of objectively exploring what the person reports); and (b) we really don't know enough at this stage about how the brain works to judge the probability or improbability of strangely abnormal experiences.

The second background consideration involves an adherence to the neoperceptual model of introspection. Researchers who endorse the idea that self-awareness – and, hence, self-report – is mediated by a mechanism for generating metarepresentations of (first-order) mental states and processes will always be confronted by questions of the reliability of this mechanism. Moreover, in cases of abnormal or unusual reports,

the possible breakdown of this mechanism must seem a reasonable hypothesis (particularly when combined with the Humean reasoning reviewed earlier). Nevertheless, while this model of introspection is still very popular in philosophical and psychological circles, it is at least optional. Other models, such as the one I've sketched, show promise of accounting for the phenomena we wish to explain in a more economical and finally less mysterious fashion when it comes to conceptualizing the subject of experience, let alone experiences themselves. For embracing an RE model of self-reflective activity means forgoing the analogy between observing the world and observing (or even tracking) inner conscious states, as if these states can and do exist independently of reflecting, judgemental, verbally adept individuals as they are (somehow) in themselves. The RE model is therefore congenial with a more straightforward way of characterizing how subjects are in themselves: They *are* simply their rich and changing conscious experiences – their consciously attended ways of taking the world. Moreover, when they give expression to these ways of taking the world, they show how it is with them subjectively. Their verbal articulations – ‘reports’, if you like – thus constitute a rich source of *data* for consciousness research, to be explained in terms of what individuals are consciously experiencing, rather than constituting *evidence* for what is actually experienced, to be assessed in terms of how well a given individual is thought to introspect what is putatively an object for them as much as it is for researchers ‘on the outside’.

In this last section, I turn to what may seem to be an additional reason for discounting subjective reports in the particular case of autism. My aim here is partly to demonstrate that particular philosophical views are not innocent in the way they enter into and affect empirical research programs.

This additional reason concerns the empirically demonstrable difficulties that autistics have in understanding *mental* states and properties – perceiving them in others, let alone reasoning about them – and so explaining and/or predicting others’ behaviour by attributing such states. If autistics cannot attribute/perceive/understand the mental states of others, then surely it’s reasonable to expect they’ll have equal difficulty with their own mental lives. As cognitive psychologists Uta Frith and Francesca Happé have recently suggested:

The logical extension of the ToM deficit account of autism is that individuals with autism may know as little about their own minds as about the minds of other

people. This is not to say that these individuals lack mental states, but that in an important sense they are unable to reflect on their mental states. Simply put, they lack the cognitive machinery to represent their thoughts and feelings *as* thoughts and feelings. (Frith and Happé, 1999)

What *kind* of cognitive machinery is this? According to Frith and Happé, it is the kind of cognitive machinery that would allow individuals to reflect on their inner experience. In particular, it is the kind of machinery that generates representations of mental states – hence, second-order representations – that are about first-order states and processes, themselves about physical things in the world. As the authors elaborate:

It seems plausible that the mechanism that keeps (second-order) representations of mental states separate from (first-order) representations of physical states is the same for self and other attribution. Even if the appreciation of others' mental states results in representations that are more error prone than the representations of own mental states, this difference becomes trivial if one is unable to represent mental states at all. (Frith and Happé, 1999, pp. 4–5)

In reviewing the details of Frith and Happé's proposal, it seems clear that they embrace a version of the neoperceptual model of introspective self-awareness where the unspecified cognitive mechanism or process that produces second-order representations of first-order states (whether of self or other) is replaced by the so-called theory of mind (or ToM) module. As they say themselves, '[I]f the mechanism which underlies the computation of mental states is dysfunctional, then self-knowledge is likely to be impaired just as is the knowledge of other minds' (*ibid.*, p. 7). Moreover, since they propose that this mechanism is what allows for the representation of any kind of mental state – sensory, as well as intentional (e.g., beliefs and desires) – we are left with an error theory to account for abnormal – or even absent – sensory reports: It is the result of autistics' failure to form adequate second-order beliefs about their own sensory experience. As Frith and Happé explain:

If low-functioning autistics are unable to reflect on their inner experiences, then they would be unable to develop over time the richly connected semantic and experiential associations which normally pervade our reflective consciousness. Observation by parents suggests that the *awareness* of sensations and experiences may be peculiar in children with autism. Anecdotal reports of abnormal sensory and pain experiences are on occasion quite extreme. . . . Abnormal response to heat and cold, as well as hypo- and hyper-sensitivity to sound, light or touch are frequently reported. . . . Such responses might be expected if there was an inability to reflect on inner experiential states. Of course, normal pain perception is greatly affected by attribution and expectation. *These individuals might feel immediate pain*

in the same way as everyone else [my emphasis], but would not be able to attribute to themselves the emotional significance that normally accompanies pain. This might explain why they do not complain about it. (Frith and Happé, 1999)

Compare this, now, with an autistic subjective report:

My insensitivity to pain was now as good as total . . . nothing hurt at all. And yet I felt – my actual feelings were not shut off – because when I was aware that I had injured myself somewhere, I could *sense something* [my emphasis], a non-pain, which branched out into my body from the place where the injury was. But the fact was, it didn't hurt. (Gerland, 1997, pp. 54, 157)

Clearly, in this autistic individual, there is *some* kind of awareness of the experience she calls 'pain' – a term, understood neutrally here, to designate that experiential state (whatever it may be) produced by bodily injury. What kind of awareness is this? It does not seem to be behavioural (i.e., based on how the person finds herself reacting to injury), since Gerland claims that she 'senses' something radiating out from the point of injury. So it must be some kind of inner experiential awareness. In light of this fact, does it make sense to say that Gerland 'feel[s] immediate pain in the same way as everyone else', but lacks the capacity to introspect, and so to become properly aware of her perfectly ordinary pain experience? The explanation seems a bit forced. At the very least, we need to rule out the more straightforward possibility that her actual 'pain' experience is abnormal – that is, that what's abnormal for her is her direct and immediate experience of bodily injury (this possibility is also suggested by Raffman, 1999).

However, this option is not so easily available to Frith and Happé since it involves significant modification to their account of introspective awareness. If autistic individuals are capable of being *fully, reflectively self-aware* of their sensory states, their relative incompetence with self- and other attribution of *intentional* states (like beliefs and desires) cannot be accounted for in terms of the breakdown of some general mechanism for generating second-order representational states – states that are, in their words, about mental, as opposed to physical, states of affairs. What, then, might Frith and Happé say?

I see two possibilities. The first is to hold onto a neoperceptual model of introspective awareness and multiply cognitive mechanisms for generating second-order representations of first-order states. One of these mechanisms, ToM_{sen}, is responsible for producing second-order representations of sensory-perceptual states and is spared in autism; but the other, ToM_{int}, responsible for producing second-order representations

of intentional states like belief and desire, is absent or malformed. (Note that this means both mechanisms must be operative in normal cognition to explain normal competence in the full range of self-reports.)

There are, however, some disadvantages to taking this option: In the first place, complicating a functional theory by multiplying mechanisms is an expensive step, requiring one of two justifications: either (a) there is no other way to explain the relevant phenomena, or (b) there is some independent evidence providing collateral support for the proposed innovation. I think neither of these conditions is met in this case. A second disadvantage concerns the overall explanatory value of introducing the ToM_{sen} mechanism to account for abnormal sensory/perceptual reports in autism. While such reports are now explained directly in terms of abnormal sensory/perceptual experiences, there is no explanation of how such experiences might relate to the social-cognitive deficits in autism. Indeed, this account would suggest that they are quite unconnected. Of course, it's always possible to explain their comorbidity in terms of diffuse brain abnormalities, where the pattern of spared and affected capacities has a purely neuromechanical explanation. However, it seems unlikely that such unusual and preoccupying sensory experiences would have no impact on autistic cognitive functioning. In any case, and this is a third disadvantage for Frith and Happé, in taking this option they defeat one of their main objectives in making the proposal that autistics have impaired introspective awareness: namely, to strengthen the claim that a so-called theory of mind deficit is in fact the central deficit in autism, capable of accounting for a range of seemingly unrelated – and yet comorbid – autistic abnormalities.

There is a second possibility that Frith and Happé have open to them. They can take the more radical theoretical step of jettisoning the neo-perceptual model of introspective awareness altogether and embracing something like the reflective-expressivist model instead. Why is this recommended?

Notice, first, that establishing the reliability of autistic self-report in the case of their sensory/perceptual experiences is not so difficult for proponents of the RE model: On this approach, if there are no signs of failure of understanding or attention, gross linguistic difficulties, confusion and so forth, then there is no reason to doubt that self-reporting autistics are directly and reliably expressing how they are experiencing things. Looking again at the now quite extensive sample of autistic autobiographical writings, it's hard to find any signs of these sources of error. While there are certainly oddities of style and content, the capacity these individuals

have to write about themselves – and, in particular their sensory experience – in such vivid and articulate detail belies the idea of confusion, linguistic incompetence, lack of understanding or attention and so forth. Hence, their reports should not be written off as unreliable.

If the evidence suggests that autistic self-reports are reliable in the case of sensory-perceptual matters, then adopting the RE model means that there is no need to postulate a *second* mechanism, also putatively operative in normal cognition, for generating second-order beliefs that are exclusively about first-order sensory-perceptual states. Since no such mechanism is required to explain a person's self-reporting competence in general, there is no need to multiply mechanisms to account for why autistics are capable in one domain while failing in another (the domain of self- and other attribution of intentional states). Hence, this option avoids the disadvantage of introducing any further complications into the functional theory.

Nevertheless, there are difficulties for this approach. For instance, we still need to explain the relative *incompetence* that autistics show with intentional mentalistic attribution and reasoning – in particular, as this involves (false) beliefs and (complex) emotional states. How is this hurdle to be overcome?

Recall that on the RE model, competence in intentional self- (and other) attribution does not depend on a capacity to form second-order beliefs about oneself (or others) that are then expressed in intentional reports – reports such as 'I believe that *p*' or 'she believes that *p*'. It does not depend on any sort of mechanism for generating such beliefs. Rather, it depends on knowing how and when the norms governing folk-psychological concepts (belief, desire, hope, fear) apply to our dispositions to interact with the world: For instance, if someone is disposed to unhesitatingly judge that *p*, then, according to the norms of folk-psychology, this amounts to someone's believing that *p*. Hence, if I'm aware of this norm, and find myself or anyone else unhesitatingly disposed to judge that *p* (whether in word or in deed), then I know we each believe that *p*, and can competently attribute this state to either one of us.

Of course, learning and applying such norms takes a lot of practice; they are much more nuanced than this crude sketch implies, and their conditions of correct application depend on many subtle features of the way humans interact with the world, constituting myriad patterns of normalcy and exception. For instance, if someone confidently insists that *p* while breaking into a light sweat and nervously tapping a foot, we tend to think, 'Aha, this person probably does not believe that *p*, but wants me to

believe that *p*' – or even more subtly, 'This person wants me to think that *he* or *she* believes that *p*', or even more subtly still, 'This person wants me to think that he or she is *lying* about *p*', and so on. How many day-to-day interactions with others, both humdrum and exceptional, does it take to acquire this kind of dedicated expertise, in knowing both the norms of folk-psychology and in knowing the ways these norms may be kept and broken in various linguistic and nonlinguistic ways?

Acquiring this knowledge may, of course, be crude and partial, and one's capacity to understand and use folk-psychological attribution and explanation correspondingly circumscribed. Much will depend on one's motivation and opportunity for learning, especially as this capacity begins to develop through childhood. So, for instance, congenitally blind and congenitally deaf children (particularly deaf children of hearing parents) are delayed on standard theory of mind tests, presumably not because they lack some special neuro-cognitive mechanism for generating second-order beliefs about intentional states but because their sensory deficits deprive them of the opportunity of engaging in rich, reason-giving interactions with parents and peers, possibly affecting their motivation for engaging in such interactions as well. Hence, they are slow to develop normal folk-psychological expertise (Hobson, 1993; Brown et al., 1997; Peterson and Siegal, 1998; Peterson and Siegal, 1999; Peterson et al., 2000). Could something like this be true for autistic individuals? Could their very profound sensory-perceptual abnormalities sufficiently inhibit their interactions with others, both through development and in an ongoing way, so as to seriously compromise their ability to acquire folk-psychological expertise? I think it could and argue extensively for this possibility elsewhere (McGeer, 2001).

If this explanatory hurdle can be overcome in the manner just suggested, then embracing the RE model of reflective self-awareness and self-report, far from leaving us with no account of autistic social-cognitive problems, opens up new explanatory possibilities and creates new research opportunities. In particular, in keeping with Frith and Happé's original motivation, it suggests that there *is* an important connection between reported sensory/perceptual abnormalities and social-cognitive problems in autism, thereby promising a unified account of these abnormalities. However, against Frith and Happé's dysfunctional ToM hypothesis, it suggests a different kind of connection and an alternative developmental possibility to what they propose. In particular, it suggests that abnormal sensory *reports* are not the contemporary consequence of a selectively damaged innate capacity for cognizing mind. Rather, it

suggests that deep and extensive sensory/perceptual abnormalities are both developmentally prior to and importantly implicated in the failure to acquire normal social-cognitive expertise. In other words, it is to suggest that research in autism is best approached from a neuro-constructivist perspective. As Karmiloff-Smith writes:

It is clear that disorders like autism and SLI [Specific Language Impairment] have a genetic origin and that evolutionary pressures have contributed to whatever is innately specified. This is a truism. The question is whether, on the one hand, the deficit results from damage to a domain-specific starting point at the cognitive level, as a result of evolution pre-specifying dedicated processing systems for grammar, theory of mind, and so forth, or whether, on the other hand, evolution has specified more general constraints for higher-level cognition and there is a more indirect way for genetic defects to result in domain-specific outcomes as a function of development. (Karmiloff-Smith, 1998)

In support of this broader developmental approach, consider one final passage from Therese Jolliffe's "personal account" of autism:

Reality to an autistic person is a confusing interacting mass of events, people, places, sounds and sights. There seem to be no clear boundaries, order or meaning to anything. A large part of my life is spent just trying to work out the pattern behind everything. . . .

Objects are frightening. Moving objects are harder to cope with because of the added complexity of movement. Moving objects which also make a noise are even harder to cope with because you have to try to take in the sight, movement and further added complexity of the noise. Human beings are the hardest of all to understand because not only do you have to cope with the problem of just seeing them, they move about when you are not expecting them to, they make varying noises and along with this, they place all different kinds of demands on you which are just impossible to understand. . . .

It is the confusion that results from not being able to understand the world around me which I think causes all the fear. This fear then brings a need to withdraw. Anything which helps reduce the confusion . . . has the effect of reducing the fear and ultimately reduces the isolation and despair, thus making life a bit more bearable to live in. If only people could experience what autism is like just for a few minutes, they might then know how to help! (Jolliffe et al., 1992)

Experiencing what autism is like is out of the question for those of us not similarly afflicted. But I do think that paying close attention to autistic self-report can yield substantial insights. In particular, I think it lends credence to the developmental hypothesis I have suggested. But, of course, this is just a hypothesis, requiring further investigation and empirical support. My hope is that once any philosophical biases towards thinking of self-attribution and self-report in broadly perceptual terms

have been removed, this line of research will become tempting for cognitive scientists to explore in detail.

Acknowledgments

This work was supported by the McDonnell Foundation through a grant received from the McDonnell Project in Philosophy and the Neurosciences. I would also like to thank Daniel Dennett, Philip Gerrans, Philip Pettit and especially Kathleen Akins for helpful comments on earlier drafts of this chapter.

References

- Armstrong, D. (1968, 1993). *A Materialist Theory of the Mind*. London, Routledge.
- Bar-On, D. (2000). Speaking my mind. *Philosophical Topics* 28(2): 1–34.
- Bar-On, D., and D. C. Long (2001). Avowals and first-person privilege. *Philosophy and Phenomenological Research* 62(2): 311–335.
- Baron-Cohen, S. (1996). *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA, MIT Press.
- Baron-Cohen, S., H. Tager-Flusberg, and D. J. Cohen, Eds. (1993). *Understanding Other Minds: Perspectives from Autism*. Oxford, Oxford University Press.
- Baron-Cohen, S., H. Tager-Flusberg, and D. J. Cohen, Eds. (2000). *Understanding Other Minds: Perspectives from Developmental Cognitive Neuroscience*. Oxford, Oxford University Press.
- Brown, R., R. P. Hobson, and A. Lee (1997). Are there ‘autistic-like’ features in congenitally blind children? *Journal of Child Psychology and Psychiatry* 38(6): 693–703.
- Churchland, P. S., V. S. Ramachandran, and T. J. Sejnowski (1994). A critique of pure vision. In C. Koch and J. Davis (Eds.), *Large Scale Neuronal Theories of the Brain*. Cambridge, MA, MIT Press: 23–60.
- Dennett, D. C. (1991). *Consciousness Explained*. Boston, MA, Little, Brown and Company.
- Dennett, D. C. (1992). ‘Filling in’ versus finding out: A ubiquitous confusion in cognitive science. In H. L. Pick, P. van den Broek, and D. C. Knill (Eds.), *Cognition: Conceptual and Methodological Issues*. Washington, DC, American Psychological Association: 33–49.
- Frith, U. (1989). *Autism: Explaining the Enigma*. Oxford, Blackwell.
- Frith, U., Ed. (1991). *Asperger and his Syndrome*. Cambridge, Cambridge University Press.
- Frith, U., and F. Happé (1994). Autism: Beyond ‘theory of mind’. *Cognition* 50: 115–132.
- Frith, U., and F. Happé (1999). Theory of mind and self-consciousness: What is it like to be autistic? *Mind and Language* 14(1): 1–22.
- Gerland, G. (1997). *A Real Person: Life on the Outside*. London, Souvenir Press.
- Gerrans, P., and V. McGeer (2003). *Theory of Mind in Autism and Schizophrenia*. Brighton and New York, Psychology Press.

- Goldman, A. I. (1997). Science, publicity and consciousness. *Philosophy of Science* 64: 525-545.
- Goldman, A. I. (2000). Can science know when you're conscious? Epistemological foundations of cognitive research. *Journal of Consciousness Studies* 7(5): 3-22.
- Green, D., G. Baird, A. L. Barnett, L. Henderson, J. Huber, and S. E. Henderson (2002). The severity and nature of motor impairment in Asperger's syndrome: A comparison with Specific Developmental Disorder of Motor Function. *Journal of Child Psychology and Psychiatry and Allied Disciplines* 43(5): 655-668.
- Happé, F. (1996). Studying weak central coherence at low levels: Children with autism do not succumb to visual illusions: A research note. *Journal of Child Psychology and Psychiatry* 37(7): 873-877.
- Happé, F. (2000). Parts and wholes, meanings and minds: Central coherence and its relation to theory of mind. In S. Baron-Cohen, H. Tager-Flusberg, and D. J. Cohen (Eds.), *Understanding Other Minds: Perspectives from Developmental Cognitive Neuroscience*. Oxford, Oxford University Press: 203-221.
- Harman, G. (1990). The Intrinsic Quality of Experience. In J. Tomberlin (Ed.), *Philosophical Perspectives*. Atascadero, CA, Ridgeview 4: 31-52.
- Hobson, R. P. (1993). Understanding persons: The role of affect. In S. Baron-Cohen, H. Tager-Flusberg, and D. J. Cohen (Eds.), *Understanding Other Minds: Perspectives from Autism*. Oxford, Oxford University Press: 204-227.
- Hume, D. (1975). *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*. Oxford, Clarendon Press.
- Jolliffe, T., R. Lansdown, and C. Robinson (1992). Autism: A personal account. *Communication* 26(3): 12-19.
- Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Science* 2(10): 389-398.
- Klin, A., F. R. Volkmar, and S. S. Sparrow, Eds. (2000). *Asperger Syndrome*. New Haven, CT, Yale University Medical Center.
- Mayes, S. D., and S. L. Calhoun (2001). Non-significance of early speech delay in children with autism and normal intelligence and implications for DSM-IV Asperger's disorder. *Autism* 5(1): 81-94.
- Mayes, S. D., S. L. Calhoun, and D. L. Crites (2001). Does DSM-IV Asperger's disorder exist? *Journal of Abnormal Child Psychology* 29(3): 263-271.
- McGeer, V. (2001). Psycho-practice, psycho-theory and the contrastive case of autism. *Journal of Consciousness Studies* 8(5-7): 109-132.
- McGeer, V. (in preparation). *Mind in Time: The Developmental Character of Human Intentionality*, under contract to MIT Press.
- Ozonoff, S., M. South, and J. N. Miller (2000). DSM-IV-defined Asperger syndrome: Cognitive, behavioral and early history differentiation from high-functioning autism. *Autism* 4(1): 29-46.
- Peterson, C. C., J. C. Peterson, and J. Webb (2000). Factors influencing the development of a theory of mind in blind children. *British Journal of Developmental Psychology* 18(3): 431-447.
- Peterson, C. C., and M. Siegal (1998). Changing focus on the representational mind: Concepts of false photos, false drawings and false beliefs in deaf, autistic and normal children. *British Journal of Developmental Psychology* 16: 301-320.

- Raffman, D. (1999). What autism may tell us about self-awareness: A commentary on Frith and Happé. *Mind and Language* 14(1): 23–31.
- Ramachandran, V. S., and E. M. Hubbard (2001a). Psychophysical investigations into the neural basis of synaesthesia. *Proceedings of the Royal Society of London, B* 268: 979–983.
- Ramachandran, V. S., and E. M. Hubbard (2001b). Synaesthesia: A window into perception, thought and language. *Journal of Consciousness Studies* 8(12): 3–34.
- Rensink, R. A. (2000). The dynamic representation of scenes. *Visual Cognition* 7(1–3): 17–42.
- Robbins, T. W. (1997). Integrating the neurological and neuropsychological dimensions of autism. In J. Russell (Ed.), *Autism as an Executive Disorder*. Oxford, Oxford University Press: 21–53.
- Shoemaker, S. (1996). *The First Person Perspective and Other Essays*. Cambridge, Cambridge University Press.
- Tager-Flusberg, H., Ed. (1999). *Neurodevelopmental Disorders*. Boston, MA, MIT Press/Bradford.
- Willey, L. H. (1999). *Pretending to Be Normal: Living with Asperger Syndrome*. London, Jessica Kingsley Publishers.
- Williams, D. (1999). *Autism and Sensing: The Unlost Instinct*. London, Jessica Kingsley Publishers.

Cognition and the Brain

The Philosophy and Neuroscience Movement

Edited by

ANDREW BROOK

Carleton University

KATHLEEN AKINS

Simon Fraser University



CAMBRIDGE
UNIVERSITY PRESS

(SW)
QF260
.5
.C6326
2005

CAMBRIDGE UNIVERSITY PRESS

Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press
40 West 20th Street, New York, NY 10011-4211, USA

www.cambridge.org

Information on this title: www.cambridge.org/9780521836425

© Cambridge University Press 2005

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2005

Printed in the United States of America

A catalog record for this publication is available from the British Library.

Library of Congress Cataloging in Publication Data

Cognition and the brain : the philosophy and neuroscience movement / edited by
Andrew Brook, Kathleen Akins.

p. cm.

Includes bibliographical references and index.

ISBN 0-521-83642-5 (hardback : alk. paper)

1. Cognitive neuroscience – Philosophy. I. Brook, Andrew. II. Akins, Kathleen.

III. Title.

QP360.5.C6326 2005

612.8'233 – dc22 2004022551

ISBN-13 978-0-521-83642-5 hardback

ISBN-10 0-521-83642-5 hardback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party Internet Web sites referred to in this publication and does not guarantee that any content on such Web sites is, or will remain, accurate or appropriate.