

## Understanding Fictional Minds without Theory of Mind!

Daniel D. Hutto

Palmer's extremely rich and wide ranging target paper defends a bold thesis; that *social* minds exist. Apparently, a social mind is not just a set or collection or conglomerate of individual minds. A group of individuals that just happen to think the same thoughts, even all at the same time and in perfect synchrony – whether by accident or design – would presumably not qualify as a social mind. To be a social mind literally (not metaphorically) requires forming a more cohesive sort of unity.

There's a lot that could be said about the very idea of a social mind, so conceived. Palmer highlights the proposal that minds sometimes extend. Defenders of this view try to show that when the necessary completion of some task requires cognitive processing involving the manipulation of external resources, there is no good reason to think that the boundary of mind does not include these external resources. If so, what should we say about cases in which the external resources include other minds – for example, as when couples rely on one another to complete trains of thought? Such activity surely requires forming a more cohesive union than simply co-cognizing with another – i.e. thinking along the same lines as them or thinking thoughts with the same content as theirs. But, unless one also tries to include ordinary conversations under this heading (which would be a stretch at best) it is not clear how widespread such forms of coupling are – either in fiction or 'real' life. And although Palmer does not claim that all social minds are of a piece, it is worth observing that his detailed analysis of the so-called Middlemarch mind reveals that it is not a case of a social mind of the sort just described. Indeed in stressing its norm-engendering role when reflecting that society's view of things, the latter looks as if it might be better understood as a kind of cognitive niche that reliably shapes individual minds in its purview rather than as literally being a mind with thoughts of its own.

In short, while I find Palmer's suggestion intriguing – and I believe I have some handle on what is required to qualify as a social mind – it is clear that the exact nature of such minds (if any exist) and the requirements for being one needs further and more detailed explication. This is not to say that exploring this conceptual terrain is not a worthy occupation: quite the opposite.

I say no more about this in what follows, as I wish to concentrate on some other important claims made in the target essay. In arguing for his main thesis Palmer tells us that 'social minds exist in storyworlds because they exist in real life'. This conclusion rests on the premise that fictional minds (whatever their properties) are – in general – analogous to actual minds. For if we also assume that *at least some* actual minds are social minds then we get Palmer's desired result.

Are fictional minds – in general – analogous to actual minds? This would follow if the former are modeled on the latter. It is hardly contentious that they *sometimes* are. Indeed, it should not be contentious that they often are – but, always and everywhere? This is an exciting question, one that is already exercising many in the fields of aesthetics and narratology. Art often mimics life, but must it? This is a view that Palmer appears to endorse in a sensible variant. However fantastical they become – whatever rules they attempt to bend or break, whatever their experimental character – even the most outlandish or unnatural narratives must start from and build upon familiar resources. Our concepts may indeed be flexible and fuzzy – including our concept of mind – but they are not *completely* flexible and fuzzy. Stretched beyond certain limits

they breakdown, fail to make sense, become incapable of designating anything recognizable at all. Exceptions need rules. As interesting as this debate is – and *it is* – Palmer only needs the much weaker claim for his purposes: fictional minds are modeled on actual minds, *at least sometimes*.

This puts us in a position to consider (1) the agenda and (2) other background assumptions behind Palmer's thesis. He aims to initiate a new approach to studying literary fiction – one that puts social minds, and an examination of the means by which we make sense of them, centre stage. This is because he is of the view that 'fictional narrative is, in essence, the presentation of mental functioning'. Plots necessarily revolve around minds; storied happenings only matter in so far as they matter to the minds of characters. Moreover, 'We follow the plot by following the workings of fictional minds'.<sup>1</sup> Even more extravagantly it is said that 'To make sense of a text the reader has to collect together all of the isolated references to a specific proper name and *construct a consciousness* that continues in the spaces between all the mentions of the character with that name'.

These are very strong claims – and are sure to raise some hackles. For myself, I am not convinced that *all* narratives must necessarily be about or involve people or their surrogates. I am not convinced that it is *always* the case that an understanding of how events unfold in a fictional world *must* revolve around or refer to – if only tacitly – characters' thoughts or feelings. And, I am hugely sceptical of the idea that in order to enjoy fiction we *must* be, quite literally, imaginatively generating storyworlds and populating them in the way alluded to above. It strikes me that Palmer can, and should, support his main thesis with weaker and more modest claims of a far less controversial nature. For it is difficult to deny that *a large and important class of* fictional narratives are primarily concerned with the presentation of minds or that following the workings of fictional minds *helps us* to follow plots; and so on. These truths are enough to justify looking at how we make sense of actual minds in ordinary settings in order to understand how we make sense of fictional ones. These truths suffice to support the idea that when dealing with certain kinds of narratives, 'like it or not', consumers of fiction will bring the same sorts of skills (or at least a subset of them) to bear that they use when dealing with actual minds. Let us call this the 'Same Resources Thesis'.

I believe the 'Same Resources Thesis' is true. But this is because I defend the view that engaging in narrative practices is the normal developmental route through which children acquire the capacity to make sense of what it is to act for a reason. If so, narratives are what *provide crucial resources* for dealing with actual minds – at least those of a certain sophisticated sort. This idea challenges the familiar view that we inherit our 'mindreading' or 'theory of mind' capacities from our evolutionary ancestors – i.e. that such resources emerged to enable our non-linguistic forebears to deal with other minds in much more primitive settings. Palmer apparently endorses this view when he approvingly cites Carruthers' supposition that we evolved our theory of mind capacities in order to keep tabs on what others think, want and feel.

Without refinements, these claims are simply not compatible. It cannot both be the case that engaging in narrative practices provides the key resources for making sense of and attributing reasons and that it is also true that we inherited these *very same* capacities from our hominid ancestors. Along with many others, I have been at pains to argue that what we inherited from our evolutionary forebears cannot be properly characterized as theory of mind or mindreading abilities. Rather we come factory-equipped with capacities to respond to and keep track of basic states of mind in more embodied ways. Why don't such basic capacities for attending to and minding minds count as mindreading capacities?

Mindreading, like theory of mind, is a much used and abused term. If we are doing serious analysis of its role in social cognition it pays to be very clear about exactly what it picks out and what it doesn't. Mindreading is sometimes used by researchers to designate our social cognitive abilities, whatever characteristics they may have. Operating with such a broad and inclusive understanding of mindreading undermines the possibility of disagreement about what making sense of others involves. In the literature that seeks to evaluate such proposals the label is therefore used restrictively and precisely to denote a particular kind of activity – i.e. the attribution of mental states and their contents. With this use in play, interesting explanatory proposals have been advanced about the means by which mental states are ascribed. Hence the term mindreading is neutral between Theory and Simulation-based proposals – see, e.g. Nichols and Stich 2003. Both parties in this on-going debate assume that we mindread, they disagree about how we do it.

Sceptics about the pervasiveness of mindreading doubt that all forms of attending to and keeping track of another's mental states require the use of mindreading capacities understood in the restrictive sense just described. They doubt that our basic ways of engaging with other minds requires making any conceptually-based mentalistic attributions at any level at all. Neither the attribution of mental state concepts nor the attribution of mental state contents plays any part in basic ways of responding to and keeping track of others' psychological attitudes. If this is correct then there are embodied and enactive ways of relating to others and attending to their states of mind that do not constitute acts of mindreading for the simple reason that they do not involve making mentalistic attributions.

Of course, and this complicates matters, this proposal about the non-mindreading ways that we relate to other minds in basic engagements is entirely compatible with the possibility that those who are capable of making mental attributions (e.g. mature users of folk psychology) can make use of the basic mind minding abilities as a reliable (enough) heuristic for assigning mental states or contents to others; at least those of a very basic sort.

Against this backdrop there is a lively, on-going debate in the field of social cognition about whether we inherit any mindreading abilities or not (witness the spate of special issues reviewing the existing empirical evidence and arguments on both sides). It might be thought that one could decline from taking an interest in this debate and opt instead to employ the label theory of mind as a generic term for *whatever* resources we use when making sense of others. At one point Palmer proposes we do just this, maintaining that we really have no choice but to embrace the theory of mind label as it is 'here to stay'. I take the opposite view. The theory of mind label is treacherous. This is because it is not *just* a label but a characterization, one with robust associations and connotations that spawns confusion. Its continued use (1) interferes with our understanding of the phenomenon to be explained and (2) unwarrantedly privileges one possible explanation of that phenomenon above others. Therefore, whatever practical problems we may face in revising our language surrounding this topic, for scientific and philosophical purposes, such reform is non-negotiable.

Witness Palmer's account of what he takes theory of mind to denote by those working in the field. On the one hand, it is said to be equivalent to 'folk psychology' and 'intersubjectivity' being used 'to *describe* our awareness of the existence of other minds ... our ability to make sense of other people's actions by understanding the reasons for those actions'. The first thing to note is that the term is used here to *describe* not merely to label. Moreover, possessing intersubjective capacities and 'being aware of the existence of other minds' are quite different things from being folk psychologically competent and having the 'ability to make sense of

actions in terms of reasons'. Immediately, without refinement, the theory of mind label used just as a descriptor, is bound to cause referential mischief leading to bouts of equivocation. Evidence of this comes from Palmer's very next sentence, in which he switches over to its more common use – i.e. to denote a substantial explanatory hypothesis. For, we are told that 'we are able to make attributions of others *because* we have a theory of mind'. Here the theory of mind label is surely no neutral tag; it designates a much-contested claim, one that requires serious defense. And it appears in this guise often in Palmer's essay which contains many remarks such as the following: 'A minimal level of mindreading and theory of mind is required for the characters to understand each other and makes everyday life possible' (p. 14).

I don't have the space to rehearse my full set of arguments to show why we don't *ever* use a theory of mind (for details of other relevant publications interested readers should consult Hutto 2004, 2005, 2007a, 2007b, 2008a, 2008b, 2008c, 2008d, 2009a, 2009b, 2009c, forthcoming-a, forthcoming-b). But I should say something about Palmer's remark that I endorse the view that we do use theory of mind (understood in the 'narrow' sense) when faced with some puzzling third personal cases. I don't say this, but I understand why I might seem to.

We can surely invoke theory to help when speculating about someone's reasons for acting but this is not to use a theory *of mind*.

I accept that those of us who are folk psychologically competent (unlike members of other species; those under the age of three years old; those with severe ASD, etc.) make mental state attributions. We do so frequently, using different heuristics – these include (1) asking others directly – i.e. the 'horse's mouth principle'; (2) calling on knowledge or theory about the general tendencies of humans, or with more grist – of specific groups of humans, of individuals, etc.; (3) putting imaginative or simulative routines into play; and (4) employing other more unreliable means too – e.g. astrology. Anyone capable of making use of these heuristics in order to attribute mental states in speculative situations must *already* understand how such mental states can inter-relate. But 'theory of mind', as it is used narrowly in the philosophical literature, is meant to denote just that sort of structured understanding or its equivalent – i.e. whatever plays the role of enabling us to understand how mental states inter-relate.

I've argued at length in my publications that having that ability – which I call being folk psychologically competent – does not imply or require the existence of any mindreading or theory of mind mechanisms. But it is even more important to note that the heuristics for attributing mental states and their contents must be employed against the backdrop of, and in addition to, whatever accounts for our core folk psychological competence. This is accepted on all sides, as becomes clear when we consider the nature of the possible proposed alliances between Theory Theorists and Simulationists when they offer hybrid accounts: If theory is thought to be required for our core folk psychological competence then simulation helps with the attribution heuristics, or vice versa. This shows the need to distinguish: A. core folk psychological competence; and B. the additional heuristics that can be used when applying it in certain contexts.

When we are thinking about sophisticated cases it is crucial to note that any interesting case of reason attribution requires going beyond what can be gleaned from another's outward expression. Thus it goes against the grain to think of the attribution heuristics mentioned above as being in the service of fast, efficient and reliable mindreading. When, in ordinary parlance, we talk about reading another's mind 'like a book' we typically imply that the other is somehow transparent to us (as they often are). Talk of mindreading is therefore most appropriate in

precisely those cases in which we *don't* have to guess, speculate, or even to ask the other what they are thinking or feeling. Compare these two examples from literary fiction:

My father spoke more through his gestures and facial expressions than he did with his mouth. I knew when not to disturb him and when it was ok to ask a question ... I knew if he was happy, sad, hot, cold. I knew when he had a headache – I recognized the headache frown. I could even pinpoint whereabouts, more frequently it clustered in the middle of his forehead and stretched out to his temples, but on the odd occasion I could see a heavy ache creep up from the base of his skull and spread out. A migraine. And I even knew how long it would be before he had to lie down and close his eyes.

Yes, if there was any slight discomfort, slight frustration, the slightest beginning of an emotion on my father's face, I could read it. But could I guess correctly what he was writing about? (Mackie 2010, p. 271-2).

My father took this as a sort of sign of nerves and whispered, 'It'll be alright, son,' in my ear. 'I feel sick,' I told him.

My father was clueless. I wanted to turn and pinch him hard on the leg. 'Wake up, for fuck sake, wake up! *You think you can read my mind?* You think I'm just nervous about meeting some crumby old teacher? Why don't you ask me what's wrong? And I'll tell you I've just seen Nigel Kerr almost blow his head off (Mackie 2010, p. 185, emphasis added).

The first case is a wonderful description of psychological ascription based on attending to another's expression. The second case is a fine reminder of why understanding a person's larger narrative is needed to understand their reasons; i.e. of why we can't read reasons off facial expressions or other behaviours except in standard contexts where such questions don't arise (e.g. Why did the waiter take our order?). With this in mind, I think we should reserve the label mindreading for cases in which we make mental attributions based on our fairly reliable access to another's state of mind by means of their expressions; i.e. those cases in which theoretical or speculative inference is *not* required. We could, without prejudice, call other cases – those in which we are forced to speculate or theorize about the state or contents of another's mind – cases of mind guessing. So understood, to the extent that we mindread at all, it is likely that we – i.e. those with the appropriate linguistically scaffolded abilities to make mental attributions – rely on our basic mind minding capacities to do so. So theory only comes into play when we mind guess, but theory *of mind* doesn't come into it at all, neither when we deal with actual or fictional minds.

## References

- Hutto DD. 2004. The Limits of Spectatorial Folk Psychology. *Mind and Language* 19:548-73
- Hutto DD. 2005. Starting Without Theory: Confronting the Paradox of Conceptual Development. In *Other Minds: How Humans Bridge the Divide Between Self and Others*, ed. B Malle, S Hodges. New York: The Guilford Press
- Hutto DD. 2007a. The Narrative Practice Hypothesis: Origins and Applications of Folk Psychology. *Narrative and Understanding Persons*. Royal Institute of Philosophy Supplement 82:43-68
- Hutto DD. 2007b. Folk Psychology without Theory or Simulation. In *Folk Psychology Reassessed*, ed. DD Hutto, M Ratcliffe, pp. 115-35. Dordrecht: Springer

- Hutto DD. 2008a. First Communion: Mimetic Sharing without Theory of Mind. In *The Shared Mind: Perspectives on Intersubjectivity*, ed. J Zlatev, T Racine, C Sinha, E Itkonen, pp. 17-38. Amsterdam/Philadelphia: John Benjamins
- Hutto DD. 2008b. *Folk Psychological Narratives: The Socio-Cultural Basis of Understanding Reasons*. Cambridge, MA: MIT Press
- Hutto DD. 2008c. Limited Engagements and Narrative Extensions. *International Journal of Philosophical Studies*. 16:419-44
- Hutto DD. 2008d. The Narrative Practice Hypothesis: Clarifications and Consequences. *Philosophical Explorations* 11:175-92
- Hutto DD. 2009a. Folk Psychology as Narrative Practice. *Journal of Consciousness Studies*. 16:9-39
- Hutto DD. 2009b. Lessons from Wittgenstein: Elucidating Folk Psychology. *New Ideas in Psychology*. Special Issue: Mind, Meaning and Language: Wittgenstein's Relevance for Psychology 27:197-212
- Hutto DD. 2009c. ToM Rules, but it is not OK. In *Against Theory of Mind*, ed. A Costall, I Leudar, pp. 221-38. Basingstoke: Palgrave
- Hutto DD. forthcoming-a. Presumptuous Naturalism: A Cautionary Tale. *American Philosophical Quarterly*. To be translated into French and reprinted in a special issue of *Recherches sur la Philosophie et le Langage*.
- Hutto DD. forthcoming-b. Elementary Mind Minding, Enactivist-style In *Joint Attention: New Developments in Philosophy, Psychology, and Neuroscience*. Seemann, A. (ed). Cambridge, MA: MIT Press.
- Mackie E. 2010. *And This is True*. London: Hodder and Stoughton
- Nichols S, Stich S. 2003. *Mindreading: An Integrated Account of Pretence, Self-Awareness and Understanding of Other Minds*. Oxford: Oxford University Press
- Palmer, A. *Social Minds*. Style (this issue).

## Endnotes

<sup>1</sup> Here Palmer is building on and adding to what is known as 'the cognitivist' tradition (a label with unfortunate connotations for his project especially as he explicitly rejects a number of the individualist assumptions associated with standard versions of cognitivism).