

Why pretend?

Peter Carruthers

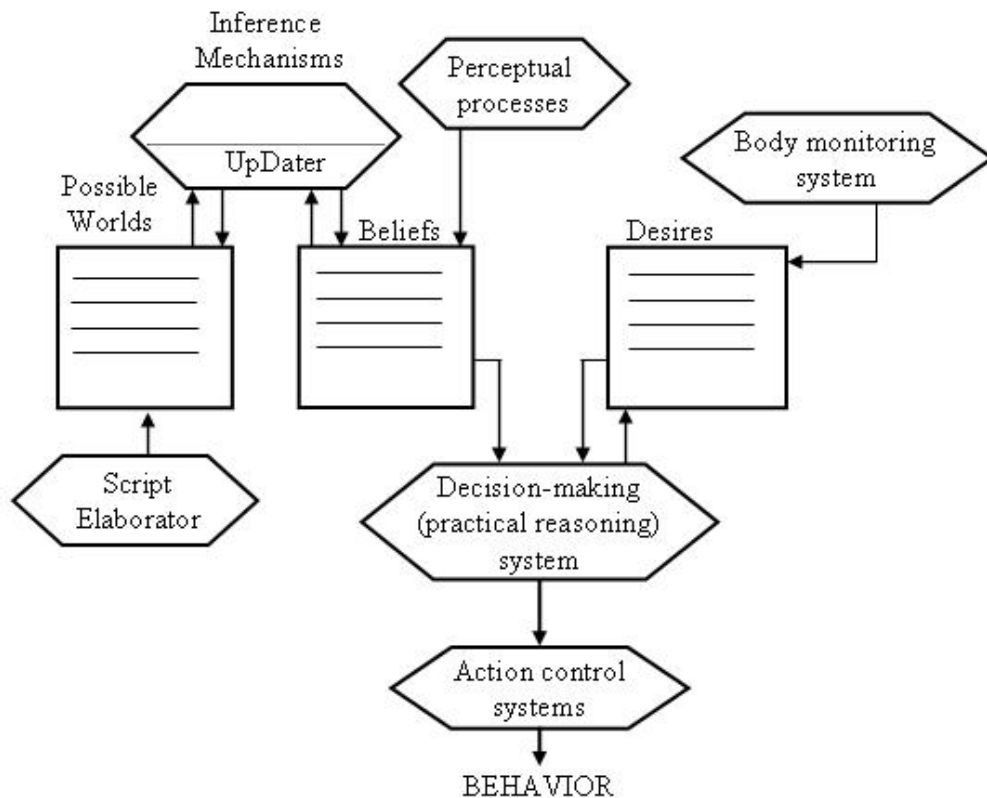
In this chapter I shall take up the question of children's motivations to engage in pretence.¹ I shall use the account provided by Nichols and Stich (2003) as a stalking horse: arguing that they are correct about much of the basic cognitive architecture necessary to explain pretence, but wrong on the question of motivation. Following a discussion of the views of Currie and Ravenscroft (2002) on this issue, I shall draw on Damasio's (1994) description of the way in which emotions enter into practical reasoning involving mental rehearsal. This will enable me to defend a novel explanation of the motivations underlying pretence.

1 Nichols and Stich on the cognitive architecture of pretence

Nichols and Stich (2003 – hereafter N&S) argue that a number of additions need to be made to the standard cognitive architecture of belief, desire, and practical reason, in order to explain the phenomenon of pretence. Most importantly, they argue that we need to recognize the existence of a distinct type of attitude, alongside belief and desire – namely the attitude of *supposing*. When children pretend, they are *supposing* that something is the case (e.g. that the banana is a telephone, that they have an invisible friend called 'Wendy', etc.), and they act out their pretence within the scope of that supposition. Moreover, *supposing* can't be reduced to believing, nor to desiring, nor to any combination thereof (nor can it be reduced to any sort of planning or intending). It therefore needs to be assigned its own 'box' within a functional boxology of the human mind.

N&S suggest, in fact, that *two* new boxes need to be added to the standard architecture (see Figure 1). One is a mechanism for generating novel suppositions. The other is a working-

¹ The question that forms my title is ambiguous: it can be taken in either a proximate or a distal sense. When we ask why children engage in pretence, this can be taken as a question about children's desires and motivations (proximate), or it can be taken as a question about the evolution and adaptive function of pretence (distal). Both questions are interesting and important. But I focus almost entirely on the former (proximate) question here. For discussion of the evolutionary function of pretence, see Carruthers (2002 and forthcoming).

Figure 1: Nichols and Stich's account of pretence

memory system within which those suppositions get elaborated. (These elaborations take place partly by inference, partly by drawing on those of the subject's beliefs that aren't inconsistent with the initial suppositions.) They label the former the 'script elaborator'. But this is misleading, since the mechanism they envisage doesn't just add suppositions to an on-going supposed scenario, but is also meant to *initiate* episodes of supposing. I shall therefore refer to this mechanism in my own discussion as the 'supposition generator'. The working memory system they label the 'possible worlds box', since what it does is contain a partial representation of the possible world in which the initial suppositions are true (as elaborated by inference and by drawing on the subject's existing beliefs).

The possible worlds box, in N&S's model, interacts with a set of inferential mechanisms. These are the very same mechanisms that interact with existing beliefs to generate novel beliefs. Amongst these mechanisms is what N&S call the 'updater'. This is normally employed to update our beliefs as our circumstances change, and as we acquire novel beliefs, so far as possible eradicating inconsistencies that may have been introduced by the adoption of new beliefs. But

the updater is also employed to ‘screen out’ beliefs that are inconsistent with the current contents of the possible worlds box, preventing them from being added to it.

N&S’s account of what happens cognitively during an episode of pretence, then, is this. An initial supposition is placed in the possible worlds box by the supposition generator. That supposition is then worked on by the subject’s inferential mechanisms, drawing on the subject’s existing beliefs to elaborate and extend the initial supposition. But at the same time the updater screens out any of the subject’s existing beliefs, or their consequences, that are inconsistent with the initial supposition, in such a way that the possible world representation being elaborated in the possible worlds box remains internally consistent. At any time, moreover, yet further suppositions can be added to the possible worlds box by the supposition generator, resulting in a further round of inference and elaboration.

All of this is very plausible, and I propose to accept it for the purposes of the discussion that follows. (For a discussion of some different ways in which N&S’s architecture might be implemented, see Carruthers, forthcoming.) I shall be focusing on, and criticizing, what N&S have to say about the way in which the possible worlds system interacts with the desires box. And in the present context, that means focusing on N&S’s account of the motivations that underlie pretend play.

2 Nichols and Stich on the motivation to pretend

N&S’s idea is this: what the pretender wants is to behave in the way that some person or thing represented in their possible worlds box would behave. Put differently, someone who pretends that *P* does so because he wants to behave more or less as he would (or might) behave if *P* were the case. And they postulate that such desires are both *basic* (that is, intrinsic, or non-instrumental) and *innate*, in much the same way that desires for food or sex are basic and innate.²

² Nichols (in correspondence) distinguishes between the *general* motivational question (‘What motivates people to engage in pretence at all?’) and the *specific* motivational question (‘What motivates someone to engage in some particular pretend behavior?’), and claims that the N&S account was intended only to address the former. On this reading, the discussion that follows in this section is about the different ways in which the N&S account might be extended to answer the specific motivational question. My own view, in contrast (to be elaborated in Section 5) is that it isn’t anything motivational that is innate (as N&S claim), but rather a species-specific disposition to mentally rehearse supposed actions. These rehearsals then interact with our motivational systems, giving rise to new desires in the way that mental rehearsals normally do.

These claims have a certain initial plausibility. For it seems quite natural to say of the child pretending that the banana is a telephone that she is acting as she does because she wants to behave in the way that she *would* behave if the banana *were* a telephone (e.g. by making a telephone call to grandma). And likewise it is natural to say that the child who is pretending to be a steam train is acting as he does because he wants to behave in the way that he *would* behave if he *were* a steam train (e.g. moving his arms in circles and going, ‘Chuff, chuff, choo, choo’). Moreover, episodes of pretence are characteristically engaged in for their own sake – or ‘for fun’ – and pretend play is a universal species-specific behavior. So the desires that serve to motivate pretence must normally be intrinsic, rather than instrumental, even if they aren’t outright innate.³

In light of the above, it is plain that N&S need to add a further arrow to their box-and-arrow diagram of the mechanisms involved in pretence. Either they need an arrow from the ‘script elaborator’ / ‘supposition generator’ to the desires box; or they need an arrow from the possible worlds box to the desires box. And the former of these suggestions seems to be the most plausible. For otherwise they would need to provide some account of *which* representations in the possible worlds box are used to generate a corresponding desire. On this interpretation, then, the idea is that whenever a novel supposition is generated – that the banana is a telephone; that I am a steam train; that I have an invisible friend called ‘Wendy’; and so on – this causes a novel desire to come into existence; the desire, namely, to behave as one would behave if the supposition were true. This would happen at the outset of an episode of pretence, and would also occur whenever supplementary suppositions are added during the course of the pretence. Each such supposition would give rise to a novel desire: the desire to behave as one would behave if that supposition were true.

2.1 *Why don’t all suppositions give rise to desire?*

But now we (or rather, N&S) have a problem. For as they themselves acknowledge, the supposition generator and possible worlds box are *also* employed for purposes of hypothetical and counter-factual reasoning quite generally, and not just for purposes of pretence. And it just

³ But those desires needn’t *always* be intrinsic, of course. A child taking part in a psychological experiment might pretend to be a steam train, not for its own sake, but in order to comply with a request from the experimenter. And likewise a child might engage in a novel pretence (like pretending to be a dead cat), not for its own sake, but in order to make his parents laugh.

isn't the case – obviously – that I have a desire to behave in the way that I would if my supposition were true *whenever* I make a novel supposition. In the course of a discussion of the abortion issue with students, for example, I might say, 'Suppose that *Roe versus Wade* were overturned'. This surely doesn't cause in me any desire to behave as I would behave if *Roe versus Wade* were overturned (e.g. carrying a placard in front of the Capitol). Likewise, when considering potential house repairs I might think, 'Suppose we had the roof re-done this year, and then had the outside painted next year'. But this surely doesn't cause in me any desire to behave as I would behave if we did have the roof renewed this year (e.g. visiting the bank for a loan).

The obvious suggestion to make on N&S's behalf is that the arrow that we have said needs to be added between the supposition generator and the desires box is there to represent a *channel* for information, which can either be open or closed. When it is open, each novel supposition gives rise to a corresponding desire; but when it is closed, no new desires get created. And the proposal could be that episodes of pretending occur whenever a novel supposition is made while the channel between the supposition generator and desires box is open; whereas in all other forms of suppositional thinking and reasoning, that channel is closed. The problem now, however, is to give some non-arbitrary account of what might cause the opening or closing of the postulated channel.

One proposal might be that the status of the channel (as open or closed, that is) is age-specific. Perhaps it is open in childhood and closed in adulthood. But this is highly implausible. Children are perfectly well capable of entertaining novel suppositions when engaging in hypothetical or counter-factual reasoning (Harris, 2000), without thereby being motivated to engage in some sort of pretence. And likewise adults are perfectly well capable of pretending – and not just on instrumental grounds, either, but for its own sake. (Think here of a mother playing at dolls with her daughter.)

Another proposal might be that the channel between the supposition generator and the desires box only gets opened by specific types of behavioral cue, such as someone giving expression to the supposition in question with a characteristic 'let's pretend' voice intonation. For there does at least seem to *be* a characteristic tone of voice that is used to initiate games of pretence (Harris, 2000), somewhat as the play-bows of dogs and wolves are used to initiate games of chase (Bekoff and Allen, 1998). And it might be suggested that we are hard-wired to

respond to such a tone of voice by creating an appropriate motivation. Moreover (and in response to the obvious objection that people frequently engage in solitary forms of pretence), it might be suggested that imagined utterances in ‘inner speech’ with the same sort of (imagined) voice-intonation can *also* serve to open the channel between the supposition generator and the desires box.

This proposal isn’t very plausible, either, however. One problem is that infants begin to engage in pretend play while still quite young – around the age of 18 months – at which age many of them are barely talking, and almost certainly aren’t yet mentally rehearsing episodes of *inner* speech. Another problem is that, even with somewhat older children (who are certainly *capable* of inner speech), it seems highly implausible that every episode of solitary pretence should begin with some sort of speech act (whether overt or mentally rehearsed) using an appropriate voice intonation. Consider the child pretending that the banana is a telephone: might not the episode begin when the child is *reminded* of a telephone by its similarity in shape with the banana that he sees? And this surely needn’t require that the child should articulate to himself in English, ‘That is a telephone’, or anything similar. Or the episode might begin when the child visualizes himself lifting the banana, making dialing movements, and placing it to his ear. Surely no forms of verbalization (inner or outer) need be required.

The most plausible suggestion that I can make on N&S’s behalf is that the channel between the supposition generator and the desires box is opened by a pre-existing desire to pretend *something*. When one is in the mood for pretence, as it were, then novel suppositions will automatically give rise to desires to behave as one would behave if those suppositions were true. *One* way of getting into the right mood might be by hearing (or mentally rehearsing) utterances made with a ‘let’s pretend’ voice intonation. But perhaps boredom, or other sorts of stimuli or background circumstance can give rise to a generalized desire to pretend, as well.

It seems to me doubtful, however, whether every episode of pretence first requires the activation of some sort of generalized desire to pretend, before it can begin. Granted, it is sometimes the case that one first wants something to pretend, and then casts around for some suitable supposition to make. But I doubt whether this has to be so. Mightn’t the child have had *no* pre-existing desire to pretend anything when he spotted the resemblance between the banana and a telephone handset, for example? And yet we can’t say that *spotting resemblances* is one of the cues that causes a generalized desire to pretend, either, since we are perfectly well capable of

seeing a resemblance without being motivated to engage in pretence.

2.2 *How to explain the patterning of pretence?*

I have been arguing that it is difficult for N&S to explain why only *some* suppositions should give rise to the desire to behave as one would if those suppositions were true. But there is a further (if not unrelated) problem with their account. This is to explain the distinctive *patterning* that we see in forms of pretence.

For the most part, people pretend to be, or to do, things that they find in some way admirable or valuable. Children who admire farmers and farm machinery pretend to drive tractors and harvesters; children who don't, don't. Likewise children who admire soldiers, and who take an interest in weapons and warfare, pretend to fight battles and shoot one another; whereas children who admire homemaker roles pretend to bathe babies and make mud pies. These facts are puzzling if (as N&S seem bound to maintain) any sort of supposition that gets made (in the right kind of pretend mode) gives rise to a desire to behave as one would behave if that supposition were true,.

Of course, children *can* (and sometimes do) pretend to be or do almost anything, whether admirable or not. They can pretend to be lying in hospital badly injured and in pain; they can pretend to be eating their most hated kind of food; or they can (in N&S's memorable example) pretend to be a dead cat. But the fact is that children don't *often* engage in these forms of pretence. And when they do, their pretence can plausibly be explained as being instrumentally, rather than intrinsically, motivated. Children engaged in forms of co-operative role-play (such as doctors and patients) may have to take their turn at playing the less desirable role, for example. And children can engage in strange and unexpected forms of play in order to make other people laugh.⁴

I think the only real option for N&S at this point is to claim that background concerns and interests tend to influence the sorts of suppositions that get entertained in the first place. While any supposition, once made (and made in the presence of a generalized desire to pretend, perhaps), will give rise to a desire to act as if that supposition were true, perhaps only those suppositions that are in some way interesting to the subject will tend to get entertained at all.

⁴ It is noteworthy that the child who pretended to be a dead cat began by drawing the attention of an audience, saying, 'Hey, look at me: I'm a dead cat', before lying down to enact the role (N&S, 2003).

Hence the patterning that we see in children's pretence can be explained: it is because the child is interested in guns and warfare that he entertains the supposition, *I am creeping with my [toy] gun behind enemy lines*; and it is because the child is admiring of home-making roles that she entertains the supposition, *This [doll] is my baby and it is time for the baby's bath*.

Although at one level this account can work, at another level it just introduces a new set of mysteries. For *how* do background interests and values make it the case that some suppositions, but not others, are entertained? For suppositions aren't *actions* that might be *chosen* in the light of one's desires or interests. So how are background interests supposed to influence them? – And not just *influence* them, but bring it about that suppositions that are incongruent with one's interests and values are hardly ever entertained. (For if they *were* entertained, then they would automatically give rise to a desire to act as if that supposition were true, on the account that we are considering. In which case we would expect to see a great deal more value-incongruent pretence than we actually do.)

N&S might reply, with some justice, that it is mysterious where novel thoughts appear from in any case. So it is nothing really to be wondered at if it should remain mysterious how subjects' interests might channel the suppositions that they entertain. Now, I think this would be a fair point if there were no *other* account of the motivations involved in pretence that can do better on this score. But as we shall see shortly, there is.

Notice, too, that the account sketched above has some implausible consequences. If the influence of background interests and values occurs only at the point of supposition generation, then we should predict that any child who can be induced to entertain another's supposition should feel motivated to pretend accordingly. So if a little girl who cares nothing for fighting can be induced to *suppose* that she is creeping behind enemy lines with a gun in her hand, she should *want* to behave as she would if that supposition were true. And if a little boy who hates even to be in the house during the daytime can be induced to *suppose* that his sister's doll is his baby, then he should likewise *want* to behave as if that supposition were true. To me these predictions don't carry the ring of truth.⁵

⁵ Sometimes, of course, children have social desires that might lead them to want to play *at something* with another child. In order to fulfill this desire, they might have to engage in forms of pretence that wouldn't otherwise interest them.

2.3 *How to explain the connections with other uses of imagination?*

I have argued that N&S have difficulty in explaining why it isn't the case that *every* supposition gives rise to a desire to behave as if that supposition were true. And I have argued that N&S have difficulty in explaining the ways in which pretence tends to be *patterned* in accordance with the agent's interests and values. But this is by no means the end of their problems. For their account can't easily be extended to account for the ways in which motivations are involved in other uses of imagination, either. And yet it seems plausible that the phenomena in question should admit of a common explanation.

N&S claim that the *cognitive* components in their account of pretence (roughly, everything on the left hand side of Figure 1) can be regarded as a good account of imagination in general. (What is distinctive about pretence, in their view, is just that it is a kind of imagination that is paired with desires of the sort that issue in pretend behavior; desires, namely, to act as if the imagined states of affairs were the case.) Hence their claim is that the very same apparatus of supposition generator, possible worlds box interacting with inferential systems, and so forth, is also involved in fantasy, in novel-writing and novel-reading, in suppositional reasoning about possible courses of action, and so on. And in the abstract, this claim seems right. It is highly plausible that the very same capacities to generate and to reason with novel suppositions lies at the core of each of these activities.

As is familiar to everyone, however, fantasy, novel-reading, and mental rehearsal of action can all of them engage with our motivational systems, giving rise to emotions. Thus imagined sex can make you sexually aroused, imagined insults and sleights can make you angry, imagined dangers can make you afraid, and so on. Similarly, people experience a whole range of emotions when reading carefully crafted novels, and often have desires regarding the fates of the characters therein. And likewise, when I mentally rehearse what might happen when I go to visit my boss to demand a raise, the results can bring me out in a cold sweat. All of this suggests very strongly that there ought to be an arrow direct from the possible worlds box to the systems that generate desires and emotions.⁶ (Only 'body-monitoring systems' get marked on N&S's

⁶ Note that it is an arrow from the possible worlds box that is needed here, and not just from the supposition generator. For it isn't the supposition that I visit my boss to ask for a raise itself that brings me out in a sweat, but rather the further predictions generated by my inferential systems operating on the initial supposition once it is placed in the possible worlds box.

diagram, in this regard, but plainly there must in any case be others. Not all desires result from changed states of the body.) It looks as if the contents of the possible worlds box can be received as input by the systems that are responsible for generating emotions and desires, and that the latter are thereby keyed into action very much as if those contents were the contents of current perception or current belief.

All of this is missing from N&S's account, however.⁷ And this is a significant problem. For as I remarked above, it seems very plausible that the same sorts of mechanisms should underlie pretence as underlie both fantasy and the mental rehearsal of action. For all are instances of a kind of *supposing*. And as we shall see in Section 5 below, once it is recognized that the contents of the possible worlds box will routinely be received by the emotion-generating systems, then we will have to hand a very different and more plausible explanation of the motivations that underlie pretence behavior. But before I develop my own approach to this question, I shall discuss the account given by Currie and Ravenscroft (2002 – hereafter, C&R). For they are quite explicit in proposing a theory that will explain all of the many uses of the imagination, and not just pretence.

3 Currie and Ravenscroft on desire-like imagining

C&R argue that propositional (as opposed to perceptual / sensory) forms of imagination come in two different varieties: belief-like and desire-like. When we entertain *suppositions* we enter into states that are significantly belief-like: these states guide our reasoning, and to some extent our acting (in pretence) in ways that are similar to the influence of belief. *Supposing* that the banana is a telephone, and drawing the conclusion that it can be used to call grandma, is not very different from *believing* that I am confronted with a telephone, and drawing the very same conclusion. And likewise *acting* on the supposition that the banana is a telephone is not so very different from acting on the belief that there is a telephone on the hall table in front of me. So, while the suppositions that are involved in pretence *aren't* beliefs, they are belief-like – they have an inference-guiding and action-guiding role similar to that of belief.

⁷ In fact N&S do briefly mention that the contents of the possible worlds box might be received as input by the emotion systems (2003, p.33); but they do nothing further with this suggestion. In later work Nichols (forthcoming, a, b) develops this idea much more fully, but doesn't turn it in the direction of explaining the motivations for pretence.

C&R argue that there is, in addition, a variety of imagination that is *desire-like*. It consists of states that, while they *aren't* desires, can motivate both practical reasoning and action in a manner that is similar to that of desire. Thus the child who is pretending that the banana is a telephone may form a desire-like state to call his grandma. It is this state that motivates the subsequent sequence of pretend actions (making dialing movements while holding the banana, putting the banana to her ear, and then talking as if grandma had answered).⁸ But it is only *desire-like* – it isn't a full-blown desire to call grandma, according to C&R. For that would send the child heading to the hall where the real telephone is, or send her looking for her mother to seek her permission to place a call. Note that the content of the state, however (just like the content of a real desire to call grandma), is: *that I call grandma*. The content isn't this: *that I want to call grandma*. To have a pretend desire, on C&R's account, isn't to pretend *that I have a certain desire*. Hence pretend desires don't generally have contents that are higher-order in character. To have a pretend desire is, rather, to enter into a state with a first-order world-directed content (e.g. *calling grandma*) whose causal role is significantly *like* (while also being partly *unlike*) the causal role of desire.

In terms of this account it is now easy to explain why it isn't the case that *all* episodes of supposing or imagining lead to action, in the way that pretending does. For only those belief-like imaginings that are accompanied by a suitably related desire-like imagining will lead to pretend actions. So this is why *supposing* that *Roe versus Wade* is overturned doesn't lead to *pretending* (acting as if) *Roe versus Wade* is overturned – it is because this belief-like imagining occurs in the absence any suitably related desire-like imagining, such as a desire-like state to make a protest at the change. What is distinctive of pretence, according to C&R, is that belief-like imaginings are paired with desire-like imaginings in just the right sort of way to cause (pretend) actions.

Such an account can also be put to work to explain the commonalities between the ways in which desire-like states figure in pretence and the ways that they figure in other uses of the imagination, such as our responses to fiction. In all of the various domains of imagination (fantasy, fiction, pretence, and so on) desire-like states are implicated, according to C&R. While

⁸ While C&R do commit themselves to the existence of desire-like imaginings, they don't fully commit themselves to this account of the motivations underlying pretend actions. (See C&R, 2002, pp.130-131.) For present purposes, and to simplify my discussion, I shall treat them as if they did.

watching a performance of *Othello*, for example, I might *want* Desdemona to survive. While this state is desire-*like* – perhaps causing me to cry when she doesn't – it isn't a real desire. For at the same time I want the play to be performed as it was written, and take delight in the tragedy of it; and this means wanting that Desdemona should *not* survive. And neither, of course, do I leap up onto the stage to help her.

What of the patterning that we see in children's pretence, however? Can C&R's account explain that? This is much more problematic. For what needs to be explained is why children's *real* desires, values, and interests should constrain their adoption of states that are desire-*like*. Why is it that for the most part children only pretend (for its own sake) to be or to do what they admire, value, or want? For there are no similar constraints on the adoption of belief-*like* states. Children routinely suppose things, in their pretending, that conflict with what they actually believe. Thus the child has no difficulty in supposing that the banana is a telephone, although she really knows, of course, that it is just a banana. Indeed, children have no difficulty entertaining suppositions that are actually impossible, such as the supposition that they can fly faster than a bullet (like Superman), or the supposition that they are steam trains or airplanes.

C&R notice this phenomenon, and stress that it figures importantly in our responses to fiction, too. (They term it 'imaginative resistance', following Moran, 1994.) They point out that we are often highly resistant to adopting an imaginative perspective that involves desires or moral values that are alien to our own. It is hard for us to identify with a character in a novel whose main desire is to kill and cook little children. And novelists will have to devote considerable effort and skill if they are to induce us to take a story seriously that requires the adoption of an alien moral system as one of its central background assumptions. In contrast, it is easy for us to adopt alien beliefs while imagining. (Consider how easy it is for us to become immersed in a work of science fiction, in which people can transform their bodily size and shape as they wish – e.g. turning into an insect – or can travel faster than the speed of light.) C&R acknowledge these points, but make no real attempt to explain them. Their problem, in short, is to explain why we can't, in our imaginings, adopt contrary-to-desire suppositional desire-*like* states at will, in the way that we *can* adopt contrary-to-belief suppositional belief-*like* states at will.

Now notice, as is quite familiar, that imagination can certainly evoke real *emotions*. Imagined insults can make you angry; imagined danger can make you afraid; the death of a

character in a novel or film can make you cry; and so forth. So why shouldn't we also accept that imagination can evoke real *desires*? And indeed, imagined delicacies can make you hungry (wanting food), just as imagined sex can make you sexy (wanting sexual relief). Our account can then be that suppositions (belief-like imaginings) aren't just taken as input by a suite of inferential mechanisms that would otherwise be employed in generating new beliefs from old, or in practical reasoning, as C&R claim, but that they are also taken as input by a variety of desire-generating and emotion-creating mechanisms. Hence we can claim that the desire-like states that occur in imagination are actually *real* desires, produced by the normal operations of such mechanisms, but in this case responding to suppositional input. And we can claim that this is the *only* way in which such states can be generated – passively, in response to belief-like and perception-like imaginings.

But how can these be real desires if they don't lead to real actions? (Although frightened by the film, I don't run out from the theatre; and although saddened by a character's death, I don't write a condolence note.) The answer to this is easy. It is that real desires will normally only lead to real actions *when interacting with real beliefs*. (I shall return to consider the case of pretend action in Section 5.) We are allowing that suppositions and belief-like imaginings *aren't* real beliefs. They differ from real beliefs in crucial aspects of their functional role. For example, the deduced consequences of suppositions are themselves merely suppositions, and aren't stored in memory and reactivated in the manner of beliefs; and the practical reasoning that takes place within the scope of belief-like imaginings doesn't normally give rise to action, nor directly to intentions to act. So it is easy to allow that the desire-like states that occur during episodes of imagining are genuine desires, while explaining why they don't have all of the usual functional consequences of desires. This is because those desires aren't, during the episode of imagining, interacting with real beliefs.⁹

This account can readily explain why it is so difficult for us to adopt alien desires or values in imagination. This is because there isn't any desiderative equivalent of *supposing*, or because there is no such thing as desire-like imagining. What a novelist or playwright has to do,

⁹ Notice, however, that once you *finish* fantasizing about the meal that you might order during your next visit to Paris – in the course of which you haven't *really* tried to call a waiter, of course – the real hunger that you have generated may send you heading to the kitchen for a snack. And children who play at making telephone calls will often say things like, 'Now let's *really* call grandma', after the episode of pretence has finished.

in order to evoke in us a motivational response that we wouldn't normally have, is to manipulate our belief-like imaginings in such a way that a response of that kind might naturally be created. We are induced to suppose that Othello and Desdemona are real people and (later) to suppose that Desdemona, although innocent of any wrong-doing, has just died because of Iago's deceit. This latter supposition is received as input by our emotion-generating systems and processed in the usual way, creating real sadness (and perhaps real anger). Hence although *we* aren't fooled into thinking that anyone has really died, our emotion systems *are* so fooled.

And I suspect, too, that what an author has to do in order to bring us to full imaginative acceptance of an alien moral system, is to encourage us to take on in imagination a rich network of normative *beliefs* ('It is good for the weak to suffer', 'The strong ought to express their dominance over the weak', or whatever). The author then relies on the fact that such attitudes are designed to straddle the belief / desire divide in order to produce in us at least some echo of the corresponding emotions and motivations. (In general, when I come to *believe* that I ought to do something, I also have some corresponding motivation to do that thing. See Carruthers, 2004; Sripada and Stich, 2006.)

But *why* should there be this sort of difference between beliefs and desires in respect of their suppositional counterparts? Arguably, the explanation derives from the role of supposition in the mental rehearsal of action, and in reflective practical reasoning more generally, as we will see in the next section. If we assume that mental rehearsal of action is what the human suppositional capacity was originally *for* (in evolutionary terms), then it is easy to understand why the *inputs* to supposition-based processes should all of them derive from the factive (belief-like or perception-like) side of the mind. For it is by supposing that I do something, or by imagining myself doing something, that such rehearsals get started; and it is important to the success of such rehearsals that my motivational reactions should be reliably similar to the ways in which I *would* really respond if the envisaged events were to occur.

4 Damasio on mental rehearsal and somatic monitoring

Damasio (1994, 2003) develops an account of human practical reasoning and decision making that is both body-centered and feeling-involving.¹⁰ The idea is that when reasoning about what to

¹⁰ A related but somewhat different account is presented by Schroeder (2004). (See also Rolls, 1999.) According to Schroeder, what we monitor isn't our *somatic* responses to our emotional and motivational states, but rather the

do, we envisage performing the various actions available to us, together with their likely outcomes. These suppositions are received as input by the various emotion-generating and desire-generating systems, which respond as appropriate, and which set in train the suite of neural and chemical changes distinctive of the emotion that gets created. Thus if the option under consideration is processed as fearful, a partial fear-response is initiated: muscles tighten, heart-rate and breathing-rate increase, adrenaline is released into the bloodstream, and so on. These changes are monitored by our somasensory systems, and our motivation to perform the action in question is adjusted accordingly. That action gets *marked* as being desirable or undesirable.¹¹

According to Damasio (1994), the somatic marking system is inoperative in patients who have undergone damage to certain parts of their frontal lobes. Such people are capable of *reasoning* very sensibly about practical matters. Moreover, they still respond emotionally to stimuli whose emotional salience is unlearned (like sudden loud noises) or that give rise to real beliefs whose emotional significance has been learned prior to their brain-damage (as when they are attacked by a bear or a mugger). But they no longer respond emotionally when they *imagine* acting one way rather than another. And they are incapable of learning to respond emotionally to new stimuli (as when they experience repeated losses from a particular deck of cards in a gambling game). As a result, their lives as practical agents are a terrible mess. Few of them can hold down a job or maintain their social relationships for very long. And while they can plan sensibly for the future, maintaining and acting on those plans proves beyond them.

Here is how the somatic marker account can be mapped onto a development of N&S's

pleasure and/or displeasure that is occasioned by such states. He argues that the basic determinants of desire satisfaction and desire frustration (which he identifies with the traditional psychological properties of *reward* and *punishment* respectively) are unconscious, and realized in activity deep amongst the more ancient parts of the brain. But these states are *represented* in the frontal cortex, in an area that appears to be the neural seat of pleasure and displeasure. And it is these hedonic states that we monitor when making decisions. For my purposes it doesn't matter much which of these two accounts is endorsed. Whether what is monitored are hedonic states of pleasure and displeasure, representing states of desire satisfaction and frustration, or rather a wider set of bodily responses to satisfaction and frustration (or both), essentially the same sort of self-monitoring decision-making architecture would seem to be involved. In what follows I shall make use of Damasio's framework for convenience.

¹¹ Damasio (1994) envisages that for much of the time we actually make use of a faster 'as if' neural circuit, which simulates and anticipates our bodily responses without really needing to create them. I shall, for the most part, ignore this complication in what follows.

architecture. When we are reasoning about what to do, the various actions open to us are entertained as suppositions and entered into the possible worlds box. There they are elaborated using a battery of different inferential systems, and some likely consequences are predicted. All of these contents (the initial supposed actions, together with their predicted consequences) are made available as inputs to the desire-generating and emotion-generating mechanisms. The latter set to work processing that input and produce a suite of emotional / bodily reactions. These are monitored, and our motivations towards executing the envisaged actions are adjusted up or down as a result, depending upon the valence (positive or negative) of the emotions in question.

The emotions generated by this process are real ones, not just emotion-*like*. But they are directed towards one or more of the suppositions contained in the possible worlds box, not towards some real (believed-in) state of affairs. Hence of course they don't normally issue in the actions that might be appropriate for those emotions, since there aren't the appropriate beliefs to guide them. For example, the fear that I feel when I mentally rehearse walking into my boss' office to ask for a raise is real enough – my heart beats faster, my mouth goes dry, and my hands start to sweat. But it doesn't (of course) lead me to stop and retreat, since I don't believe that I am *actually* walking into my boss' office to confront him. Rather, the awareness of fear makes me less inclined to do so, perhaps to the point that I abandon the idea altogether.

Once a mechanism of this sort has evolved, it is easy for it to become parasitized in fantasy and in literature. I can entertain the supposition that I walk into a Paris bistro and order my favorite French dish, not as part of an episode of practical reasoning, but for its own sake. For that supposition, together with a number of further elaborations and consequences (e.g. carried forward to the point where I am imagining that the dish in question has just been placed on the table before me), will be received as input by my various emotion-generating and desire-generating systems – in this case producing some of the emotional satisfaction that would attend actually *smelling* the distinctive garlic and wine sauce, for example, as well as producing real (albeit faint) hunger.

Likewise when I enter the theatre to watch a performance of *Othello*: I entertain the supposition that these are real people, really saying the things and doing the things that are acted out on the stage. Those suppositions are received as input by my emotional and desire-generating systems, which set to work to produce many of the same feelings that they would create if the events being witnessed were real. Some of these feelings are directly pleasurable (as when Iago

gets what he deserves). But even where the feelings caused would otherwise be aversive (such as sadness at the death of Desdemona), I retain an important element of control over them. For at any moment, by reminding myself that this is only a play, and that no one has really died, I can close down the pretend inputs to my emotion systems, hence shutting down or modulating their response. And this might (in a different way) be pleasurable.¹²

N&S's possible-worlds-box architecture, then, when combined with the somasensory-monitoring architecture proposed by Damasio (1994), has the capacity to explain many of the commonalities between pretence, fantasy, literature, and mental rehearsal of action for purposes of practical reasoning. In each case the initial suppositions and their consequences are received as input by the agent's motivational systems, and real desires or emotions are produced as a result. And at the same time we can explain the phenomenon of emotional resistance. The reason why we cannot take on alien desires or values at will, is because there is no such thing as a desire-supposer, distinct from the belief-like supposition generator. On the contrary, the only way that desires and emotions can enter into imaginative episodes is by being produced 'passively' by the motivational systems operating on belief-like suppositional input.¹³

5 Explaining pretence

This sort of account can explain how emotions are caused *during* episodes of pretence, plainly – as when a little boy playing war games becomes frightened at the thought that his hiding place is about to be discovered. And so our account of why people are motivated to engage in *some* forms of pretence can parallel an account of why people should watch tragedies and attend horror-movies. In both sorts of case the attraction will be a combination of aversive emotional reactions with the element of control that comes with the knowledge that nothing of the sort pretended or envisaged is really taking place (and/or with admiration at ourselves for responding

¹² There is a long-standing puzzle about how it is possible for us to take pleasure in negative emotional responses to theatre, literature, or music, of course (Levinson, 1997). And while some writers do seek to explain this pleasure in terms of *control*, in something like the manner sketched in the text (Eaton, 1982; Morreall, 1985), this is only one possibility amongst others. Another idea consistent with those being developed in the present chapter would be that the pleasure that I take in negative emotional responses to tragedy results from thinking myself to be in various ways *admirable* for being susceptible to those emotions (Levinson, 1982; Feagin, 1983).

¹³ Of course I can always imagine *that I desire something*. But this is a (belief-like) supposition that takes a desire as its *content*, rather than a desire-like supposition of the sort that C&R envisage.

in the ways that we do). But while this kind of account might explain why children engage in games of hide-and-seek, or play war-games, it can't fit many other cases, nor explain the simple forms of pretence in which infants engage.

So can the hybrid N&S–Damasio account suggested above be extended to explain why children should engage in simple forms of pretence at all? What sort of desire is it that explains the child's pretence, if not the desire to act in the way that they would act if their suppositions were true (as N&S suppose)? And can the account explain, moreover, the sorts of *patterning* that we see in children's pretence? I shall address the former pair of questions first.

5.1 *Why pretend?*

Let us suppose that infants are capable of non-verbal mental rehearsal. Then consider the child who is prompted to entertain the supposition, *That [banana] is a telephone*, perhaps caused by the similarity in shape between the banana in question and a telephone handset. This supposition is placed in the child's possible-worlds box, and taken as input by a variety of inferential systems, one of which delivers the thought, *That [banana] can be used to call grandma*. The child then mentally rehearses the action of using *that* [banana] to call grandma, and this supposition together with its further likely consequences (e.g. talking with grandma) are received as input by the child's motivational systems. Since she loves her grandma, and also loves talking with her grandma, she experiences some positive emotions as a result. These are noted by the somasensory monitoring system, and used to index the envisaged action as desirable. Since the child now *wants* to use *that* [banana] to call grandma, she goes ahead and executes the envisaged action schemata (making dialing movements; placing the banana to her ear; beginning to talk). It is a real desire (to call grandma on *that*) combined with a supposition (that *that* is a telephone) that explains the child's actions.¹⁴

What is it that is innate and/or species-specific, on this account, sufficient to explain the species-specific character of pretend play? I am confident that in other animal species, too, worldly similarities of various sorts can lead to partial activation of concepts, together with some of the resulting inferences. Thus although no animal is likely to possess the concept *telephone*, a

¹⁴ The description provided in this paragraph has been modified slightly – away from what I actually think takes place – to make it fit more naturally within N&S's framework. But these modifications aren't relevant to the main themes and arguments of this chapter. For a fuller account, see Carruthers (forthcoming).

chimpanzee perceiving a ball might have her concept *coconut* partially activated as a result. And I am myself confident that some other animals engage in mental rehearsal of action (Carruthers, forthcoming). What is distinctive of human children may be just that they are disposed to use the former partial-activations as a basis for activating and mentally rehearsing some of the related action-schemata.

But why does the child actually *do* anything, however? Why doesn't she stick to fantasy? For the rewarding emotions involved in talking with grandma will *already* have been experienced, resulting from her mental rehearsal of the action. After all, adults are perfectly well capable of mentally rehearsing an action (such as entering a Paris restaurant and eating dinner) and enjoying the emotions involved, without thereby starting to do anything. So why should the child be any different? A number of distinct (but mutually consistent) answers to these questions can be given.

One answer is that the use of mental rehearsal to initiate action may be the default mode of its operation. That is to say, a mentally rehearsed action that leads to positive emotional rewards will automatically lead to the execution of (or intention to execute) that action unless something intervenes to prevent it.¹⁵ And then the difference between the child and the adult may just be that the adult has learned to use his knowledge that he is not really in Paris and so forth to pre-empt the action that would otherwise be initiated (e.g. calling the restaurant to make a reservation).

Another answer is that the child's imaginative capacities may be insufficient for her to derive significant pleasure from the act of *imagining* talking to grandma alone. As we all know, fantasy can be hard work. Attention has to be carefully focused, and the various imaginary scenarios have to be envisaged in significant detail and with considerable vividness. So it may be that young children are incapable of holding their imagined actions *steady* and/or *vividly* enough in mind to reap significant emotional rewards from them.¹⁶ It may then be that *acting out* the

¹⁵ Where the emotional rewards of a mentally rehearsed action are sufficiently negative, the result will be that the action is *not* performed. The imaginative episode will only lead to action if it is extended to initiate the rehearsal of some *other* action, which does issue in positive rewards.

¹⁶ What the child does may nevertheless be enough for the action in question to be indexed as desirable. For as Damasio (1994) emphasizes, the operations of somasensory monitoring can be both extremely swift, and operate below the level of consciousness, while still having its effects on our goals and choices. Indeed, it can also operate in a purely simulatory 'as if' manner, on his account, in which case no real emotions will be experienced (whether

pretence is necessary for the child to get a vivid enough representation of herself talking with her grandma for her to experience much real emotion.

Relatedly (and even if the initial mental rehearsal of the action *does* yield significant enjoyment) it is likely that acting out the pretence will reinforce, enhance, and extend that enjoyment. For the child's physical movements will be performed under the (supposed) description, *talking to grandma*. (Remember, the action-schema being executed is: *using that [banana] to call and talk to grandma*.) There will therefore be vivid perceptual representations of various sorts that get subsumed under that description, and that get made available as input to the child's desiderative and emotional systems. So when she hears herself chatting into the 'telephone' this will be conceptualized in a manner coherent with her initial suppositions (that the banana is a telephone, and that she is now talking to grandma). So those experiences will be received as input by her emotional systems with the content, *I am talking to grandma* attached, and a much more intense and temporally extended sequence of rewarding emotion is likely to result.

Our N&S–Damasio hybrid can explain why the child engages in pretence, then. But what does it tell us about the *goal* of her action? What is she aiming at when she begins the sequence of movements in question? Is she aiming, in particular, to undergo a rewarding emotional experience? This might be one natural way of interpreting the import of Damasio's (1994) account, issuing in a kind of *hedonism*. The child mentally rehearses talking to grandma on *that [banana]* and experiences a rewarding emotion. So she then executes the mentally rehearsed action in order to undergo that experience again, or more vividly. So the goal of her action is to experience a rewarding emotion.

Although this sort of hedonistic reading of Damasio (1994) might be tempting, I believe that it is wrong. For he talks of the emotions that result from mental rehearsal as *indexing* the actions rehearsed (as desirable or undesirable), not as themselves becoming the *objects* of desire. And this is all to the good, since there are well-known difficulties with hedonistic accounts of agency (Feinberg, 1985). Rather, the account is this: the child mentally rehearses the action-schema, experiences a positive emotion, and thereby comes to desire the execution of that action-schema. So her goal, when she acts, is *to talk to grandma on that [banana]*. This is a goal that both exists within, and only makes sense within, the scope of the initial supposition, that *that*

[banana] *is a telephone*, of course. (This is why the child isn't disappointed when her grandma doesn't answer; for she retains the knowledge that you can't really talk to grandma on a banana.)

5.2 *The limits of pretence*

But how are we to explain the limits that children place on the acting-out of their pretence? Why don't they 'go all the way'? For example, consider a child who is playing at domestic roles, and who acts-out the action-schema *cooking [mud] pies to eat for dinner*. If her goal is *to cook and eat [mud] pies*, why doesn't she follow through and *really* eat the mud? A pair of related, and mutually consistent, replies can be given to this question.

One is that the action-schema that the child rehearses is unlikely to include all of the elements involved in the corresponding real (non-pretend) action. Although at the most abstract level the schema rehearsed is *cooking pies and eating them*, the more detailed implementation of that schema that gets rehearsed in the child's imagination is unlikely to include actual chewing and swallowing. For of course the child retains the knowledge that the 'pies' are made of mud, and that mud doesn't taste good; and this knowledge is likely to guide the child's construction of a detailed action-schema.

The second answer (and an answer that might provide us with one explanation of how the first can be true) is that the child might rehearse a number of more-or-less complete action-schemata, one of which *does* involve putting the 'pies' into her mouth, chewing, and swallowing them. When this schema is mentally rehearsed in combination with the knowledge that the pies in question are made of mud, the sensations and tastes that get predicted are likely to produce a strong negative emotional reaction. This tags that action-schema as undesirable. We can therefore suppose that the action-schema on which the child acts is a sort of compromise between an attractive abstract description ('cooking and eating pies for dinner') and an aversive detailed implementation of that description ('putting those mud objects in my mouth, chewing, and swallowing'). The aversiveness of the latter will shape just how far the child will go in acting-out the former.

5.3 *The patterning of pretence*

Besides its other advantages, the hybrid N&S–Damasio account sketched above provides an elegant explanation of the patterning of pretence. The reason why, in general, children only

pretend to do or to be things that they find in some way valuable or admirable, is that only in these cases will the mental rehearsal of the pretend actions give rise to the sorts of positive emotional reactions that index them as desirable. Let me work through a couple of our earlier examples to illustrate the point.

When the boy who finds guns and warfare exciting and admirable supposes that he is a US marine, and mentally rehearses the action schema, *creeping along with my [toy] gun behind enemy lines*, he experiences emotions of excitement, which tag the action as desirable. And when he then acts out the pretence, and creeps along behind the living room sofa while representing what he is doing under that description, he not only experiences some of that same excitement, but throughout the episode he will be representing himself *as* something that he finds admirable, namely a US marine.

Likewise when a boy who finds steam-trains to be admirable entertains and acts on the supposition, *I am a steam train*, he will be representing himself *as* something that he admires. So the perceptions of his current actions will be made available to his motivational and emotional systems tagged with the content, *I am a steam train going 'chuff, chuff, choo, choo'*, and so forth. Since he admires steam trains, and is representing himself to *be* a steam train, the episode is likely to be emotionally rewarding.¹⁷ (We all like to think that we are admirable, or to be reminded of admirable features of ourselves, of course!)

In contrast, when a boy entertains and acts on the supposition, *I am a dead cat*, he is unlikely to find the episode intrinsically rewarding. For he is unlikely to find dead cats to be especially admirable. We can thus explain why children often pretend to be steam trains (and the like) while rarely pretending to be dead cats (and the like). The kinds of pretence in which children engage will be a direct reflection of their interests and values. So the patterning of children's pretence is satisfying explained.¹⁸

6 Conclusion

¹⁷ This sort of episode might also – or instead – be rewarding for other reasons, of course.

¹⁸ Of course it isn't a problem for this account that a child might *sometimes* pretend to be a dead cat while finding nothing intrinsically admirable about dead cats. For acts of pretence, like any other types of action that are normally intrinsically motivated (like eating or sex), can sometimes be performed for instrumental motives. In this case I hypothesize that the child's goal was to do something funny.

In this chapter I have used the views of N&S and C&R to triangulate my own position. From the former I have taken the basic architecture of supposition-generator and possible worlds box. And from the latter I have taken the idea that the motivations involved in pretence should somehow be of-a-piece with those involved in fantasy, in literature and theatre, and in the mental rehearsal of action. I have suggested that we can meet all of the main *desiderata* for a theory of pretence by combining N&S's architecture with Damasio's (1994) account of mental rehearsal and of somasensory monitoring of emotional reactions to those rehearsals. In short, children pretend because they find both the mental rehearsal and the performance of the pretend actions (under suppositional descriptions) to be emotionally rewarding, reflecting their standing desires, values, and interests.¹⁹

References

- Bekoff, M. and Allen, C. (1998) Intentional communication and social play: how and why animals negotiate and agree to play. In M. Bekoff and J. Byers, *Animal Play*, Cambridge: Cambridge University Press.
- Carruthers, P. (2002) Human creativity: its evolution, its cognitive basis, and its connections with childhood pretence. *British Journal for the Philosophy of Science*, 53, 1-25.
- Carruthers, P. (2004) Practical reasoning in a modular mind. *Mind and Language*, 19, 259-278.
- Carruthers, P. (forthcoming) *The Architecture of the Mind: massive modularity and the flexibility of thought*.
- Currie, G. and Ravenscroft, I. (2002) *Recreative Minds: imagination in philosophy and psychology*. Oxford: Oxford University Press.
- Damasio, A. (1994) *Descartes' Error: emotion, reason, and the human brain*. London: Papermac.
- Damasio, A. (2003) *Looking for Spinoza: joy, sorrow, and the feeling brain*. New York: Harcourt.
- Eaton, M. (1982) A strange kind of sadness. *Journal of Aesthetics and Art Criticism*, 41, 51-63.
- Feagin, S. (1983) The pleasures of tragedy. *American Philosophical Quarterly*, 20, 95-104.
- Feinberg, J. (1985) Psychological egoism. In J. Feinberg, *Reasons and Responsibility*, New York: Wadsworth.

¹⁹ Thanks to Shaun Nichols for his comments on an earlier draft.

- Harris, P. (2000) *The Work of the Imagination*. Oxford: Blackwell.
- Levinson, J. (1982) Music and negative emotion. *Pacific Philosophical Quarterly*, 63, 327-346.
- Levinson, J. (1997) Emotion in response to art: a survey of the terrain. In M. Hjort and S. Laver (eds.), *Emotion and the Arts*, Oxford: Oxford University Press.
- Moran, R. (1994) The expression of feeling in the imagination. *Philosophical Review*, 103, 75-106.
- Morreall, J. (1985) Enjoying negative emotions in fiction. *Philosophy and Literature*, 9, 95-103.
- Nichols, S. (forthcoming, a) Imagining and believing: the promise of a single code. *Journal of Aesthetics and Art Criticism*.
- Nichols, S. (forthcoming, b) Just the imagination: why imagining doesn't behave like believing. *Mind and Language*.
- Nichols, S. and Stich, S. (2003) *Mindreading: an integrated account of pretence, self-awareness, and understanding other minds*. Oxford: Oxford University Press.
- Rolls, E. (1999) *The Brain and Emotion*. Oxford: Oxford University Press.
- Schroeder, T. (2004) *Three Faces of Desire*. Oxford: Oxford University Press.
- Sripada, C. and Stich, S. (2006) A framework for the psychology of norms. In P. Carruthers, S. Laurence, and S. Stich (eds.), *The Innate Mind: culture and cognition*, Oxford: Oxford University Press.