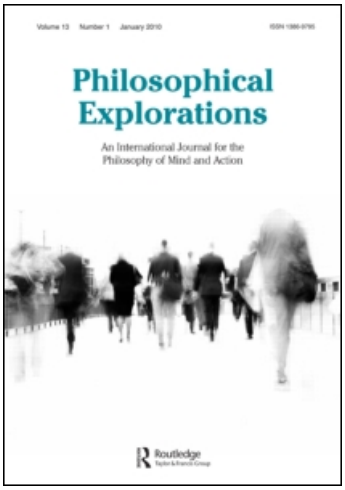

Philosophical Explorations
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713706422

Folk psychology without principles: an alternative to the belief-desire model of action interpretation

Leon C. de Bruin[a]; Derek W. Strijbos[b]
[a] Institute for Philosophy, Ruhr-University Bochum, Germany [b] Department of Philosophy, Radboud University Nijmegen, The Netherlands

Online publication date: 20 September 2010

To cite this Article de Bruin, Leon C. and Strijbos, Derek W.(2010) 'Folk psychology without principles: an alternative to the belief-desire model of action interpretation', Philosophical Explorations, 13: 3, 257 — 274
To link to this Article: DOI: 10.1080/13869795.2010.501899
URL: http://dx.doi.org/10.1080/13869795.2010.501899

Philosophical Explorations
Vol. 13, No. 3, September 2010, 257–274

Folk psychology without principles: an alternative to the belief–desire model of action interpretation

Leon C. de Bruin[a] and Derek W. Strijbos[b]*

[a]*Institute for Philosophy, Ruhr-University Bochum, Germany;* [b]*Department of Philosophy, Radboud University Nijmegen, The Netherlands*

In this paper, we take issue with the belief–desire model of second- and third-person action interpretation as it is presented by both theory theories and cognitivist versions of simulation theory. These accounts take action interpretation to consist in the (tacit) attribution of proper belief–desire pairs that mirror the structure of formally valid practical inferences. We argue that the belief–desire model rests on the unwarranted assumption that the interpreter can only reach the agent's practical context of action through inference. This assumption betrays a deep-seated bias toward disengaged, observational interpretation strategies. On our alternative picture, the interpreter can start off on the assumption of a shared practical context and proceed to reason discourse in those cases in which this assumption runs aground. Following Brandom's non-formalist account of reason discourse, we suggest that interpreting other people's actions in terms of reasons is not a matter of following the principles of formally valid practical syllogisms, but of endorsing practical material inferences that are correct in virtue of a shared practical world.

Keywords: folk psychology; belief–desire model; theory theory; simulation theory; Robert Brandom

1. Introduction

We normally perform actions for reasons in the course of our daily affairs. In order to get along with each other, it is often important that we find out about each other's reasons for action. And we are quite successful in getting at those reasons when and where it is important that we do so. In the course of three decades of intense debate, numerous positions have been carved out in order to explain what this kind of action interpretation[1] consists in, how it works and how people acquire it. From the start, there has been a tendency to frame this discussion in terms of the so-called belief–desire (BD) model. Accordingly, making sense of other people's actions requires the (re)construction of the action under consideration in terms of a desire toward some goal and a belief regarding the means. This requirement mirrors the 'central action principle' of folk psychology: 'if A wants p and believes that doing q will bring about p, then ceteris paribus, A will q' (Borg 2007, 6).[2]

*Corresponding author. Email: d.strijbos@phil.ru.nl

ISSN 1386-9795 print/ISSN 1741-5918 online

DOI: 10.1080/13869795.2010.501899
http://www.informaworld.com

Yet consider for a moment what we *say* when we give reasons for our actions. 'Why are you getting up so early?' 'Got an early client this morning'. 'Why didn't you call me?' 'I didn't want to ruin your trip'. 'Why is he making such an effort?' 'He thinks he's in for promotion'. We practically never give and get answers to such questions in proper syllogistic format. According to the BD model, the routines of BD integration are supposed to be at work nonetheless: not necessarily explicitly or consciously, but tacitly, subconsciously. The result is a 'principled' picture of action interpretation, one according to which every genuine act of action interpretation eventually follows something akin to the above central action principle, embodied in some or other 'mindreading' procedure. On this principled picture, all instances of action interpretation minimally take this same route, independently of what has actually been said in discourse. In case of 'truncated' answers such as the above, in order to really understand them and appreciate them as genuine reason explanations, the interpreter needs to silently fill in the missing syllogistic premises.

In this paper, we take issue with the principled picture of action interpretation and the BD model that gives rise to it. Our primary aim is to show that there is a serious, unprincipled alternative to this model. We start by discussing principled BD accounts of both traditional theory (TT) and simulation theory (ST) varieties in section 2. While the ST varieties oppose to the idea that the interpreter needs to *represent* action principles, they nevertheless agree with their TT cousins that the simulation procedure needs to *mirror* such principles. In section 3, we argue that a commitment to principled mindreading routines of both TT and ST format rests on an implicit assumption to the effect that in making sense of the agent's action, the interpreter needs to *infer* the agent's practical context of action. This assumption is not argued for and betrays a deep-seated bias toward disengaged, observational interpretation strategies. It is also an implausible assumption, given the availability of a cognitively less demanding interpretation strategy. According to this alternative, the interpreter can start off on the assumption of a *shared* practical context between herself and the agent and normally proceed to reason *discourse* in those cases in which this assumption runs aground.

Taking people's participation in second-person conversations about reasons as the point of departure, we draw on Robert Brandom's (1994, 2000) account of reason discourse and propose to conceive of this process not as a (tacit) reconstruction of *formally valid* inferences mirroring the *principles* of practical syllogisms, but as a matter of proposing and endorsing *materially correct* inferences that reflect *patterns* in a largely shared practical world (section 4). Sketching the contours of our alternative account in section 5, we hint at the suggestion of construing third-person 'mindreading' as the extension of a primary, second-person practice of giving and asking for reasons. We end this paper by drawing some interesting parallels with other non-standard accounts of action interpretation.

2. The BD model: varieties of principled action interpretation

It is important to keep in mind, here and throughout, that our focus is on action interpretation in terms of the agent's reasons for action, the kind of interpretation that Davidson (1963/2001) was interested in: full-blown propositional action interpretation with a practical demand to distinguish among all the reasons the agent possibly could have (had) the one she in fact has (did have). We had better be explicit about this, since there is a growing body of literature on the so-called 'low-level' forms of interpretation, in terms of 'face-based emotion recognition' (Goldman 2006), 'empathic resonance' (Iacoboni 2005; Slors 2007), 'primary and secondary intersubjectivity' (Trevarthen and Hubley 1978; Trevarthen 1979; adopted by Gallagher and Hutto 2008). While we think this

renewed interest in low-level social cognition is of crucial importance for an inclusive picture of interpretation, and while we agree with many of these authors that appreciation of these forms of interpretation in daily human interaction severely restricts the scope of 'high-level' interpretation, for the purpose of this paper we wish to put these issues to one side.

When it comes to characterizing high-level action interpretation as specified above, there are three positions that need to be mentioned: the TT and ST and the fairly recently established 'narrative theories' (NT). Although most theories are framed in terms of the BD model, it would be wrong to conclude that they are *necessarily* committed to this model and the resulting principled picture of action interpretation. Virtually all versions of TT explicitly endorse the BD model, but there is nothing in principle about *theory* theory that requires such endorsement (think of e.g. 'behaviorist' versions of TT). With respect to ST, there is a minority of theorists who do not explicitly frame their theory in terms of the BD model. Here we are thinking of Gordon (1986, 1996) and his 'radical' ST and Heal's (1986, 1998) account of 'replication' or 'co-cognition'. These theories may be regarded as hostile to the principled view of action interpretation, although this has not always been explicitly argued for. And even though Hutto's (2008a, 2008b) 'narrative practice hypothesis', the most elaborate version of NT to date, seems to adopt the core of the BD model (unfortunately in our view, see section 6), it is actually targeted against the principled picture that goes with it.

Our issue is thus not with these theories *per se*. Far from it: there are aspects of Gordon's and Heal's ST that will be incorporated in our Brandomian alternative (section 3), and a narrative rendering of interpretation may in fact prove a valuable extension of the account we provide here (section 6). Our target is the BD model and the resulting principled picture as it is adopted by some of them. In the remainder of this section, we shall therefore consider versions of TT and ST that explicitly endorse the BD model of action interpretation.

On the BD model, action interpretation necessarily involves the attribution of different kinds of propositional attitudes to the agent under consideration: predominantly beliefs and desires and minimally a desire toward some end and a belief regarding the means. Obviously, for it to be a genuine case of interpretation in terms of reasons, the beliefs and desires attributed must meet certain constraints. The idea is that by engaging in a procedure often referred to as 'mindreading', we discover the BD pair that actually played a central role in bringing about the action that is being interpreted. This is so whether it involves past, present or future actions: in the case of past or present actions, we normally work back from the action to the BD pair, whereas in the case of a future action, we normally start with the two interlocking propositional attitudes and work our way toward a predicted or anticipated behavioral outcome.

According to the theory theory of action interpretation, making sense of each other's actions in terms of reasons is normally carried out through the use of a psychological theory that specifies how beliefs and desires combine to give rise to (intentions and) actions.[3] The tenets of this psychological theory are often made explicit in terms of folk-psychological generalizations. Gopnik and Meltzhof (1997, 126), for example, claim that the theory '(…) has many complexities but also a few basic causal tenets (…) These tenets are perhaps best summarized by the "practical syllogism": if a psychological agent wants event y and believes that action x will cause event y, he will do x'. More is needed, of course, in order to select the contents of the attitudes over which the theory quantifies in a particular situation. On TT proposals, this requires more theory about how beliefs and desires relate to perceptions, bodily expressions, (verbal) behavior and other mental

states. Although some psychological generalizations can be made explicit, most theory theorists agree that they are largely stored and drawn upon tacitly.

Importantly, all these folk-psychological generalizations crucially depend for their accuracy on *ceteris paribus* clauses. To be of any practical use, it is therefore vital that the context of the interpreted behavior be taken into account.[4] There may be other mental states to be inferred from behavioral evidence and environmental cues, or situational factors, character traits, personal histories and behavioral limitations that exceed these clauses and make the generalization less adequate or even useless for interpretation purposes in the particular context of action. For TT to be an adequate account of our interpretative skills, folk-psychological generalizations should be embedded in extensive know-how concerning their context-sensitive application (cf. Gallagher 2004). If we stay within the framework of TT, this know-how should itself be governed by yet another layer of knowledge of rules specifying the conditions of application of mentioned folk-psychological generalizations.[5]

There are two worries here. The first concerns computational complexity: applying folk-psychological principles to specific situations requires computation over a host of additional clauses to determine whether the appropriate background conditions hold and whether there are countervailing factors in play (cf. Bermudez 2004). Although this hardly provides a knock-down argument against TT, it might tip the balance in favor of cognitively more parsimonious accounts. Another closely related but deeper worry is that TT here faces a precursor to the frame problem: the problem of determining relevance. As Spaulding (2010, 136) writes, the problem is 'how one determines which general principle to apply in a particular case given that the relevant information for determining which principle is appropriate is in principle unlimited and could come from any domain'.

ST has traditionally been considered the main rival of TT. Simulation theorists believe that our default procedure in making sense of actions in terms of reasons consists in *simulating* the process responsible for the action. According to Goldman's (1989, 2006) cognitivist version of ST, mindreading normally consists in a process of feeding pretend beliefs and pretend desires into one's own offline practical reasoning system and attributing the resulting pretend decisions (and corresponding actions) to the agent who is being simulated.[6]

It has always been one of ST's main arguments against TT that ST gives a more parsimonious explanation of mindreading: where most versions of TT are committed to tacit knowledge of FP principles and the rules of their application, ST holds that the simulation routine is 'process-driven' rather than 'theory-driven' (Goldman 1989). Since our own decision-making process does not require *psychological* knowledge concerning the propositional attitude*s* – it only requires unfolding the *contents* (propositions) involved – putting this *same* process to use in interpreting other people's actions does not seem to require such knowledge either.[7]

But even if process-driven simulation is not knowledge-based: How does the simulator know *which* beliefs and desires to process in pretend mode? Goldman goes hybrid here and proposes that selecting the correct mental states to be processed in pretend mode requires the use of an additional folk psychological theory. Thus, he explains that 'In a decision-prediction task, an attributor would use theoretical reasoning to infer the target's initial states (desires and beliefs), for which the corresponding pretend states are constructed. The pretend states are then fed into the decision making mechanism, which outputs a decision. The first step of this sequence features theorizing, whereas the remaining steps feature simulating' (2006, 44). As concerns backtracking of reasons for actions already performed, Goldman suggests that 'This type of mindreading might be approached via a

generate-and-test strategy. [. . .] The generate-and-test strategy employs simulation at a crucial juncture but also relies on theorizing. Theorizing seems necessary to generate hypotheses about states responsible for the observed effects, hypotheses presumably prompted by background information' (2006, 45). On this construal, however, ST faces problems of computational complexity in the selection phase similar to those of TT. And the problem of relevance appears to be equally pertinent: 'For the Simulation Theory, the problem is with identifying what the relevant mental states of the target could have been in order to cause the observed behavior when the information relevant to identifying the correct mental states is in principle unlimited and could come from any domain' (Spaulding 2010, 136).

Both TT and cognitivist ST subscribe to the idea that making sense of other people's performances in terms of reasons requires interrelating desires and beliefs that are supposed to play a central role in the bringing about of the action under consideration. Their dispute concerns the precise nature of this cognitive procedure. Given the commitment to the BD model of action interpretation, both theories *must* take this procedure to at least *mirror* folk-psychological principles as explicated in practical syllogisms. Goldman's ST variant is no exception here: even though the central simulation routine might not require tacit knowledge of central action principles, it still needs to *follow* the cognitive route marked by them: no pretend-decision and hence no action interpretation without integrating the pretend-belief with the pretend-desire. Therefore, as Goldman admits, even if we think of simulation as being process-driven, such a process still requires that 'some elements inside the attributor causally mediate between his explicit premises and conclusions, and that the causal structure of these elements mirrors the logical structure of psychological theory' (2006, 33). It is in this sense that both accounts are 'principled': the cognitive procedure must meet logical constraints that apply across the board, for every instance of proper action interpretation in terms of the agent's reasons.

For quite some time now, the TT and ST varieties discussed have been dominating the debate. Accordingly, there seems to have been a silent consensus among many participants that when it comes to high-level interpretation in terms of reasons for action, principled BD accounts of action interpretation slice up the theoretical options without remainder.

3. Reasons in context

As a first step toward an alternative conception of action interpretation, it may prove useful to trace the origins of the BD model. Perhaps, we only have to go back as far as Davidson (1963/2001), who argued that we must revert to the notion of causality in order to account for the 'because' in 'she did *a* because *r*'. There are several candidate reasons that rationalize a particular action, but normally only one of them renders the statement true. According to Davidson, this reason must be the reason that caused the action. Davidson termed such reasons 'primary reasons'. A primary reason why an agent performed an action (under a certain description) consists of a pro attitude of the agent toward actions with a certain property and a belief of the agent that the action performed (under that description) has that property. He notoriously claimed that 'In order to understand how a reason of any kind rationalizes an action it is necessary and sufficient that we see, at least in essential outline, how to construct a primary reason' (2001, 4).

Few today would hold on to Davidson's sufficiency claim, and usually the BD model construes desires as directed toward the agent's ends (rather than an action type) and the beliefs as directed toward means to ends (rather than a token action of the desired type). Yet the idea that ascribing a BD pair is a necessary ingredient of the interpretation

process appears to follow straightforwardly from the Davidsonian picture of action interpretation: in order to single out *the* reason why an agent performed an action, it is necessary that the interpreter attribute the primary reason that caused the action.

But this is too quick. For who is the 'we' in the above quote who is supposed to 'see, at least in essential outline, how to construct a primary reason'? As far as Davidson's claim goes, this is the *action theorist*. Davidson provides a *causal analysis* of intentional action, not a *cognitive hypothesis* regarding folk psychological interpretation. All Davidson's proposal strictly amounts to here is that the cause of an intentional action can always be described correctly in terms of an appropriately structured BD pair. But from the claim that a reason explanation *logically implies* an appropriately structured BD pair on the part of the agent, it does not follow that explaining the action *psychologically requires* ascribing such BD pair on the part of the interpreter. So why is the latter so widely endorsed nonetheless? In our view, this has to do with the fact that on both TT and cognitivist ST accounts, the interpreter takes a thoroughly disengaged stance toward the agent to be interpreted. Disengaged in the sense that the interpreter is assumed not to *share* a practical context with the agent.[8]

Heal (1996) argued that TT faces a version of the frame problem if it tries to account for our folk-psychological ability to think about other people's thoughts in purely theoretical terms. One of the advantages of simulation as 'co-cognition' – 'a fancy name for the everyday notion of thinking about the same subject-matter' (Heal 1998, 483) – is that the interpreter can use *her own frame of mind* in order to predict and explain other people's thoughts. Instead of theorizing over the inferential relations between different thought contents and how these should be *applied to* a particular context, the co-cognizer can simply use her own ability to cognize these contents herself *within* her own particular context to track the contents of other people's thoughts. Since the interpreter's and interpretee's background overlap significantly, this strategy normally yields an adequate prediction of the contents of the interpretee's mind. Even in the case the interpreter starts with a content that is not among her beliefs – say when she is already been informed that the interpretee has a false belief or when she wonders what the interpretee will think in a hypothetical situation – she will frame this content within her own background knowledge and use her own cognition skills to figure out the interpretee's thoughts.

Heal did not use the same argumentative strategy for the interpretation of *action*, but we think the same considerations apply here. Consider TT first. If the central action principles could be framed by the *interpreter's own view on the world*, additional generalizations that specify the context-dependent application of the central action principles would no longer be needed. The interpreter could simply depend on her own recognition capacities in order to determine when and where to apply the principles. But once this is considered a viable alternative, there is nothing to prevent us from accounting for (some of) the clauses of the central action principles *themselves* in terms of the interpreter's own practical context as well. If the interpreter can rely on her own background beliefs, desires etc., to frame the relevant practical syllogism, than why not let, for example, the relevant instrumental belief be part of that background as well? If this is possible, as it surely is in those many situations that interpreter and agent *share* the relevant belief, than why have the interpreter *represent* the relevant belief as part of a central action principle?

Now consider cognitivist ST. Goldman's reliance on TT in order to account for the selection of the relevant BD pair to be processed in pretend mode is rather surprising given the availability of a broadly simulationist alternative. Why couldn't the interpreter simply rely on her own background beliefs and desires in order to embed the relevant BD pair to be processed in pretend-mode? The standard objection to this move is that

this would make simulation too unreliable for practical purposes. But this seems to assume that the interpreter's and agent's background beliefs and desires are systematically and significantly different. This assumption is not argued for and moreover not very plausible. Of course, there are plenty of cases in which the standard assumption of a shared background runs aground, but there is no reason at all to take these cases *as default*. Rather, these are cases when *default strategies no longer work* and when the next best thing is to *start asking the agent* for reasons why (see below). But if it is conceded that the assumption of a shared background is a viable alternative default strategy, then why suppose that the relevant BD pair need always be processed in *pretend* mode? If, for example, the instrumental belief is one that the interpreter and the agent *share*, then to insist that the interpreter needs to take *this* belief offline is to beg the question as to why *only* this belief needs to be taken offline. Biting the bullet here, one could say that in fact the interpreter takes *all* his background beliefs offline and attributes all of them to the agent in simulating him. But this is duplication beyond necessity. In those cases where interpreter and agent share the relevant instrumental belief, then, it seems the interpreter can suffice by taking only the relevant desire in pretend mode, using her own frame of mind to figure out what decision follows. And when interpreter and agent share both belief and desire, it seems the interpreter only needs to *look out into the world* to figure out what the agent is up to next (cf. Gordon 2000).

The disengaged picture of folk psychological interpretation of mentioned TT and cognitivist ST accounts lies in the assumption that the context of action must be *inferred*. This, in turn, only makes sense if it is assumed that interpreter and agent do not share their practical context. This assumption is not argued for. It does explain, however, why mentioned TT and ST accounts rely on the central action principle so much. Our discussion of Davidson at the beginning of this section revealed the *logical* thesis that an action performed for a reason can *in principle* be described correctly in terms of a primary reason (or in terms of an ends-directed desire and a means-directed belief). On the disengaged picture sketched by mentioned accounts, the interpreter is assumed not to rely on her own practical context, a context that she by default shares with the agent. That is why she needs (cognitive routines mirroring) principles that apply *across different contexts*. From this perspective, it is tempting to transform the above *logical* implication into a *psychological* requirement on the part of the interpreter. The resulting picture is one according to which the interpreter starts with *context-free* principles or principled routines, which then need to be *applied to* an *inferred* context of action.

But if we allow for the possibility of interpreter and agent sharing a practical context to some significant degree, the need for applying FP principles or running cognitive routines mirroring such principles is obviated. On the assumption of a shared context, action interpretation neither requires an 'inference from me to you' as it does on Goldman's account (Gordon 1995), nor an inference from the agent's environment and behavior to the contents of his mind, as on standard TT accounts. More in particular, if we allow the interpreter to frame the agent's action in terms of her own view on things, we may start sketching an alternative picture, one on which making sense of another agent can rely on unprincipled, essentially context-sensitive patterns of practical thinking.[9]

It might be objected that by applying one's own frame of mind, the interpreter fails to account for the *differences* in beliefs and desires between herself and the agent. No doubt interpreter and agent can differ with respect to the contents of their mental states, but it is again symptomatic of a thoroughly disengaged picture of folk-psychological interpretation to assume that the interpreter has to account for such differences *all by herself* – privately speculating about the possible contents of the agent's mind. The standard accounts direct

their attention exclusively at *third-person* contexts of action interpretation, situations in which the interpreter is a bystander, someone observing the agent without interacting with him. *Second-person* action interpretation is simply left out of the picture. Characterizing interpretative practice from a detached point of view, it seems that the only way to properly explain and successfully predict other people's behavior in cases of non-shared contexts is by starting to speculate, by means of either theorizing or simulating.

Yet why think that people are so good at third-person mindreading in the first place? What appears to have been systematically overlooked is that when and where it really matters to know the *agent's* side of the story in such puzzling cases, people *do not trust* their own hypotheses: there simply are *too many* plausible candidates (cf. Hutto 2004). Rather, interpreters tend to ask someone, preferably the agent himself. Privately speculating about the reasons of others is something people normally engage in only when nothing important hangs on it or when asking is inappropriate, inconvenient or impossible. And even on such occasions, it should be noted that there are often *other* people around who can help the interpreter understand the agent's reasons. They may know him better; they may have spoken to him about the action or to someone else who did. Think of a practice like gossiping. Strictly speaking, it does not constitute a third-person context of interpretation: here we have one interpreter *asking* another interpreter for the agent's reason. Compared to the answers we get from other people, the explanations and predictions drawn from reason speculation are generally far less reliable for interactive purposes. In short, there are good reasons to think that when and where it is important to get the agent's perspective into clear view in the case of divergent practical contexts between interpreter and agent, the practically most reliable and effective way to achieve this is *by participating in reason discourse*.

Now consider again the short examples given in the introduction: 'Why are you getting up so early?' 'Got an early client this morning'. 'Why didn't you call me?' 'I didn't want to ruin your trip'. 'Why is he making such an effort?' 'He thinks he's in for promotion'. Starting from such ordinary instances of second-person action interpretation, it makes *prima facie* good sense to take the reason discourse exhibited at *face value*. This, in short, is what we propose in the following two sections. On this approach, people often have something, or rather someone, at their disposal who bypasses the need (on the side of the interpreter) to employ mindreading routines for the tacit reconstruction of BD pairs: the agent himself. On many occasions, it is the agent who actively takes part in the contextualizing work, by simply answering the interpreter's questions when the latter does not follow the explicit answers already given. In the following two sections, we will show how successful participation in reason discourse can be characterized as the construction and endorsement of practical material inferences within a 'deontic context', the defining feature of such inferences being that they do *not* necessarily contain well-formed BD pairs.

4. A non-formalist conception of reason discourse

People normally start asking each other for reasons when their actions deviate from social norms or expectations. This is often appropriately characterized as a form of challenge: the interpreter challenges the propriety of the action performed. Giving a reason in such cases can be regarded as providing a justification, one the interpreter will hopefully accept. Crucially, the agent needs to *make himself understood* by providing answers the interpreter can *follow*.

According to Brandom's (1994, see also 2000) model of 'deontic scorekeeping', discursive practice is to be characterized as a game in which various interlocutors keep score of each other's 'deontic status', a status attributed by fellow interlocutors (and oneself) that specifies what one is committed and entitled to, and what one is precluded from, saying or doing, primarily in light of social standards and what one has oneself said or done in the past. The performance of a speech act or an action has the practical significance of making a move in the game of giving and asking for reasons. It changes deontic scores by altering the constellation of further moves one is committed and entitled to make and those that are incompatible with one's commitments. By changing deontic scores performances of speech acts and actions can *provide a reason* for making certain 'committive' or 'permittive' moves at some future moment and for refraining from making others. But they may also *stand in need of reasons* in case they constitute a move to which fellow scorekeepers do not attribute entitlement. Under such circumstances, giving further reasons may be required in order to vindicate one's entitlement. Now the issue at present is this: what is it that cements the game of giving and asking for reasons? What determines the specific inferential relations of commitment, entitlement and incompatibility associated with a speech act or an action?

Principled approaches to action interpretation, as we have seen, appear to take detached, third-person interpretation as the default strategy. Accordingly, people first have to acquire mastery over the principled cognitive routines required for this interpretation strategy. With these procedures in place, they can then successfully engage in reason discourse, by applying them to the allegedly *truncated* answers others tend to give in response to their questions. According to this picture, genuine action explanations always include a proper BD pair. When one or both of the attitudes are left out of the answer, the suppressed partners in motivational crime must be tacitly accounted for by means of underlying cognitive procedures that mirror the structure of formally valid inferences.

This is reminiscent of what Sellars (1953, 313) called 'the dogma of formalism', according to which 'the inference which finds its expression in "it is raining, therefore the streets will be wet" is an enthymeme'. An enthymeme is an informally stated syllogism with an unstated conditional that must be true for the premises to lead to the conclusion. According to this dogma, whenever an inference such as 'it is raining, therefore the streets will be wet' is endorsed, it is because of a belief in the conditional 'If it rains, then the streets will be wet'. With this tacit premise supplied, the inference is an instance of conditional detachment of the formally valid inference 'If it rains, then the streets will be wet. It is raining; therefore the streets will be wet'. On this view, the *psychology* of inference thus has to match the *logical* articulation of inference and the know-how of the proprieties of inference thus requires mastery of the *principles* of logic. Inferences like the above 'it is raining, therefore the streets will be wet' are consequently treated as derivative. Drawing from Sellars's observations, Brandom wants to reverse the order of explanation by treating the latter category as *primitive* and mentioned formal principles as *expressions* of inferences belonging to this category.[10] He proposes to start with *materially* correct inferences, inferences that are correct in virtue of their *content*, not their form. The material inference 'it is raining, therefore the streets will be wet' will normally be treated as correct *because* streets get wet under conditions of rain.

Brandom extends Sellars' account of 'theoretical' material inferences to *practical* ones. As we read Brandom, he treats practical inferences as transitions from reasons to intentions (and resulting actions).[11] Intentions correspond to *practical commitments*: commitments to act. Like their 'doxastic' counterparts, practical commitments can be challenged: scorekeepers can withhold attribution of entitlement to it. By providing reasons, the agent can

vindicate entitlement to his commitment. These reasons then function as antecedents in practical material inferences, with the intention or practical commitment under consideration as their consequent. Consider Brandom's example: 'Only opening my umbrella will keep me dry, so I shall open my umbrella' (1994, 245–9). Brandom objects to the 'Davidsonian' view (in our terminology: 'principled' view) of practical reasoning on which the transition from a reason to an intention must always be *formally valid* in order to count as a genuine practical inference. On such view, the above inference 'only opening my umbrella will keep me dry, therefore I shall open my umbrella' is not considered as a genuine inference by itself. It only *seems* to be an inference, in virtue of a *suppressed* further premise, e.g. 'I want to stay dry'. Brandom, in contrast, proposes to view practical inferences of this form as *materially* good ones. If a scorekeeper endorses an inference of the sort 'only opening my umbrella will keep me dry, therefore I shall open my umbrella' it will normally be in virtue of the fact that he takes part in a social practice in which people open their umbrellas when opening their umbrellas will keep them dry. These inferences are 'material' in the sense that they reflect certain patterns in *a practical world*. Importantly, material inferences do *not* borrow their propriety from formal principles. Following Brandom, we could consider these formal principles to be *explications* of material inferences. As such, the intelligibility of these principles of inference is crucially *dependent* on the propriety of the material inferences they express.

Consider again the practical inference 'only opening my umbrella will keep me dry, so I shall open my umbrella'. This inference expresses a belief in the antecedent and an intention in the consequent. Imagine the situation in which John opens his umbrella going out into the rain, Jack asks him why he does so and John responds that only opening his umbrella will keep him dry. We argue that John's response is *a proper reason explanation* of why he opened his umbrella. No other tacit premise is needed to rationalize his action, e.g. that he does not want to get wet. Now Brandom's example, although it serves its purpose, is not a very representative one. 'Only opening my umbrella will keep me dry, so I shall open my umbrella' is a phrase we never utter in ordinary social practice. Consider the following example: Jack asks Jerry why John is carrying his umbrella. Jerry responds: 'because it's going to rain'. Here Jerry's factive response does not include ascription of either belief or desire. Yet it can still be considered as the antecedent of the practical material inference 'it is going to rain, therefore John should carry his umbrella'. If Jack *follows* Jerry's answer, that is if he endorses the material inference of which it is the antecedent, it will probably be in virtue of a robust pattern in their shared practical world, viz. *that people carry umbrellas when it is going to rain*. Obviously, this material inference does not equal some sort of practical syllogism minus the major premise. But why would it need to? The interesting feature of material inferences is that they are *based on* social practices and as such do not have to meet such formal requirements *before* being endorsed *in* a particular reason conversation.

According to this unprincipled picture of reason discourse, reason discourse about actions is a process of constructing and endorsing practical material inferences, correctness of which is founded in the normative patterns of our everyday social practices and depends on the assessment of their propriety by fellow interpreters (and oneself) in a game of deontic scorekeeping. Propriety of material inferences can be challenged, the reasons provided can be rejected and the corresponding intentions and actions can be withheld attribution of entitlement. Such challenges themselves constitute new moves in the game of giving and asking for reasons, mirroring patterns of their own. Asking for reasons is an attempt to follow the practical patterns that reflect the agent's considerations, giving reasons is an attempt to take the interpreter to those aspects of one's practical world that moved one to act.

5. Action interpretation as deontic scorekeeping

Consider the following conversation:

A: OK bye! See you at the party tonight!
B: No, I'm not coming.
A: Why not?
B: My ex is coming.
A: No he's not, C didn't invite him.
B: Yes she did, D told me.
A: Well, even if he is coming, you can't keep hiding from him.
B: I really don't want to face him right now, OK?!
A: OK. Too bad though, it's going to be a hell of a party.
B: Yeah, too bad. See you tomorrow.
A: Bye.

It is possible to construct something like the following 'practical syllogism' on the basis of this conversation: 'B doesn't want to face her ex right now, she believes that by not coming to the party tonight she will not have to face her ex right now, therefore B should not come to the party tonight'. But as far as one can tell from the conversation itself, this is a cumbersome *reconstruction* of the patterns exhibited in this piece of reason discourse.

We propose to analyze the conversation roughly as follows. B expresses an intention/practical commitment not to come to the party tonight. A withholds attribution of entitlement to B's intention in asking B for her reason. B complies by appealing to the presumed fact that her ex is coming, thereby constructing the material inference 'My ex is coming, so I shall not come to the party tonight'. A does not accept B's answer, she challenges the propriety of A's inference by withholding entitlement to the antecedent: B's doxastic commitment that her ex is coming to the party tonight. B tries to vindicate entitlement to this doxastic commitment that her ex is coming to the party tonight by invoking the authority of D, thereby attempting to justify her original intention. Moving away from this little doxastic dispute, A presents a new challenge and proposes a material inference that expresses B's practical commitment to the contrary: B cannot keep hiding from her ex, so she should come to the party tonight. In another attempt at vindicating her intention not to come, B constructs a new material inference: 'I really don't want to face him right now, therefore I shall not come to the party tonight'. A accepts this answer; she endorses the material inference expressed by it and finally attributes entitlement to B's commitment not to come to the party tonight.

The temporally extended psychological processes underlying this conversation can properly be described in terms of consecutive construction, rejection and endorsement of material inferences. Even though it is possible to logically reconstruct a practical syllogism from the raw materials of conversation, *this does not mean that it has any psychological reality* in ordinary reason discourse.

Now suppose E meets up with A and asks her whether B is coming to C's party tonight.

E: Is she coming to the party tonight?
A: No, she thinks her ex is coming. She really doesn't want to face him right now.
E: I see.

In this scenario, A clearly expresses both a belief and a desire she attributes to B. Doesn't this show that reason discourse can, at least sometimes, express a principled routine of BD integration? First of all, it should be clear that this conversation does not

have the ingredients for the construction of the *proper* BD pair. Consider how unnatural it would be if A would answer: 'No, she thinks that by coming to the party she will have to face her ex and she doesn't want to face him right now'. So A's answer does not meet the strict demands of BD psychology. Second, our claim is not that reasons can *never* include both a belief and a desire, only that they *need* not and in fact often *do* not. There are no constraints on the material inferences proposed and endorsed *outside* the context of a particular instance of reason discourse, but within such a context it may require quite an elaborate answer to show the patterns that explain the agent's actions. There may be huge individual differences and some people may find it more amusing to express themselves formally than others. What is important is that there is no principle that can be invoked in order to settle these issues and mark the constraints of reason interpretation *tout court*.

But of course there are a lot of *contextual* constraints. The interpreter must be able to follow the agent if the latter wants to be attributed entitlement to his action. Whether the answer he provides counts as a proper reason depends on his deontic score, as kept by the interpreter. The interpreter needs to endorse the material inference proposed by the agent. But this is only possible against the agent's 'deontic background': the agent's world as carved out by his commitments and entitlements, reflecting his beliefs, desires and intentions. Here it may seem as if assessing propriety of the proposed material inference must involve the *integration* of the material inference with the agent's background beliefs, desires and intentions. Such integration process, in turn, would require just those principled procedures of BD psychology we want to avoid.

But this is to misconstrue the nature of material inferences. As stated, material inferences mirror normative patterns in the agent's practical world. Such patterns cannot exist *in vacuo*, however: they are instantiated *in* the agent's world and *in* his practices. The scorekeeper needs to see the pattern mirrored by the material inference proposed by the agent in order to assess its propriety and its status as a proper reason explanation, but he cannot do this without some degree of familiarity with the practical world *in which* this pattern is instantiated. Deontic scorekeeping is a process of such familiarization: by keeping track of the agent's commitments and entitlements the scorekeeper structures and re-structures the agent's world so that new patterns can be distinguished in it, reflecting newly proposed material inferences in reason discourse.

This is to say that material inferences have no *ceteris paribus* clauses.[12] Material inferences reflect patterns *in* the agent's world and his practices. These patterns are unprincipled because they are *local*, bound to a *particular* world. An 'all else being equal' clause is only called for if it is assumed that in making sense of an action, we must generalize over *different* worlds with *similar* patterns. Now it might seem that we simply must look past the differences between the worlds of individual agents. For how else would we be able to interpret agents without great familiarity with each and every one of them? This line of thought again reveals a deep-seated detached view on folk psychological practice. To repeat our main message from section 3: What has been left out is the fact that agents and interpreters *share* much of their practical worlds, especially if they belong to the same socio-cultural group. The average person is indeed able to follow many of the patterns exhibited by the actions and answers of people they hardly know. These patterns are instantiated in a practical world in which *both* the agent and the scorekeeper live. It is a shared world, but a *particular* one nonetheless.

People can start off in the game of giving and asking for reasons on the simple assumption that contexts are shared and proceed as far as it goes. Of course, there are many situations in which scorekeepers do not share contexts to this degree. But why automatically

assume that these differences in context need to be tacitly accounted for? The point at which the agent's action does not follow a pattern the scorekeeper can distinguish is easily discerned: it is the point where the scorekeeper starts asking him for his reasons. And it is normally also quite clear whether or not the material inference proposed by the agent highlights such a pattern: if it does not, the scorekeeper will probably start asking for further reasons until the agent's answer does hit on a pattern the scorekeeper can distinguish. Answers given may reflect back on a shared world, but they may also go beyond it and highlight patterns specific to the agent's practices or those of the socio-cultural group to which he belongs. Our proposal accounts for the context sensitivity of action interpretation by characterizing it in terms of inferential practices that do not rely on *ceteris paribus* considerations. Material inferences can do the work that *ceteris paribus* 'practical syllogisms' are supposed to do on principled accounts, because the interpreter *puts the agent to work* when the material inference proposed by the agent *fails to work* for the interpreter. Material inferences do not have to be *applied to* a particular context, for they are again and again *formulated within* a particular context.

Successfully engaging in reason discourse in an appropriate manner crucially depends on the capacity to keep track of each other's deontic scores. The point we want to make here is that this is not a puzzle the interpreter need to solve. To the extent that we perform our actions in a shared practical world, there is no need to peer into the minds of fellow agents in order to understand their answers. And to the extent that their actions follow patterns of personalized extensions of this common ground, we can simply follow their lead and let them make themselves understood.

At the end of this section, let us briefly address cases of third-person interpretation and insincere reason discourse. We argued that action interpretation in terms of reasons for action is best conceived as the ability to construct, reject and endorse material inferences that mirror patterns in our (shared) practical worlds. Extending our account to third-person contexts, we suggest that making sense of others in terms of reasons for action likewise consists in the construction of practical material inferences within a deontic context. As argued, this context may be a default background that is shared by the scorekeeper, in which case the scorekeeper can simply use his own view on the world in order to make sense of the agent. Consider a man running to get on a just-departing train. According to Scholl and Leslie (1999), bystanders would interpret the man as 'an intentional agent, who believes that there is a just-departing train, and who wants to get on it' (131). *Pace* Scholl and Leslie, we think that interpreting the man's reason in such non-puzzling cases is simply a matter of identifying the fact to which the man is responding to, viz. that the train is just departing. In virtue of their sharing a practical context with the man, bystanders can simply look out into the world and propose a material inference for the man in question that they may or may not endorse: the train is just departing, so he should run to try to get on it.

But the agent's deontic context may also differ from the scorekeeper's own background of practical reasoning in important ways relevant for the current interpretation process. It thus has to be updated accordingly. Again, this updating process is not something we normally have to do by ourselves; the agent (or his socio-cultural peers) will have done it for us throughout our (conversational) history together. In such cases, 'mindreading' in fact consists in keeping track of people's practical worlds on the one hand, their deontic scores that form the ingredients for future material inferences to be proposed, endorsed or challenged, and constructing a particular practical material inference within this context on the other. On this proposal, our capacity to interpret other people's actions in terms of reasons as bystanders can be regarded as a derivative capacity, an extension of our interpretative skills in second-person encounters. As such, it can be accounted for in a similar fashion.

What about truly puzzling third-person cases in which speculation is the only option? What about second-person cases in which people are lying or keeping hidden agendas? Such second-person cases could also be considered as instances of reason speculation, perhaps even more so than third-person interpretation: it is speculation *despite* appearances. Our suggestion is again to view these interpretation skills as an extension of and derived from those involved in sincere reason discourse. They may require quickly shifting between different hypothetical deontic contexts, each one with a certain degree of plausibility given the agent's performances in the past, his personal and socio-cultural background and the reports of others. But interpretation, we suggest, still comes down to the construction and endorsement of material inferences within such deontic contexts.

At this point, some may want to regard our focus in second-person contexts of reason discourse as a heuristic tool, a kind of Wittgensteinian ladder: once the unprincipled alternative to the BD model is in clear view, the means of getting there become irrelevant. Although this conclusion would serve our general aims in this paper quite well, there is a conceptual and an empirical reason why we want to resist it. The conceptual reason is that we do not see how the notion of a material inference can be fleshed out in abstraction from the practice of giving and asking for reasons. The empirical reason is that focus on second-person contexts puts forward an interesting and in our view plausible developmental hypothesis, viz. that the normal developmental route for children is to acquire their high-level interpretation skills by being pulled up, by their caregivers, into the game of giving and asking for reasons and by being taught what counts as a reason for what relative to one's deontic scores. Unfortunately, we will have to leave it at this. Save to say that this hypothesis shows obvious parallels with Hutto's (2008a, 2008b) 'narrative practice hypothesis', according to which children normally learn to master the skills of interpretation by engaging with their caregivers in folk-psychological story-telling (about agents and their reasons).

6. Conclusion and comparison

What do people do when they interpret other people in terms of reasons for action? According to the widely endorsed BD model of both TT and ST varieties, interpreting other agents in terms of their reasons for action requires the reconstruction of the action in terms of proper practical syllogisms in the form of appropriately structured BD pairs. We have argued that such principled approaches to action interpretation rest on the implicit assumption that the principles of formally valid practical inferences need to be applied to an inferred context of action. This rendering of action interpretation is motivated by an exclusive focus on detached, spectatorial interpretation strategies. Such focus, we argued, is unwarranted. To a large extent, practical contexts are shared between interpreter and agent. And when the agent's performances start to puzzle the interpreter, the most efficient and effective strategy for the interpreter to adopt is to start a conversation and ask the agent why. Building on insights from Sellars and Brandom, we proposed an alternative descriptive account of the psychology behind reason discourse that does not consist in principled routines of BD integration. Successfully participating in the game of giving and asking for reasons can be characterized as a process of undertaking and attributing commitments and (withholding) attribution of entitlements by constructing, endorsing or rejecting material inferences, all against the background of the deontic scores of the participants as these are continuously being updated by the history of speech acts and actions performed. These deontic backgrounds reflect the participants' particular worlds, shared to a large

extent, but personalized in the course of interaction. Mirroring patterns *in* particular practical worlds, and being proposed *by* agents upon demand, material inferences do not have to be applied *to* agents. Starting from the second-person practice of giving and asking for reasons, our account of action interpretation thus bypasses the principled cognitive routines required for BD integration.

Wrapping up, we would like to highlight some similarities with other non-standard accounts of action interpretation found in the literature. As indicated above, our account shows similarities with Hutto's (2008a, 2008b) 'narrative practice hypothesis'. Apart from his unnecessary adoption of the BD-model,[13] we think our accounts can work complementary. We are furthermore open to the suggestion that deontic contexts are structured in a narrative form.

We are enthusiastic about the suggestion that reason interpretation is often more accurately characterized as the regulative or educational (rather than descriptive) practice of (re)establishing norms of social interaction (McGeer 2007). In such cases, interpreting other people's actions is not a matter of passively 'reading' minds, but of actively *shaping* them (Zawidzki 2008). Our account leaves room for this focus on the normative aspect of reason interpretation. In fact, it is an inherent feature of the practice of giving and asking for reasons as we construe it: a practice in which people undertake, attribute and withhold commitments and entitlements.

Acknowledgement

We would like to thank Marc Slors, Dan Hutto and Bob Gordon for their helpful comments on earlier versions of this paper.

Notes

1. Here and throughout, we use the term 'interpretation' in a broad sense to include all kinds of sense-making activities, such as understanding, explanation and prediction. As will become clear, though, we resist focus on explanation and prediction in third-person contexts.
2. Cf. Botterill (1996, 115): 'If belief-desire psychology has a central principle, it must link belief, desire and behavior. It could be formulated like this: [*Action Principle*] An agent will act in such a way as to satisfy, or at least to increase the likelihood of satisfaction, of his/her current strongest desire in light of his/her beliefs'.
3. Fodor (1992, 283), for example, suggests that 'normal cognitive development eventuates in the child's internalization of a tacit "metacognitive" intentional psychology: specifically, in the internalization of some version of the folk psychological theory that an agent's behavior is normally caused by his beliefs and desires'. Another example can be found in Leslie, German and Polizzi (2005, 50), who state that 'a prediction of behavior question requires an additional ascription of desire, the integration of belief and desire, and the inferring of a resulting action'.
4. See, for example, Horgan and Woodward (1985, 197).
5. Some theory theorists explicitly restrict folk psychological theory to *non*-tacit knowledge of folk psychological principles represented in a mixed variety of simpler or more complex abstract *models* that, by their very nature, do not contain *ceteris paribus* clauses (e.g. Maibom 2003). The context-sensitive application of the models is delegated to other cognitive mechanisms. Yet the worries concerning complexity we raise in section three also apply to the workings of these additional cognitive mechanisms.
6. Thus, Goldman (2006, 29), for example, thinks that 'a decision-making mechanism normally takes genuine (nonpretend) desires and beliefs as inputs and then outputs a genuine (nonpretend) decision. In simulation exercises, the decision-making mechanism is applied to pretend desires and beliefs and outputs pretend decisions'. In a somewhat similar fashion, Currie (1995, 158) claims that in simulating another agent 'we tend to acquire, in imagination, the beliefs and desires an agent would most likely have in that situation, and those imaginary beliefs and desires have consequences in the shape of further pretend beliefs and desires as

well as pretend decisions that mimic the beliefs, desires and decisions that follow the real case'. These ideas are made explicit in Nichols' and Stich's boxological model of ST (e.g. Stich and Nichols 1992, 1995, 1997; Nichols and Stich 2003).

7. If, however, one believes that the simulation routine requires mastery over the concepts of the propositional attitudes, then, as Ravenscroft (2003) argued, cognitivist ST may be committed to tacit knowledge of FP principles after all. In the case of (simulated) practical reasoning, different kinds of propositional attitude concepts are involved (beliefs, desires, decisions). One needs to keep track of these concepts in order to reach a practical conclusion, that is: one needs tacit knowledge specifying the way these attitudinal 'tags' are related to each other and the contents. But how can we explain our ability to tag and keep track of these different propositional attitudes? Tacit knowledge of FP principles seems to fit the job description perfectly.
8. Wilkerson (2001) makes a point against TT and ST that is inspired by similar considerations.
9. It should be clear that worries about a 'threat of collapse' of our proposal onto principled accounts are wrong-headed. First, the tacit knowledge (if there is such) required for co-cognition does not require knowledge concerning psychological generalizations (Davies and Stone 2001). In the case of practical co-reasoning, it only does when one assumes that such co-cognition requires 'tagging' the relevant contents with psychological concepts (see Ravenscroft 2003, footnote 7). Our proposal, however, is to the effect that such practical co-reasoning on default takes place in a shared context, which is to say that the contents are precisely *not* specified in terms of propositional attitudes. Secondly, and more importantly, on our proposal, the practical co-reasoning does *not* mirror *formally* valid inferences, but consists in proposing, endorsing or rejecting *materially* correct ones (see section 4). On our unprincipled account of interpretation, the interpreter's 'knowledge' consists in *knowing how* to propose, endorse or reject such material patterns of inference. It explicitly does not involve (tacit) knowledge regarding the principles of reasoning.
10. Brandom treats the formal principles of inference as expressive tools that enable us to make explicit the proprieties of material inferences, by *saying* what was *done* in endorsing the inference. Thus, the conditional 'whenever it rains the streets will be wet' is considered as a means of saying what is being done in endorsing the material inference 'it is raining, therefore the streets will be wet'.
11. Here and throughout, we assume that intentionally performing an action at t is typically the reliable result of an intention to perform an action of that type at or prior to t. Explaining the intention with which one performs a certain action is therefore normally sufficient to explain the action. Brandom treats actions as 'language exit transitions': (reliable dispositions to respond differentially to the) acknowledging of commitments by bringing about various kinds of states of affairs (see 1994, 235).
12. This may be confusing for those familiar with Brandom's work. In claiming that material inferences do not have *ceteris paribus* clauses, we aim at a common conception of such clauses on which they are supposed to *secure* the monotonicity of the statements to which they belong. Brandom proposes to understand *ceteris paribus* clauses 'as explicitly marking the nonmonotonicity of an inference, rather than a deus ex machina that magically *removes* its nonmonotonicity' (2000, 88). We obviously do not want to object to this alternative conception of *ceteris paribus* here: material inferences being essentially contextual is precisely what their non-monotonicity amounts to.
13. According to Hutto's NPH, infants become familiar with the background norms for wielding folk psychology in practice by being exposed to 'folk-psychological narratives'. The defining feature of these narratives is that they reveal how beliefs and desires (and other propositional attitudes) interrelate and conspire to form reasons for action. For, Hutto says, 'it is not enough to imagine it as being sponsored by a singular kind of propositional attitude; one must also ascribe other kinds of attitudes that act as relevant and necessary partners in motivational crime' (2008a, 26).

Notes on contributors

Leon C. de Bruin is a Post-doctoral Fellow at the Ruhr-University of Bochum. His research interests are primarily in philosophy of mind and philosophy of psychology. He currently works on a developmental theory of intention attribution and cross-cultural differences in folk psychology.

Derek W. Strijbos is a PhD student at the Radboud University Nijmegen. His main research interests are in the philosophy of mind, philosophy of action and philosophy of psychiatry. He is currently working on meta-representation in social cognition.

References

Bermudez, J. 2004. The domain of folk psychology. In *Mind and persons*, ed. A. O'Hear, 25–48. Cambridge: Cambridge University Press.

Borg, E. 2007. If mirror neurons are the answer, what was the question? *Journal of Consciousness Studies* 14, no. 8: 5–19.

Botterill, G. 1996. Folk psychology and theoretical status. In *Theories of theories of mind*, ed. P. Carruthers and P. Smith, 184–99. Cambridge: Cambridge University Press.

Brandom, R.B. 1994. *Making it explicit*. Cambridge, MA: Harvard University Press.

Brandom, R.B. 2000. *Articulating reasons: An introduction to inferentialism*. Cambridge, MA: Harvard University Press.

Currie, G. 1995. Imagination and simulation: Aesthetics meets cognitive science. In *Mental simulation: Evaluations and applications*, ed. M. Davies and T. Stone, 151–69. Oxford: Blackwell.

Davidson, D. 1963. Actions, reasons, and causes. *Journal of Philosophy* 60, no. 23: 685–700. Repr. in Davidson 2001.

Davidson, D. 2001. *Essays on actions and events*. New York: Oxford University Press.

Davies, M., and T. Stone. 2001. Mental simulation, tacit knowledge and the threat of collapse. *Philosophical Topics* 29, nos. 1/2: 127–73.

Fodor, J. 1992. A theory of the child's theory of mind. *Cognition* 44, no. 3: 283–96.

Gallagher, S. 2004. Understanding interpersonal problems in autism: Interaction theory as an alternative to theory of mind. *Philosophy, Psychiatry, and Psychology* 11, no. 3: 199–217.

Gallagher, S., and D. Hutto. 2008. Understanding others through primary interaction and narrative practice. In *The shared mind: Perspectives on intersubjectivity*, ed. J. Zlatev, T. Racine, C. Sinha and E. Itkonen, 17–38. Amsterdam: John Benjamins.

Goldman, A. 1989. Interpretation psychologiz. *Mind and Language* 4, no. 3: 161–85.

Goldman, A. 2006. *Simulating minds: The philosophy, psychology and neuroscience of mindreading*. New York: Oxford University Press.

Gopnik, A., and A.N. Meltzoff. 1997. *Words, thoughts, and theories*. Cambridge, MA: MIT Press.

Gordon, R. 1986. Folk psychology as simulation. *Mind and Language* 1, no. 2: 158–71.

Gordon, R. 1995. Simulation without introspection or inference from me to you. In *Mental simulation*, ed. M. Davies and T. Stone, 53–67. Oxford: Blackwell.

Gordon, R. 1996. 'Radical' simulationism. In *Theories of theories of mind*, ed. P. Carruthers and P. Smith, 11–21. Cambridge: Cambridge University Press.

Gordon, R. 2000. Sellars's Ryleans revisit. *Proto Sociology* 14: 102–14.

Heal, J. 1986. Replication and functionalism. In *Language, mind and logic*, ed. J. Butterfield, 135–50. Cambridge: Cambridge University Press.

Heal, J. 1996. Simulation, theory and content. In *Theories of theories of mind*, ed. P. Carruthers and P. Smith, 75–89. Cambridge: Cambridge University Press.

Heal, J. 1998. Co-cognition and off-line simulation: Two ways of understanding the simulation approach. *Mind and Language* 13, no. 4: 477–98.

Horgan, T., and J. Woodward. 1985. Folk psychology is here to stay. *Philosophical Review* 94, no. 2: 197–226.

Hutto, D.D. 2004. The limits of spectatorial folk psychology. *Mind and Language* 19, no. 5: 548–73.

Hutto, D.D. 2008a. *Folk psychological narratives: The sociocultural basis of understanding reasons*. Cambridge, MA: MIT Press.

Hutto, D.D. 2008b. The narrative practice hypothesis: Clarifications and implications. *Philosophical Explorations* 11, no. 3: 175–92.

Iacoboni, M. 2005. Understanding others: Imitation, language, empathy. In *Perspectives on imitation: From cognitive neuroscience to social science: Mechanisms of imitation and imitation in animals*, ed. S. Hurley and N. Chater, Vol. 1, 77–100. Cambridge, MA: MIT Press.

Leslie, A.M., T. German, and P. Polizzi. 2005. Belief–desire reasoning as a process of selection. *Cognitive Psychology* 50, no. 1: 45–85.

Maibom, H. 2003. The mindreader and the scientist. *Mind and Language* 18, no. 3: 296–315.

McGeer, V. 2007. The regulative dimension of folk psychology. In *Folk Psychology Re-assessed*, ed. D. Hutto and M. Ratcliffe, 137–56. Dordrecht: Springer.

Nichols, S., and S. Stich. 2003. *Mindreading. An integrated account of pretence, self-awareness, and understanding of other minds*. Oxford: Clarendon Press.

Ravenscroft, I. 2003. Simulation, collapse and Humean motivation. *Mind and Language* 18, no. 2: 162–74.

Scholl, B.J., and A.M. Leslie. 1999. Modularity, development and 'theory of mind'. *Mind and Language* 14, no. 1: 131–53.

Sellars, W. 1953. Inference and meaning. *Mind* 62, no. 247: 313–38.

Slors, M.V.P. 2007. Intentional systems theory, mental causation and empathic resonance. *Erkenntnis* 67, no. 2: 321–36.

Spaulding, S. 2010. Embodied cognition and mindreading. *Mind and Language* 25, no. 1: 119–40.

Stich, S., and S. Nichols. 1992. Folk psychology: simulation or tacit theory? *Mind and Language* 7, no. 1: 35–71.

Stich, S., and S. Nichols. 1995. Second thoughts on simulation. In *Mental simulation: Evaluations and applications*, ed. M. Davies and T. Stone, 87–108. Oxford: Blackwell.

Stich, S., and S. Nichols. 1997. Cognitive penetrability, rationality and restricted simulation. *Mind and Language* 12, nos. 3/4: 297–326.

Trevarthen, C. 1979. Communication and cooperation in early infancy: A description of primary intersubjectivity. In *Before speech: The beginning of interpersonal communication*, ed. M. Bullowa, 321–47. Cambridge: Cambridge University Press.

Trevarthen, C., and P. Hubley. 1978. Secondary intersubjectivity: Confidence, confiding and acts of meaning in the first year. In *Action, gesture and symbol: The emergence of language*, ed. A. Lock, 183–229. London: Academic Press.

Wilkerson, W.S. 2001. Simulation, theory and the frame problem: The interpretative moment. *Philosophical Psychology* 14, no. 2: 141–53.

Zawidzki, T. 2008. The function of folk psychology: Mind reading or mind shaping? *Philosophical Explorations* 11, no. 3: 193–210.