

Folk Psychology and the Biological Basis of Intersubjectivity

MATTHEW RATCLIFFE

1. The Usual Story

Recent philosophical discussions of intersubjectivity generally start by stating or assuming that our ability to understand and interact with others is enabled by a 'folk psychology' or 'theory of mind'. Folk psychology is characterized as the ability to attribute intentional states, such as beliefs and desires, to others, in order to predict and explain their behaviour. Many authors claim that this ability is not merely one amongst many constituents of interpersonal understanding but an underlying core that enables social life. For example, Churchland states that folk psychology 'embodies our baseline understanding' of others (1996, p. 3). Currie and Sterelny similarly assert that 'our basic grip on the social world depends on our being able to see our fellows as motivated by beliefs and desires we sometimes share and sometimes do not' (2000, p. 143). And, as Frith and Happé put it, 'this ability appears to be a prerequisite for normal social interaction: in everyday life we make sense of each other's behaviour by appeal to a belief-desire psychology' (1999, p. 2).

As there is general consensus concerning what folk psychology *is*, the focus of recent debates has been on *how* it is accomplished. 'Theory-theorists' claim that the term 'theory of mind' should be taken literally. Attribution of intentional states is enabled by a largely tacit, systematically organized body of knowledge concerning intentional states and their relations. 'Simulation-theorists', in contrast, maintain that our understanding of others depends upon a practical ability as opposed to an organized body of knowledge. Given the plausible assumption that most people have a similar psychological structure, the possibility arises of using one's own mental states and processes as a model for others, predicting what they would do by putting oneself in their situation or psychological predicament. Many different positions are encompassed by these two general approaches and others borrow elements from both¹.

¹ See Carruthers and Smith (eds.) (1996) and Davies and Stone (eds.) (1995a, 1995b) for various theory, simulation and hybrid approaches.

Questions also arise concerning the biological basis, development and evolutionary origin of folk psychology. Both theory and simulation accounts tend to maintain that folk psychology is ‘modular’, meaning that it is tailored to the solution of specific kinds of environmental problems and functions in a manner that is largely autonomous of other cognitive abilities. As Sperber and Wilson put it:

Most theories of mind-reading [...] assume that it is performed not by a general purpose reasoning mechanism, which takes as premises a number of explicit hypotheses about the relationships between behaviour and mental states, but by a dedicated module. (2002, p. 10)

Modular accounts are closely associated with the view that folk psychology is an innate biological adaptation, which arose as a response to selection pressures favouring certain social abilities. As Carruthers puts it, ‘there seems little doubt that our mind-reading (or “theory of mind”) faculty has evolved and been selected for’ (2000, p. 267). Others have suggested that folk psychology, despite having an innate, modular *basis*, also affords various developmental possibilities (Scholl and Leslie, 1999). But commitment to a substantial innate component is not universal. For example, Garfield, Peterson and Perry place more emphasis on developmental processes and claim that folk psychology is supported by an ‘acquired module’ (2000, p. 502), which forms through interaction between innate capacities and social environments. And Gopnik (e.g. 1996) plays down innate abilities even further, suggesting that folk psychology develops in a manner analogous to scientific theories. However, if, as in most accounts, a substantial innate component is postulated, there is the question of which selection pressures influenced its development. The ‘core’ of folk psychology is claimed to involve an ability to recognize that another’s beliefs, desires and intentions differ from one’s own. One context in which such an understanding could be put to use is in manipulating/deceiving others or being receptive to the possibility of their manipulating/deceiving you. Hence folk psychology is complemented by the ‘Machiavellian intelligence’ hypothesis (e.g. Byrne and Whiten, 1988), according to which a primary selection pressure driving human brain development was strategic interaction, with social competition leading to increasingly sophisticated mechanisms for mentalistic interpretation.

So, in summary, although there are many disagreements, almost all accounts accept (1), (2) and (3) below and many also accept (4) and (5):

Folk Psychology and the Biological Basis of Intersubjectivity

1. Social understanding and interaction are enabled by a 'folk psychology', whose 'core' is the ability to attribute intentional states to others in order to predict and explain behaviour.
2. Folk psychological abilities are facilitated by a largely tacit theory, an ability to simulate or a combination of the two.
3. Folk psychology has a modular basis.
4. Folk psychology is largely innate.
5. Folk psychology is an adaptation for strategic social interaction.

My focus here will be on (1). Accounts of how folk psychology is facilitated, how it developed and how it evolved all presuppose an understanding of what folk psychology *is*. Hence, if (1) turns out to be mistaken or substantially incomplete, it is likely (2), (3), (4) and (5) will also have to be abandoned or significantly revised.

Despite near universal acceptance of (1)², it is not at all clear how the claim that 'understanding others consists in the attribution of intentional states in order to predict and explain behaviour' (hereafter FP) is arrived at. FP is, by definition, 'commonsense' psychology. So, although much of its underlying structure may be tacit and only accessible through scientific study, some part of it must be evident in everyday commonsense. However, if one were to ask a variety of people on the street what their understanding of others consists of, it is doubtful that one would get FP or anything like it as a consistent response. Hence FP is something that philosophers and others claim to discover *in* commonsense, rather than something that is readily apparent *to* commonsense. The question thus arises as to whether it is indeed something internal to individuals, which is discovered, or whether it is an external systematisation, imposed by philosophers (Stich and Ravenscroft, 1996). In other words, is FP how we do think about others or is it just a way that some people think about how we think about others? If the latter is so, searching for underlying FP abilities or contemplating their origins and development would be misguided.

I will suggest that, although FP is generally regarded as an ability possessed by individuals, it is actually an abstraction from a much richer context of social understanding and interaction, which has no psychological reality as an autonomous ability. This abstraction has its source in an over-intellectualisation of social life. FP assumes that, in understanding others, we *observe* their behaviour and

² An exception is Gordon (e.g. 1996), who claims that an understanding of intentional states is generated by more fundamental intersubjective abilities.

employ some internal cognitive process in order to *postulate* intentional states as internal regulators of that behaviour. Hobson (1993a) and Gallagher (2001) both point out that the assumptions of detached observation and postulation emphasize the first- to third-person stance, where one looks upon another as a 'he', 'she' or 'it' and contemplates them from a distance. The 'first- to second-person' stance, through which another is encountered as 'you' is very different. In an 'I-you' scenario, one is engaged in complex interactions with others and the cognitive structure of interaction may be very different to that of non-participant observation.

Of course, one might claim that interaction involves exactly the same cognitive processes as observation, regardless of phenomenological differences. However, many practical activities are quite clearly not just a matter of applying internal cognitive abilities in an engaged, rather than observational, context. Clark gives the example of a jigsaw puzzle:

Completing a jigsaw puzzle [...] involves an intricate and iterated dance in which 'pure thought' leads to actions which in turn change or simplify the problems confronting 'pure thought'. (1997, p. 36)

The ability to complete the puzzle is inextricable from one's ability to move and reposition the pieces, to perceive the results and further manipulate them. It is not that some internal capacity is *manifested* through interaction with an environment. The ability to complete the puzzle is indissociable from an ability to interact with and reconfigure the environment. One acts to reshape the environment and receives perceptual feedback, which changes the nature of the problem faced. Now compare completing a jigsaw to interacting with other people. Feedback from others in a social situation is far more complex than that gained from one's manipulation of an inert environment. There is intricate interaction of word, gesture, action, expression, gaze and tone. In what follows, I will argue that such interactions constitute a framework for interpersonal understanding. The ability to understand others is generated through one's interactions with them, rather than through internal capacities that are deployed upon others in contexts of interaction. I will start by focusing on the affective dimension of interaction and suggest that an affective, perceptual, practical grasp of others is central to interpersonal understanding. Furthermore, it does much of the work that postulation of internal states underlying behaviour is claimed to do. This might suggest that FP is an incomplete rather than mistaken account of human social ability. However, I will go

Folk Psychology and the Biological Basis of Intersubjectivity

on to argue that an understanding of intentional states is inextricable from contexts of interaction. Thus FP misinterprets the structure of intersubjectivity, construing an essentially practical, self-engaging process as detached and observational. I will conclude that intersubjectivity is consequently not founded on a domain-specific module in the brain but on a plethora of abilities.

2. Sartre on Intersubjectivity

In order to claim that FP inadequately describes commonsense interpersonal understanding, one cannot appeal to scientific studies of subpersonal processes, given that such things are not part of commonsense. However, one cannot simply ask people on the street either, given that much of commonsense is weakly tacit (by which I mean that it is ordinarily implicit but can, in principle, be made explicit). In my view, a good place to start is phenomenology, given that a goal of phenomenology is to make explicit the ordinarily taken for granted structure of experience. In this section, I will look at Sartre's phenomenological account of intersubjectivity or 'Being-for-others' in his *Being and Nothingness*. In so doing, my aim is not to provide a comprehensive summary or critique of Sartre's position but to draw from his phenomenological descriptions the view that our primary sense of others is perceptual, affective and interactive. Having done so, I will turn to sources other than phenomenology in order to support and elaborate this view.

Sartre employs the example of 'shame' to illustrate how others are encountered as *others*, as opposed to inanimate entities. For example, one can be peeping through a keyhole, spying on someone else's private pastimes (p. 259). One hears a step on the stair behind and is suddenly aware of being looked upon by another. This awareness of another's presence does not take the form of a detached inference but of a self-altering feeling. Sartre describes the phenomenology of shame as follows:

I have just made an awkward or vulgar gesture. This gesture clings to me; I neither judge it nor blame it. I simply live it. [...] But now suddenly I raise my head. Somebody was there and has seen me. Suddenly I realize the vulgarity of my gesture, and I am ashamed. (p. 221)

It is clear that what Sartre calls 'shame' is a kind of affective response. It is 'an immediate shudder which runs through me from head to foot without any discursive preparation' (p. 222).

Registering the presence of another incorporates a change in one's own orientation towards the world, a *feeling* of being scrutinized that breaks up the coherence of one's prior concerns. The project of spying, in which one was previously absorbed, disintegrates. One is no longer a locus of practical projects but an entity that stands before somebody else, an object that is situated in the context of their projects. This affective re-orientation does not just accompany or facilitate the experience of another. It *is* itself one's sense of 'the Other'. Sartre does not construe affective transformation of oneself before the other as a one-off event but as a dynamic process of interaction, whereby self and other engage in a play of mutual objectification. A dance of changing affect first renders one an object before the other and then the other an object before oneself. So the experience of others involves mutual transformation. One is essentially *engaged* in this, rather than standing back as a detached onlooker. And this direct, perceptual, affective apprehension is more basic than any theoretical or detached understanding that one might also employ: 'The Other is present to me without any intermediary as a transcendence which is not mine' (p. 270).

Sartre claims that the feeling of shame has its source in 'the look' of the other, which directly elicits an affective response. By 'the look', he means something more abstract than a pair of eyes gazing at one:

Of course, what *most often* manifests a look is the convergence of two ocular globes in my direction. But the look will be given just as well on occasion where there is a rustling in the branches, or the sound of a footstep followed by silence, or the slight opening of a shutter, or a light movement of a curtain. (p. 257)

Hence 'the look' is not something in the world with a physically identifiable structure. But surely if a 'rustling in the branches' can do it, some kind of cognitive process is required in order to infer the presence of another from an inanimate stimulus? Sartre claims that our most basic sense of the presence of others arises in a context of bodily interaction and it is clear that he thinks this, at least, requires no such inference. The other is directly apprehended, not just through her gaze, but in an affect-laden perception of her dynamic body: 'The Other is originally given to me as a *body in situation*' (p. 344). However, this is not to suggest that a 'mere body' is apprehended and mental states are then postulated as the causes of bodily movements. One perceives the other's body *as* a locus of experiences and projects. This perception is inextricable from an

Folk Psychology and the Biological Basis of Intersubjectivity

affective transformation of one's own body, which one suddenly becomes aware of as a thing before the other. Hence 'the original bond with the Other', on which the objectifying interplay between self and other is based, 'first arises in connection with the relation between my body and the Other's body' (p. 361).

Sartre's depiction of Being-for-others emphasizes an irresolvable tension in interpersonal interaction. Both are locked in an objectifying dynamic; one objectifies the other or is objectified by the other. However, one might suggest that this is at best a description of certain pathological relationships and fails to capture the openness towards another that characterizes a loving relationship, for example. The claim that such relations take the form of one party objectifying the other is, at the very least, phenomenologically implausible. However, I do think that Sartre's discussion makes salient some important characteristics common to the phenomenology of all interpersonal interaction. We can reject the Sartrean emphasis on conflictual relations and build on the following insights into the structure of interpersonal experience and interaction:

1. Our sense of others as *others* is perceptual and phenomenologically direct.
2. We perceive others as animate beings, rather than as moving bodies with underlying mental states.
3. Perception of others incorporates changes in one's self-perception. It is not detached but self-engaging and dynamic.
4. Perception of others is essentially affective. One registers others through the way in which they induce affective changes in oneself.

If the above are indeed aspects of our everyday experience, then an appreciation of others is, in some respects at least, affective, practical, perceptual and direct, as opposed to detached, observational and indirect. Hence, if FP claims to comprehensively describe our everyday apprehension of others as *others*, it misinterprets a practical dynamic in terms of detached observation. However, our FP advocate could reply that FP *underlies* the affective transformations that Sartre describes. X is ashamed before Y because X believes that Y believes that spying is wrong; X also believes that spying is wrong; X believes that Y has seen X and so forth. Hence what looks like affect-laden perception is, upon reflection, supported by abilities to attribute various internal mental states. But the question arises as to why we *should* reinterpret the experience in these terms, even if we *can*. One

reason to resist such a move is that complex affective engagement with others appears to be evident in very young infants, who do not have full-blown FP abilities. Indeed some descriptions of early infant-parent interactions incorporate insights I have drawn from Sartre's description of adult phenomenology, but without Sartre's narrow emphasis on conflict.

3. Affective Interaction

It is well established that very young infants respond to emotional expressions and gestures with attention, gaze, expression and sometimes imitation. Hobson (1993a, b; 2002) appeals to numerous studies of autistic and normal children of various ages to make the stronger claim that infant-parent interactions constitute a kind of proto-dialogue, which is enabled by mutual *perception* of affect in expression, during structured interaction. He suggests that, although young infants are unable to conceptualize intentional states, they display 'capacities to *perceive* a range of overt, bodily-expressed attitudes in other people' (1993a, p. 103). Interactions do not always take the form of simple 'perception-response' exchanges. They can be complex, structured patterns, which involve distinguishable stages such as initiation, mutual orientation, greeting, play dialogue and, finally, affective disengagement (2002, p. 35). Hobson claims that early interaction involves neither 'behaviour-reading' nor 'mind-reading' on the part of the infant. One need not *infer* meaning from behaviour. One apprehends it *in* the behaviour. Infants have 'direct perception of and natural engagement with person-related meanings that are apprehended in the expressions and behaviour of others' (1993a, p. 117). Others' expressions and gestures are *perceived* to be meaningful, through the affective responses they elicit. He argues that early interpersonal abilities do not depend on an infant's internal capacities, operating without the aid of interaction. They involve a mutual receptivity that is partly constituted by and grows through affective, bodily interaction: 'It is not the case that to begin with, behaviour is perceived in a cool, detached way' (1993b, p. 214). The structure of interpersonal understanding is instead a matter of 'relations'. Parent and child together configure a framework for their exchanges, through perception of gesture, expression and affect, and affective, expressive and gestural response. Interaction with others is thus inextricable from an infant's developing ability to understand them.

Folk Psychology and the Biological Basis of Intersubjectivity

Gallagher (2001) takes such developmental claims a step further, arguing that these early abilities are not only developmentally prior to FP but remain the primary means of interpersonal understanding in adults. A perceptual and affective appreciation of others, arising through our interaction with them, is all we require in many social scenarios. There is usually no additional need to posit underlying internal states. Gallagher suggests that most interpersonal understanding incorporates an 'I-you', rather than an 'I-she' structure. It is an 'embodied practice', rather than something that incorporates an objective, detached, intellectualized stance towards them:

... in most intersubjective situations we have a direct, pragmatic understanding of another person's intentions because their intentions are explicitly expressed in their embodied actions. (2001, p. 86)

I think Gallagher is right to emphasize the importance of perceptual-affective factors in adult interpersonal interactions. A good way in which to make something explicit is to look at those cases where it breaks down. And this is, I suggest, the case here. For example, Cole (1998; 2001) addresses the contribution made by facial expression and its perception to interpersonal interaction, by exploring cases where the ability to express oneself facially or to perceive facial expressions is impaired or absent. Personal interactions ordinarily involve an intricate interplay of perception of expression and expressive response. Breakdowns of this interaction are evident in those with various facial problems, such as Möbius syndrome, a form of facial paralysis. As Cole observes, 'those with facial differences describe a loss of social relatedness leading to profound social isolation and to an impoverished sense of self' (2001, p. 478). One's ability to interpret others is substantially diminished by a breakdown of normal interaction. And one's own sense of self is altered by the absence of those reciprocal gestures and expressions by which one is ordinarily affected. Cole describes one subject with facial paralysis as follows:

... without the feedback and reinforcement between people that facial gestures provide, there was little relatedness and engagement. Her loss of facial responsiveness made her feel somehow invalidated at her very core. (1998, p. 10)

Such cases make salient the way in which an interplay of affect and expression structures interpersonal understanding; one's perception of another's expression incorporates an affective

response, which is often manifested in one's own expressions. Others respond to this and so forth. Expressions are not best interpreted from a detached, observational standpoint but through the way in which they are modified in response to one's own expressions and gestures. This interplay constitutes an openness or receptivity to each other; a dynamic framework within which the task of mutual understanding is played out. We do not have to appeal to clinical cases in order to appreciate this. Most people have had the experience of a conversation where one feels detached from the other participant, as though one has somehow failed to make contact. The dance of expression, gesture and eye contact fails to flow, the conversation breaks down and one *feels* a failure to 'connect'. Breakdowns of mutual understanding need not take the form of an inability to infer the relevant intentional states. They more often involve a *feeling* of distance, an absence of the to-and-fro of expression and gesture that constitutes a harmonious back-drop for mutual understanding. Hobson notes that such feelings of detachment can be especially pronounced when interacting with autistic people:

A person can *feel* that there is something missing when relating to someone who is autistic—it is as if one is in the presence of a changeling, someone from a different world—but this escapes the net of scientific methods. (2002, p. 49)

So why not admit that understanding others centrally incorporates perception of expression, gesture and feeling in contexts of interaction? One move would be to suggest that it is simply not possible to *perceive* the meaning of a gesture or expression; one can only perceive surface behaviour. However, this is certainly not an a priori truth and some recent findings in neurophysiology indicate that we may indeed have perceptual access to the meaning of certain expressions and gestures, and, more generally, to the teleological structure of action. I am referring to the discovery of mirror neurons in the mid 1990s. These are cells that discharge when one performs a certain kind of action and also discharge when one observes a conspecific performing a similar action. They were first found in the ventral premotor cortex of monkeys and have since been found in other cortical areas. There is also strong evidence of a more widespread mirror system in humans. In monkeys, mirror neurons are responsive to various kinds of hand actions, such as grasping, holding, tearing or manipulating. They do not discharge when the target object alone is presented or when an action is mimicked in the absence of a target object. Two different classes of

Folk Psychology and the Biological Basis of Intersubjectivity

mirror neuron have been identified. Strictly congruent neurons fire when one performs a specific action and also when a conspecific performs the same action in the same way. However, broadly congruent neurons are not sensitive to the precise manner in which the action is carried out but to the goal of the action. In other words, they are receptive to the teleological structure of action, rather than to similarity of movement. There are also broadly congruent neurons, which are sensitive to sequential actions, such as when an experimenter places food on a tray and when the monkey grabs it³. In humans, there is further evidence for gaze-sensitive mirror neurons (Fogassi and Gallese, 2002), in addition to cells sensitive to gesture and posture (Rotondo and Boker, 2002). Studdert-Kennedy (2002) hypothesizes that we also have a specialized mirror system for perceiving and imitating facial expressions. Hence the human mirror system may turn out to be more complex and differentiated than that of the monkey.

These findings have been taken by some to indicate that we are able to perceive much of the structure of action, rather than infer it from behaviour. In observing X perform action A, parts of one's own motor system are activated in the same way that they would be if one were to perform action A. Perception is structured by a proprioceptive mirroring that facilitates a perceptual awareness of agency:

... when 'reading the mind' of conspecifics whose actions we are observing, we rely *also*, if not mostly, on a series of explicit behavioral signals, that we can detect from their observed behavior. These signals may be *intrinsically meaningful* to the extent that they enable the activation of equivalent inner representations on the observer/mind-attributer's side. (Fogassi and Gallese, 2002, p. 30)

Gallese and Goldman suggest that the existence of a mirror system constitutes support for a simulation theory of FP, given that mirror neurons could be a precursor to and constituent of the ability to get oneself in the same 'mental shoes' as a target (1998, pp. 497–8). However, Gallagher points out that such an interpretation is not supported by the evidence. The motor system is activated *during* perception and 'there is no extra step involved that could count as a simulation routine' (2001, p. 102). One *perceives* the agency of others through a proprioceptive sense of one's own motor readiness,

³ See Fogassi and Gallese (2002) and Rizzolatti, Craighero and Fadiga (2002) for discussion of all these findings.

rather than by perceiving behaviour and *then* employing an ability to simulate in order to interpret that behaviour as action.

Regardless of whether such interpretations of the nature and role of the mirror system receive further experimental support, they remain significant for current purposes, in showing how self-engaging perception of agency, gesture and expression is an empirical possibility and not something to be dismissed on the basis of a priori cognitive implausibility. So it is not at all clear why this phenomenologically pervasive aspect of interpersonal understanding and interaction should be re-interpreted in FP terms. Scrutiny of commonsense suggests otherwise and scientific findings show how what seems to be perceptual, direct and self-engaging might well be precisely that⁴.

If we acknowledge that our understanding of others incorporates a kind of affective responsiveness, which is employed most effectively through contexts of interaction, emphasis on a circumscribed 'intersubjectivity' device or module starts to look misleading. A plethora of variably connected abilities play a role in affective interaction. These include expressing oneself facially and responding to facial expressions, responding to gaze, initiating and responding to gestures, and co-ordinating a variety of affective states. None of these is a peripheral accompaniment to some underlying core and they do not together constitute a single device. We encounter others as whole organisms and with our whole organism. Intersubjectivity is not a single, discrete skill.

One could respond by conceding that FP is not all-encompassing; it is not a magic box from which *all* other interpersonal abilities spring but it does play a substantial role in interpersonal understanding, which cannot be wholly replaced by an account of perception of feeling in contexts of interaction. In the next section, I will argue that this too should be rejected. FP is not a separable component of intersubjectivity but something that has been abstracted from a broader framework of social interaction and misinterpreted.

⁴ It is worth raising a note of caution here. That the mirror system facilitates perception of action does not imply that it also facilitates perception of feeling. For example, Hobson (2002, p. 56) discusses experiments which suggest that autistic people can understand the actions of others whilst failing to comprehend their emotions. However, if we have a proprioceptive awareness of others' action, gesture and expression, it is surely possible, and indeed likely, that this also applies to the perception of affect. After all, the 'mirror system' is unlikely to be a single, unified structure.

4. A Scaffold for Thought

In this section, I will suggest that linguistic exchanges between people are intricately connected with the kind of affective interaction sketched above. The nature and role/s of intentional talk can only be appreciated when considered as part of this rich context of verbal and non-verbal interaction. Thus FP is not a discrete ability but an abstraction from the realities of social life.

Brief reflection is enough to reveal that everyday face-to-face conversations are sculpted by a subtle, harmonious interplay of feeling, gesture, expression and action. For example, if one is met by a smile and or by an extension of one's previous comment, one may develop that comment further. One understands, through her smile, that another has understood and sympathized. The construction of conversational narratives is not simply a matter of having the capacity to infer intentional states and respond accordingly. It incorporates a background of affective and perceptual interaction, which serves to constrain and direct the mutually constructed narrative. The centrality of this kind of interactive narrative construction to interpersonal understanding is made salient by Bruner and Feldman (1993), who argue that the primary deficit involved in autism is not, as is often maintained, an impairment of FP, construed as a detached ability to attribute internal states. It is instead a failure to fully participate in the narratives that are ordinarily formed through interpersonal interaction. In linguistic interactions, autistic subjects fail to extend a previous speaker's comment or grasp 'where it is "going"' and their ability to tell coherent stories is impaired (p. 274). Bruner and Feldman also recognize that affect ordinarily plays a central role in narrative interaction. As they tellingly note:

Although all but one of the subjects manifestly enjoyed having the conversations, the interviewer felt she had failed. In spite of the appearance of so much talk, she nevertheless *felt* that she had been unable to make contact⁵. (p. 277)

The interactive structure of face-to-face linguistic exchanges and their inextricability from an intricate dance of expression and gesture further diminishes the role of FP. In 'I-you' interaction, others are interpreted within a mutually created context, which massively constrains interpretation. Both parties' understanding of each other is progressively shaped and focused by an evolving

⁵ My italics.

narrative, whose development is supported by interaction of gesture, movement, expression and tone. One does not need to predict what another will say, think or do from a neutral, detached perspective or assign internal states by *observing* behaviour. A shared context is constructed through *interaction* and it is within this context that one interprets. Interpretations are supported, shaped or rejected through interaction, through a frown, a laugh or a grimace. In order to understand another, one does not 'read' from afar; one acts, gestures, smiles, speaks and responds. Understanding others is a multi-faceted process, in which others' activities partly constitute the structures through which one understands them⁶. It is not a matter of detached observation of behaviour or deployment of a pre-given internal capacity. One actively investigates with word, gesture and expression. One creates the structures of interpersonal understanding through engagement.

It might be objected that such significant contextual constraints only apply to one-to-one interactions, whereas many other social encounters involve brief exchanges or prediction and explanation on the basis of observation. After all, I have focused throughout on 'I-you' interaction. But we often do understand others by observing them. Surely FP plays a significant role in such cases? Furthermore, what about exchanges of letters and e-mails, or telephone conversations, where much of the structure of face-to-face interaction is absent?

In order to placate such concerns, it should be acknowledged that interactions between two people A and B are not exclusively responsible for constraining their understanding of each other. Interpersonal exchanges do not take place in a desert but within a broader context of shared practices. This is the case from a very early age. For example, Hobson (2002, Chapter 2) notes that early adult-child interactions often involve structured play or games, where interaction is progressively constrained by established patterns of exchange and performance. These patterns are accepted by both as a background for more intricate and complex exchanges.

It is clear that not all such patterns are *created* through the interaction of parent and child. The parent already inhabits an intricate cultural framework of standardized practices and agreed patterns of interaction. Some of these are linguistically expressible and may indeed have been learned via linguistic communication. Others will involve forms of practical know-how that have never

⁶ Morton (2003) discusses at length the extent to which interpersonal understanding involves responding to others in such a way as to make oneself intelligible to them.

Folk Psychology and the Biological Basis of Intersubjectivity

been put into words. All may have originated through affective interaction between people. However, they now form a context within which interpersonal interpretation takes place, a background of institutions, prescriptive narratives, social roles, artefact roles and accepted ways of doing things. So the infant not only helps create standardized patterns of interaction but is also tuned into them through its interactions. This shared cultural context serves to constrain *all* personal interpretation. As Bruner puts it:

... human-beings, in interacting with one another, form a sense of the canonical and ordinary as a background against which to interpret and give narrative meaning to breaches in and deviations from 'normal' states of the human condition. (1990, p. 67)

In most circumstances, much of the interpretive work is done by shared context. It is ordinarily a presupposition of interpretation that we share many of the same practices and 'canonical narratives', which tell us 'what one does', 'what should be done'. 'what is to be done with artefacts of type X', or 'what those with social role Y are expected to do' in given situations. Hence first- to third-person interpretation does not require assigning internal mental states from some detached standpoint. A shared cultural context of established practices can do most of the work. Indeed, explanations are often take forms such as 'If X has social role M and is in Situation A, X will do what one does if one is an M in Situation A'. Such explanations do not involve the assignment of internal states but an understanding of normative practices and prescriptive narratives. Furthermore, one can still rely on perceptual/affective engagement in first- to third-person cases. One need not interact with someone in a complex fashion in order to be moved by her gesture and respond to her expressions.

What about those first- to second-person interactions, such as telephone and e-mail, where one is deprived, to various degrees, of gestural and expressive interaction? These need not be described in terms of an ability to infer internal mental states on the basis of observed evidence. They will still incorporate some degree of affective engagement and when such exchanges involve genuine personal understanding, rather than mere information exchange or pre-scripted professional performances, they tend to rely heavily on already established contexts of interaction. For example, when intricate narratives are constructed via telephone, conversations are ordinarily between friends or those who share a fairly specific set of interests and practices. In such cases, the 'rules of engagement'

come largely pre-formed, so that both parties are already predictable to each other. The same goes for prolonged e-mail exchanges. The lengthy and elaborate first- to second-person exchanges between strangers, which (I'm told) take place in Internet chat rooms, are substantially constrained by shared practices, interests and codes of conduct. They may also involve the development of new practices and rules of engagement, suited to that particular medium of engagement. Furthermore, whether a 'chat room' facilitates fully enriched interpersonal understanding is debatable. Perhaps part of the appeal of such things is the mystery and unpredictability of the respondent.

But when are intentional states assigned? Bruner suggests that FP is employed in a way that is inextricable from participation in practices. Most of the time, people will do what is expected. It is only when things deviate from the norms of shared practices that an FP narrative is constructed:

The function of the story is to find an intentional state that mitigates or at least makes comprehensible a deviation from a canonical cultural pattern. (1990, pp. 49–50).

Hence the assignment of intentional states is not something that we need do all the time in order to explain or predict the internal workings and behaviour of others. It is an occasional activity, whose purpose is to describe a person's unexpected actions in such a way that they make sense, given a shared background of practices and ways of doing things. It is part of a richer context of interpersonal understanding, a story told to make sense of apparently deviant activities by showing or trying to show how they accord with shared cultural frameworks through which we (or others) live⁷.

Bruner's account of FP's role is of course debatable. Indeed, it is arguable that belief-talk plays a multiplicity of different roles in everyday social life. Morton (2003, Chapter 3) goes so far as to suggest that our commonsense psychology does not even incorporate a unitary conception of 'belief'. Different senses of the term may be at play in different contexts. But regardless of precisely how or when such talk is employed, it is clear that intentional state assignment is not an ever-present core underlying all interpersonal understanding and interaction. Even if some

⁷ Goldie makes a complementary point with respect to emotion, observing that our understanding of emotions has a narrative structure; 'our thought and talk of emotions is embedded in an interpretive (and sometimes predictive) narrative which aims to make sense of an aspect of someone's life' (2000, p. 103).

Folk Psychology and the Biological Basis of Intersubjectivity

belief-talk does involve postulating internal intentional states (for some purpose or other, whose nature will need to be clarified), it is clear that the ability to do so is inextricable from a much broader collage of abilities. Assignments are massively constrained by shared cultural context. They are also aided considerably by affective response, and perception of action and gesture, most often a context of interaction. Thus FP, construed as a *discrete* ability to attribute internal states from a detached observational standpoint, is a misleading abstraction from social life. It misrepresents one part of a multi-faceted process, where interaction with others and participation in a culture are constitutive of interpersonal understanding, as a detached, observational, internal ability to assign internal states to others. It then proceeds to claim that this distorted fragment underlies everything else.

5. The False Belief Task in Context

Much of the support for a domain-specific, discrete ability to assign intentional states to others is drawn from the many variations of Wimmer and Perner's (1983) False Belief Task (FBT). These tasks involve setting up a situation and asking a question about it, the answer to which requires a child to recognize that another person possesses a belief that differs from her own (true) belief. FBTs are often taken as evidence for a domain-specific ability, which arises at around four to five years of age and enables the attribution of internal mental states. However, I suggest that they indicate no such thing and cohere equally well, if not better, with the idea that an understanding of intentional states is embedded in broader contexts of interaction.

It is interesting to note that much younger children appear to display a grasp of epistemic differences between people, when in more familiar contexts of interaction. Papafragou (2002) notes that two-year-old infants will modify their requests for a hidden object, depending on whether or not an adult in the room has seen it or not. And Bloom and German observe that three-year-olds often pass more 'pragmatically natural' variants of the FBT, with simpler or more specific questions (2000, p. 27), adding that two-year-olds participate in pretend play, understand pretences, assign goals and imitate intended or completed actions.

If, as such findings suggest, an ability to understand others is bound up with interaction in a shared social context, the FBT is an inappropriate test of that ability. Most variants of the task involve

decontextualized observation and hence strip away structures of interaction that are, I have suggested, constitutive of understanding. Indeed, it is arguable that the very design of the task and the importance ascribed to it simply presupposes that a detached ability to assign intentional states is central to interpersonal understanding. For example, Lord (1993) remarks that autism involves deficits in gaze, gesture, and verbal and nonverbal communication. Thus it does not appear to be a specific cognitive deficit. Despite this, an impaired ability to perform on FBTs is often taken not just to confirm the presence of autism but to exemplify the central underlying deficit (impaired FP). Lord indicates that this may amount to a self-fulfilling prophesy:

It is not surprising that cognitive explanations have seemed so attractive, when what has been studied has for the most part been cognitive tasks that are set up in surroundings quite different from those of naturally occurring, affect-laden settings. (p. 310)

It is by no means clear that the skills measured by the FBT are a fundamental or discrete constituent of interpersonal understanding. For example, Garfield, Peterson and Perry (2001) note that children under four years of age are already 'able to perceive a wide variety of socially meaningful objects and properties in their social environments' (p. 532). They go on to cast more general doubt on the idea of a circumscribed FP by claiming that its acquisition is 'essentially social in character, and [...] the body of knowledge represented by [FP] is inextricably bound up with broader knowledge about persons and their lives' (p. 496). Commonplace interpretations of the FBT simply assume the fundamentality of a decontextualized ability to detect beliefs that differ from one's own. Even if they do succeed in measuring a discrete ability, they do not amount to a case for its primacy. It is for such reasons that Bruner and Feldman refer to passing the test as a 'False Belief Diploma', handed out at some arbitrary 'Graduation Day' during development (p. 269). FBTs may be good *indicators* that a certain developmental stage has been reached but an ability to pass them need not itself be the essence of intersubjectivity. Analogously, a smoke detector may indicate a house fire but a house fire is not itself comprehensively described as 'that which makes a smoke alarm go off'. It is not even clear that such tasks require the same *kind* of cognitive performance as everyday interpersonal understanding and interaction. Furthermore, although the developmental achievement required to pass these tests *can* be described in terms of a distinctive cognitive ability to 'attribute an

Folk Psychology and the Biological Basis of Intersubjectivity

internal mental state that differs from one's own', this by no means implies that it *should* be. As I have already suggested, such descriptions could well involve a misleading abstraction from what is really going on. Feelings and bodily reactions may well play a key role in the child's interpretation of the task, along with numerous other factors. Hence these tasks only constitute evidence for FP if they are interpreted through the lens of FP.

6. Evolution and Intersubjectivity

The account I have proposed suggests that intersubjectivity is enabled by myriad variably interconnected abilities, as opposed to an single 'underlying core'. We are evolved organisms, whose abilities are refined through interactions with our environments, rather than by observing the world from afar. It should come as no surprise that our ability to understand something as complex as each other is bodily, practical and multi-faceted. By implication, an account of the evolution of interpersonal understanding will need to incorporate a multiplicity of factors and stages. A speculative account of human cognitive evolution, proposed by Donald (1991), complements much of what I have said concerning perceptual and bodily interaction. Donald suggests that three major evolutionary accomplishments, which appeared in succession, distinguish us from our closest primate cousins:

1. Mimesis (an ability to re-enact events).
2. Speech (with an emphasis on the ability to construct narratives).
3. The ability to use the environment as an external storage system for symbolic representations.

The first two shifts, he claims, occurred at a biological level and involved genetic changes, whereas the third involved a reconfiguration of the organismic environment, which led to changes in brain development. For the sake of brevity, I will restrict my discussion here to Stage 1. Donald's description of mimesis is similar in many respects to the structure of affective interaction described in Sections 2 and 3. According to Donald, mimesis is an ability to re-enact events and actions. It is more sophisticated than mimicry or imitation, given that mimetic performances are structured by one's goals and intentions, but it can incorporate both. Donald emphasizes the role played by a range of bodily capacities and receptivities in facilitating mimesis: 'Tones of voice,

Matthew Ratcliffe

facial expressions, eye movements, manual signs and gestures, postural attitudes, patterned whole-body movements of various sorts, and long sequences of these elements can express many aspects of the perceived world' (p. 169).

He goes on to argue that mimesis was not just an evolutionary precursor to modern human social abilities. It persists as a framework within which linguistic and other abilities are nested and structured:

No matter how evolved our oral-linguistic culture, and no matter how sophisticated the rich varieties of symbolic material surrounding us, mimetic scenarios still form the expressive heart of human social interchange. (p. 189).

Mimesis, Donald claims, plays a key role in one-to-one social exchanges, in addition to activities such as games, customs and dance. It is an 'integral skill', utilising a number of different biological components and not something that could be subserved by a single, discrete, core biological capacity (p. 186). In suggesting that language is nested within mimetic abilities, Donald emphasizes that mimesis has a complex systematic structure, which includes many of the prerequisites for spoken language, such as 'intentionality, generativity, communicativity, reference, autocueing, and the ability to model an unlimited number of objects' (p. 171).

In endorsing Donald's account of intersubjectivity as a multi-faceted, multi-staged accomplishment, I do not want to dispute the popular claim that Machiavellian intelligence may have played some role in our evolution. However, given that intersubjectivity involves a range of abilities, it is unlikely that selection pressures favouring more refined abilities to interact strategically were the only major factors in play. What's more, an ability to detect deceit in others need not fall back on a detached ability to attribute intentional states. Anomalies in the interplay of expression, gesture and gaze could well contribute to the *feeling* that another is not to be trusted. I also suggest, tentatively, that acknowledgement of the multi-faceted nature of interpersonal understanding and the extent to which it is bodily, perceptual and affective, will serve to better clarify and maybe lessen perceived differences between ourselves and other species⁸. FP suggests that the differences between our social abilities and those of our closest primate relatives are largely

⁸ See, for example, Dunbar (2000) and Whiten (2001) for survey and discussion of evolution of FP and comparisons between our own abilities and those of other primates.

Folk Psychology and the Biological Basis of Intersubjectivity

due to our ability to attribute complex internal mental states, on the basis of behavioural observations. Given that intersubjectivity is practical, affective and perceptual, we should perhaps attend to other individually unremarkable differences, such as a more diverse range of facial expressions and a more refined affective receptivity to action, gesture, expression and tone.

7. Conclusion

In summary, FP is an abstraction from a complex of perceptual, affective, expressive, gestural and linguistic interactions, which are scaffolded by a shared cultural context. When talk of beliefs and desires is considered in context, it is clear that FP does not comprise a discrete ability. Even if and when intentional states are assigned, the ability to do so incorporates frameworks of shared culture and structures of interaction. Taking FP as the 'core' of interpersonal understanding involves extracting an aspect of social interaction from its context, reinterpreting it as an autonomous, decontextualized ability to observe and postulate, and then claiming that this abstraction is in fact the foundation of social life. I can see no rationale for popular descriptions of FP. It is possible that such abstractions serve some conceivable theoretical purpose but they do not reflect the structure of social life.

References

- Baron-Cohen, S., H. Tager-Flusberg and D. J. Cohen (eds.) 1993. *Understanding Other Minds: Perspectives from Autism* (Oxford: Oxford University Press).
- Bloom, P. and T. P. German. 2000. 'Two Reasons to Abandon the False Belief Task as a Test of the Theory of Mind', *Cognition*, 77, 25–31.
- Bruner, J. 1990. *Acts of Meaning* (Cambridge Mass.: Harvard University Press).
- Bruner, J. and C. Feldman. 1993. 'Theories of Mind and the Problem of Autism', in Baron-Cohen, Tager-Flusberg and Cohen (eds.), 267–91.
- Byrne, R. and A. Whiten. (eds.) 1988. *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes and Humans* (Oxford: Oxford University Press).
- Carruthers, P. 2000. 'The Evolution of Consciousness', in Carruthers and Chamberlain (eds.), 254–275.
- Carruthers, P. and P. K. Smith (eds.) 1996. *Theories of Theories of Mind* (Cambridge: Cambridge University Press).

- Carruthers, P. and A. Chamberlain (eds.) 2000. *Evolution and the Human Mind: Modularity, Language and Meta-Cognition* (Cambridge: Cambridge University Press).
- Churchland, P. M. 1996. 'Folk Psychology', in P. M. Churchland and P.S. Churchland. *On the Contrary: Critical Essays 1987–1997* (Cambridge Mass., London: MIT Press).
- Clark, A. 1997. *Being There: Putting Brain, Body and World Together Again* (Cambridge Mass., London: MIT Press).
- Cole, J. 1998. *About Face* (Cambridge Mass., London: MIT Press).
- Cole, J. 2001. 'The Contribution of the Face in the Development of Emotion and Self', in Kaszniak (ed.), 478–82.
- Currie, G. and K. Sterelny. 2000. 'How to think about the Modularity of Mind-Reading. *Philosophical Quarterly*, 50, 143–60.
- Davies, M. and T. Stone (eds.) 1995a. *Mental Simulation: Evaluations and Applications* (Oxford: Blackwell).
- Davies, M. and T. Stone (eds.) 1995b. *Folk Psychology: The Theory of Mind Debate* (Oxford: Blackwell).
- Donald, M. 1991. *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition* (Cambridge Mass.: Harvard University Press).
- Dunbar, R. 2000. 'On the Origin of the Human Mind', in Carruthers and Chamberlain (eds.), 238–53.
- Frith, U. and F. Happé. 1999. 'Theory of Mind and Self-Consciousness: What is it Like to be Autistic?' *Mind and Language*, 14, 1–22.
- Fogassi, L. and V. Gallese. 2002. 'The Neural Correlates of Action Understanding in Non-Human Primates', in Stamenov and Gallese (eds), 13–35.
- Gallagher, S. 2001. 'The Practice of Mind: Theory, Simulation, or Primary Interaction?', in Thompson (ed.), 83–108.
- Gallese, V. and A. Goldman. 1998. 'Mirror Neurons and the Simulation Theory of Mind-Reading', *Trends in Cognitive Sciences*, 2, 493–501.
- Garfield, J. L., C. C. Peterson and T. Perry. 2001. 'Social Cognition, Language Acquisition and the Development of the Theory of Mind', *Mind and Language*, 16, 494–541.
- Goldie, P. 2000. *The Emotions: A Philosophical Exploration* (Oxford: Clarendon Press).
- Gopnik, A. 1996. 'Theories and Modules: Creation Myths, Developmental Realities, and Neurath's Boat', in Carruthers and Smith (eds.), 169–83.
- Gordon, R. 1996. "'Radical" Simulation', in Carruthers and Smith (eds.), 11–21.
- Hobson, P. 1993(a) *Autism and the Development of Mind* (Hove (UK): Lawrence Erlbaum Associates).
- Hobson, P. 1993(b). 'Understanding Persons: The Role of Affect', in Baron-Cohen, Tager-Flusberg and Cohen (eds.), 204–27.
- Hobson, P. 2002. *The Cradle of Thought* (London: Macmillan).
- Kaszniak, A. (ed.) 2001. *Emotions, Qualia and Consciousness* (London: World Scientific).

Folk Psychology and the Biological Basis of Intersubjectivity

- Lord, C. 1993. 'The Complexity of Social Behaviour in Autism', in Baron-Cohen, Tager-Flusberg and Cohen (eds.), 292–316.
- Morton, M. 2003. *The Importance of being Understood: Folk Psychology as Ethics* (London: Routledge).
- Papafragou, A. 2002. 'Mindreading and Verbal Communication', *Mind and Language*, 17, 55–67.
- Rizzolatti, G., L. Craighero and L. Fadiga. 2002. 'The Mirror System in Humans', in Stamenov and Gallese (eds.), 37–9.
- Rotondo, J. L. and S. M. Boker. 2002. 'Behavioral Synchronization in Human Conversational Interaction', in Stamenov and Gallese (eds.), 151–62.
- Sartre, J. P. 1989. *Being and Nothingness* (trans. H. E. Barnes) (London: Routledge).
- Scholl, B. J. and A. M. Leslie. 1999. 'Modularity, Development and "Theory of Mind"', *Mind and Language*, 14, 131–53.
- Sperber, D. and D. Wilson. 2002. 'Pragmatics, Modularity and Mind-reading', *Mind and Language*, 17, 3–23.
- Stamenov, M. I. and V. Gallese eds. 2002. *Mirror Neurons and the Evolution of Brain and Language* (Amsterdam, Philadelphia: John Benjamins).
- Stich, S. and I. Ravenscroft. 1996. 'What Is Folk Psychology?', in Stich. *Deconstructing the Mind* (Oxford: Oxford University Press).
- Studdert-Kennedy, M. 2002. 'Mirror Neurons, Vocal Imitation, and the Evolution of Particulate Speech', in Stamenov and Gallese (eds.), 207–27.
- Thompson, E. ed. 2001. *Between Ourselves: Second-Person Issues in the Study of Consciousness* (Thornton, UK: Imprint Academic).
- Walsh, D. (ed.) 2001. *Naturalism, Evolution and Mind* (Cambridge: Cambridge University Press).
- Wimmer, H. and J. Perner. 1983. 'Beliefs about Beliefs: Representation and Constraining Function of Wrong Beliefs in Young Children's Understanding of Deception', *Cognition*, 13, 103–28.
- Whiten, A. 2001. 'Theory of Mind in Non-Verbal Apes: Conceptual Issues and the Critical Experiments', in Walsh (ed.), 199–223.

