

[Manuscript version. Final paper to appear in S. Sedivy (ed.) *Art, Representation and Make-Believe: Essays on the Philosophy of Kendall L. Walton*. Routledge.]

The story of the ghost in the machine

Adam Toon
University of Exeter
a.toon@exeter.ac.uk

1 Introduction

Once upon a time, many philosophers might have followed Descartes in thinking of the mind as a distinct substance to the body—a spirit, soul, perhaps, or mind-stuff. Lurking somewhere deep within us, this strange and wondrous substance was the repository of our beliefs, desires, hopes and fears. When our bodies moved about the world—placing us in front of a sunset, for example, or having us bite into a rotten apple—changes resulted in this inner world—thoughts about beauty, perhaps, or feelings of disgust. In turn, changes within this inner substance might bring about changes to our body—we might resolve to linger a while in front of the view or throw the apple away in disgust. Despite these frequent interactions with the outside world, however, our inner world was quite unlike the material stuff that makes up atoms, tables or the human body. And of course, unlike our bodies, it might also hope to live happily ever after.

Nowadays, most philosophers find the existence of any kind of non-physical substance a bit too spooky to believe. Instead, most will locate the mind in a particular part of the material world—albeit an especially marvellous and complex part—namely, the human brain. And yet for many authors, the basic structure of the Cartesian picture of the mind

remains intact. The mind is still taken to be a kind of inner world that houses our mental states—our beliefs, desires, hopes, fears and the rest. The state of our bodies can affect this inner world and, in turn, this inner world can direct the movement of our bodies. Of course, since our inner world is now resolutely material, these interactions seem rather less mysterious than before. But the basic picture of the mind as an inner realm is retained. For this reason, such a view is often described—typically by its opponents—as *neo-Cartesianism*.

Gilbert Ryle's *The Concept of Mind* is a wonderfully irreverent assault upon the Cartesian view of the mind, which Ryle memorably dubbed “the dogma of the ghost in the machine” (1949, p. 17). According to Ryle, Cartesianism is absurd. It cannot make sense of our ordinary talk about the mind and it renders even our humdrum, everyday knowledge of the thoughts and feelings of others (and ourselves for that matter) utterly mysterious. Ryle's critique is normally thought merely to tell against dualism. And yet many of his arguments can also be levelled against neo-Cartesianism. For Ryle's main complaint against Cartesianism is not that it takes the mind to be non-physical, but that it takes it to be something that is hidden away somewhere inside us—locked away in some “secret grotto” (Ryle 1949, p. 114; see Tanney 2009). It makes little difference whether we take this inner grotto to be made of mysterious soul-stuff or wonderfully intricate grey matter.

For Ryle, Cartesianism is a “philosopher's myth” (1949, p. 17). It is a category mistake that philosophers have inflicted upon our ordinary language for talking about the mind. In this paper, I want to suggest an alternative view. I will argue that our picture of the mind as an inner world is not a myth, but a story. It is a story that we tell in order to make sense of people and the way they behave. Like Ryle, I think that the story is not true. Indeed, I think it cannot be true. Like much science fiction or fantasy, when we examine its plot too closely we find that it stretches our credulity and even harbours latent contradictions. And yet, I shall argue, it is not merely a story told by philosophers and foisted upon the unsuspecting folk. Instead, the story of the mind as an inner world lies at the heart of our ordinary talk about people and their mental states. Indeed, it is a story that we cannot

avoid telling.

The approach I shall defend is a form of *mental fictionalism*. There is now a burgeoning debate over the prospects for a fictionalist approach to the mind (e.g. Demeter 2013a; Toon, 2016; Wallace 2007, 2016). This debate has shown that any such approach faces serious challenges.¹ My aim in this paper is to defend a specific form of mental fictionalism and show how it might overcome some of these challenges. To do so, I will draw on Kendall Walton’s hugely influential work on make-believe (1990, 1993). In earlier work, I have drawn on Walton’s ideas to develop a theory of scientific models (Toon 2010, 2011, 2012). It is characteristic of scientific models, I have argued, that they represent the world by asking us to imagine it as other than it is—a surface is imagined to be free of friction, the earth is imagined to be a perfect sphere, and so on. In a similar vein, I now want to suggest that much of our ordinary talk involves imagining people as

¹ In light of these challenges, few authors have been willing to endorse mental fictionalism. An important exception is Tamás Demeter (e.g. 2013b), who views folk psychology as a tool for expressing an interpreter’s affective reactions. Closer to my own view is a recent paper by Bill Lycan (forthcoming). Like me, Lycan defends a form of mental fictionalism that draws on Ryle’s work, although he rejects appeals to make-believe or Walton’s framework. (Lycan also presents his view in a “fictionalist” spirit and does not endorse it.) Although not a fictionalist account, Derek Melser’s *The Act of Thinking* (2004) includes a fascinating discussion of the role of metaphor in our ordinary talk about the mind. Melser argues, as I do, that the notion of the mind as an inner world arises from taking our metaphors too literally. His discussion focuses especially on the noun “mind”, which he claims to be a nominalisation of the old-fashioned verb “to mind”. An important difference between Melser’s view and my own is that, while Melser claims that the relevant metaphors are a “mess” and should be put aside, I argue that are indispensable and lie at the heart of folk psychology. (I am grateful to Bill Lycan for drawing Melser’s work to my attention.)

other than they are: we imagine them to have mysterious inner worlds that harbour their beliefs, desires, feelings and intentions. Like models that invoke perfect spheres or frictionless planes, thinking about the mind as if it were an inner world is enormously useful, even indispensable. And yet we ought not to take it too seriously. As we shall see, however, applying this idea to our talk about the *mind*—rather than talk about planets or inclined planes—is far from straightforward. For it seems to require us to rethink central aspects of Walton’s framework itself, such as the notions of imagination and make-believe.

2 The ghost as a myth

Why does Ryle regard Cartesianism as a myth? His arguments are notoriously difficult to summarise, but Julia Tanney (2009) helpfully distinguishes three main lines of attack: ontological, epistemological and semantic. The ontological worry is most familiar: if mind and body are distinct substances, how do they interact? The epistemological worry arises because Cartesianism takes the mind to be private: if others’ minds are hidden away inside them, how can we ever know what they are thinking? No doubt people do sometimes keep their thoughts to themselves. And yet Cartesianism threatens to rob us of *any* knowledge of the beliefs, desires or emotions of others—even of the most humdrum sort. The semantic worry is similar. According to Cartesianism, the words we use to talk about others’ minds refer to things that are hidden from us, so that we can never know whether we have applied them correctly or not. And yet, Ryle argues, it is precisely because we *can* use such terms meaningfully that questions about the nature of the mind arose in the first place. It is only because we regularly identify people’s behaviour as intelligent, for example, that we find ourselves asking how people differ from mere machines (1949, p. 22).

Ryle devotes surprisingly little time to the ontological difficulties that beset Cartesianism. Instead, most of his critique focuses on its epistemological and semantic commitments. His principal aim is to show that Cartesianism cannot do justice to our ordinary

knowledge of the mind and the way that we commonly talk about it. As Tanney (2009) points out, Ryle's arguments against Cartesianism depend not so much upon the ghostliness of its conception of the mind, but its *hiddenness*. It is because Cartesianism sees the mind as a private realm locked away inside us that Ryle believes that it fails to account for our ordinary practices. As a result, many of Ryle's criticisms also apply to neo-Cartesianism. Neo-Cartesians reject dualism and locate the mind within the brain, rather than within any mysterious non-physical substance. Although this might avoid the ontological problems of Descartes' own view, it also threatens to make the minds of others hidden to us—at least in practice. After all, when we make judgements about what people think or want or feel, we aren't normally in a position to sneak a quick peek inside their skulls first.

What is the alternative to Cartesianism? Ryle's own position is difficult to characterise. He is often taken to be an analytical behaviourist. And yet in many respects this characterisation is misleading, not least because Ryle never attempts to reduce talk about the mind to talk about behaviour. In fact, in later work, Ryle describes both behaviourism and Cartesianism as “category-howlers”—although he admits that he might be taken to have “one leg and one ear” in the behaviourist camp (1979, p. 31). In Ryle's characterisation, the behaviourist is a *Reductionist* whose slogan is “Nothing But...” (*ibid*, p. 79). She might insist, for example, that thinking is nothing but talking to ourselves. By contrast, Ryle tells us, the Cartesian is a *Duplicationist*. Her motto is “Something Else as Well...” (*ibid.*) Against the behaviourist, the Cartesian quite rightly points out that thinking cannot merely be talking to ourselves, since the very same sequence of words might be uttered intelligently or merely parroted. But the Cartesian wrongly concludes from this fact that the difference between the two—between talking intelligently and merely parroting something—must lie in some hidden ghostly utterance that has remained hitherto unseen.

The solution, Ryle insists, is to avoid both Reductionism and Duplicationism. Talk about the mind is neither to be reduced to talk about behaviour nor taken to describe some inner mechanism—either a ghostly or non-ghostly one. Instead, the philosopher's task is to

describe our ordinary talk about the mind as she finds it—to chart its “logical geography” (1949, p. 17)—without mistakenly trying to shoehorn it into the mould of some entirely different domain of discourse, such as talk about the movements of muscles or machines. In a similar vein, *The Concept of Mind* rejects not only dualism but also materialism and idealism. Once we see that “mind” and “body” do not belong to the same logical categories (e.g. categories of “substance”, “stuff”, “process” etc.), we see that we must reject, not only the idea that mind and body are *distinct* substances (dualism), but also the idea that we can collapse mind into body (materialism) or body into mind (idealism). Instead, Ryle argues, we should recognise that

[it] is perfectly proper to say, in one logical tone of voice, that there exist minds, and to say, in another logical tone of voice, that there exist bodies. But these expressions [...] indicate two different senses of ‘exist’, somewhat as ‘rising’ has different senses in ‘the tide is rising’, ‘hopes are rising’ and ‘the average age of death is rising’ (1949, p. 24)

Not all have found Ryle’s attempt to dissolve the mind-body problem in this manner satisfying, however. Let us now turn to see how fictionalism differs from Ryle’s approach.

3 The ghost as a story

According to Ryle, Cartesianism is a category mistake that is mistakenly imposed upon folk discourse by misguided philosophers. Descartes himself was led to this “para-mechanical” conception of the mind, Ryle suggests, in an attempt to emulate the successes of the new science of mechanics, while avoiding the implication that human beings were mere clockwork (1949, p. 20). And yet, Ryle argues, however laudable Descartes’ intentions were—and however productive his myth might once have been—its ghost should now be exorcised. In his review of *The Concept of Mind*, Stuart Hampshire objects to Ryle’s account of the ghost’s origins:

so far from being imposed on the plain man by philosophical theorists, and

even less by seventeenth century theorists, the myth of the mind as a ghost within the body is one of the most primitive and natural of all the innumerable myths which are deeply imbedded within the structure of our languages. [...] [Ryle] represents philosophers as corrupting the literal innocence of common sense speech with alien metaphors. In this he not only greatly exaggerates the influence of philosophers [...] but (more seriously) neglects the fact, or rather the necessity, that the forms of common speech and its modes of description should be permeated with such metaphors [...] (1950, 237; emphasis in original)

As I will present it here, mental fictionalism can be seen as an attempt to follow up on Hampshire's insight. For the fictionalist, the notion of the mind as an inner realm is not merely a philosophers' myth. On the contrary, it sits right at the heart of our ordinary language as a set of central metaphors that we use for talking about the mind (cf. Melser, 2004, esp. Chapters 8 and 9).

We can develop this idea by drawing on Walton's well-known analysis of metaphor and figurative language. This analysis depends upon the notion of a game of make-believe. Walton calls the objects used within a game of make-believe *props* and its rules *principles of generation* (1990). If children are playing a game of pirates, their props might be a cardboard box and a pile of stones. The principles of generation might include rules to the effect that sitting in the box counts as being on board ship or that the stones are gold doubloons. Together the props and principles of generation create the content of the game; they make propositions *fictional*. If Anna is sitting in the box, then it is fictional that she is on board ship; if David hides the pile of stones, then it is fictional he has hidden the gold; and so on. One important feature of children's games is that the children participate verbally too. If Anna shouts "I've got the gold" then, in one sense, she is only pretending: she knows perfectly well that the stones are only stones, not gold. And yet, by pretending in this way, Anna does make a genuine assertion about the state of the props in the game: she claims that she has hold of the all-important pile of stones. If David were holding the stones instead, then Anna's claim would be false.

In cases of metaphor, we talk about one topic (our *primary domain*) in terms taken from another area (our *secondary domain*). In Walton's analysis, the primary and secondary domains are brought together by pretence within a game of make-believe. For example, if we say "David is at a crossroads", we draw upon a familiar game in which we pretend that someone's life is a journey. In this game, we imagine events in someone's life (e.g. important decisions, difficulties they encounter) to be stages in a journey (e.g. crossroads, steep paths to be climbed). In Walton's terminology, the events in someone's life are props and the rules that link our primary and secondary domain are principles of generation. Our utterance is an act of pretence within this game. In this context, when we say "David is at a crossroads" we do not claim that he is literally standing at a road junction. We only pretend to assert this, just as Anna only pretended to have hold of a pile of gold doubloons. And yet, much like Anna, we also indicate that our pretence is appropriate and thereby make a genuine assertion about our props: we claim that David must make an important decision. It is this fact about David's life that makes our pretence appropriate within the game. If David were not faced with an important decision, our pretence would be inappropriate and our assertion would be false.

We can now turn to talk about the mind as an inner world. The fictionalist claims that such talk is fundamentally metaphorical: we make sense of the mind by drawing on our understanding of other domains. As Hampshire points out, one such domain is ordinary physical objects or processes: we talk as if our thoughts and feelings "push" or "pull" us in certain directions, for example. It is this sort of talk that characterises what Ryle disapprovingly calls a "wires and pulleys" conception of the mind. Another domain that we use to make sense of the mind is the realm of public representations—most notably, spoken and written language. We often talk about the mind as if it were a private world of inner sentences or pictures that capture our thoughts, feelings or intentions—a hidden realm of representations that have content or meaning, can be true or false, accurate or inaccurate, and so on. Think of how we talk about our ideas as if they were little pictures in our heads or describe our reasoning as if it we had followed the steps of an inner argument from premises to conclusion. Or consider memory and standing belief. We often

talk about memory as if it were a kind of private, inner notebook that we keep tucked away inside our heads. As someone goes about the world, we can imagine them jotting down important information in their inner notebook. Later, when they need this information, we can imagine them looking it up in this notebook and acting accordingly. In these and many other ways, I suggest, we use the world of external representations as a rich source of metaphors for making sense of people and their behaviour.²

These metaphors can be understood using Walton's analysis. By invoking the game that treats life as a journey, we give ourselves a convenient, and perhaps more vivid, way of describing the events in someone's life. We talk about their life by making moves in the game. The metaphor of memory as a notebook can be understood in the same way. By invoking the game that treats memory as a notebook, we give ourselves a convenient—indeed, as we shall see, an indispensable—way of describing someone's behaviour. We talk about their behaviour by making moves in this game. This gives us a way to understand what we are doing when we attribute a particular standing belief to someone. For example, suppose we say that Mark believes that the number 73 bus goes to Oxford Street. When we say this, we are not claiming that he has any inner representation that bears this content; we are indicating how to pretend correctly within the memory-as-a-notebook game. In doing so, we also make an assertion about his behaviour. What makes our pretence appropriate (or inappropriate) and our assertion true (or false) is a complex set of facts about Mark's behaviour. For example, when he wants to go to Oxford Street, he jumps on the bus marked "73"; if someone at the bus stop asks him if the 73 goes to central London he says "yes"; and so on. If, on the other hand, he had cheerily let the 73 go past or answered "no" to their question, then our pretence would have been inappropriate and our assertion false.

For the fictionalist, then, the notion of the mind as an inner world is a useful pretence. Or,

² For a history of the different metaphors that have been applied to memory, see Draaisma, 2000.

to put the point slightly differently, it is a story we tell ourselves. The props of our game are people and their behaviour. Its principles of generation are the rules of folk psychology. Together, these props and principles ask us to imagine that people have a strange inner world that houses their thoughts and guides their behaviour. This is the content of our folk psychological game. Walton (1993) makes a well-known distinction between *content-oriented* and *prop-oriented* make-believe. Children's games are typically content-orientated. The children are not interested in cardboard boxes or piles of stones; these objects are merely useful means of summoning up their imagined world of pirate ships and treasure. By contrast, metaphors usually involve prop-oriented games. When we use a metaphor, our interest is first and foremost in the objects we are talking about, not the pretence we are invoking. Indeed, we often do not even notice when we are speaking metaphorically. The line between content-oriented and prop-oriented make-believe need not be hard and fast (Walton 2000, p. 94). After all, our interests can vary at different times and in different contexts. Much of the time, it would seem, the pretence we invoke to describe the mind is prop-oriented: our interest lies solely in describing people and their behaviour. In more reflective moments, however—perhaps when we are doing philosophy or thinking about psychology—our pretence will be content-oriented: our focus will be on the story we tell about our inner world, rather than merely this or that person's behaviour.

4 Myths and stories

How does fictionalism differ from Ryle's view? The key difference is one that we have encountered already: whereas Ryle takes Cartesianism to be a category mistake imposed upon ordinary language, the fictionalist sees it as an important part of that language. The fictionalist agrees with Ryle that philosophers should aim to chart the "logical geography" of talk about the mind. But she claims that its landscape takes a particular form—namely, a metaphorical one, with its central terms migrating from one area to another. One advantage of this approach is that it allows us to see how different aspects of someone's behaviour "hang together". Consider the enormous variety of different behaviours that

might lead us to conclude that Mark believes that the 73 goes to Oxford Street: he gets on the 73 when he wants to go to Oxford Street; he answers “yes” when someone asks him if the bus goes there; when a friend wants to go to Soho he advises them to get the 73 and walk down Wardour Street; he feels disappointed when he sees a 254 come round the corner instead; he is surprised when the driver announces that the bus will terminate at Tottenham Court Road; and so on. What on earth, we might wonder, do these different sorts of behaviour have in common? Why should our language count all of *these* as situations in which we should attribute this belief to Mark? Attending to the metaphorical structure of our talk about memory offers an answer to this question. Put simply, they are all respects in which Mark behaves *as if* he had “No. 73 goes to Oxford Street” written down in an inner notebook that guides his actions.

Like Ryle, the fictionalist tries to steer a middle course between Cartesianism and behaviourism. According to the fictionalist, Cartesianism’s error is that it takes a metaphor too seriously (see also Melsler, 2004). The ghost in the machine does not exist; it is merely a useful fiction. As a result, fictionalism can avoid the main difficulties that confront Cartesianism (and neo-Cartesianism). There is no mysterious mind-stuff. This means that there is no ontological puzzle about how such stuff could interact with the physical world. The epistemological problem facing Cartesianism is also avoided. For the fictionalist, the facts about someone’s mind are rooted in their public behaviour, not hidden away in some private, inner world. So there is no impenetrable barrier that prevents us from finding out about the minds of others. Finally, fictionalism also sidesteps the semantic worry that Ryle levels against Cartesianism. For the fictionalist, talk about the mind is not a doomed attempt to pick out entities in an inner realm that is screened off from us—either within a mysterious ghostly substance or within the brain; it is simply a description of someone’s behaviour, albeit a metaphorical one.

If Cartesianism takes a metaphor too seriously, behaviourism doesn’t take it seriously enough. Analytic behaviourism claims that talk about the mind can be translated into talk about behaviour. Providing such translations proves impossible, however. Fictionalism explains why it is impossible. One of the main reasons that we value metaphors is that

they expand our powers of expression, allowing us to say things that we could not say without them (e.g. Walton 1993, 2000; Yablo 1998). This means that not all metaphors can be given a literal paraphrase. For example, imagine trying to describe, without recourse to metaphor, exactly what it was about someone's behaviour—precisely what they did or said, what tone of voice they used, and so on—that led you to say that they have a chip on their shoulder. The same lesson applies to many of the metaphors we use to describe the mind. These metaphors add to our powers of expression, allowing us to say things about people that we could not say otherwise. For this reason, we cannot do without such talk by translating it into talk about behaviour (cf. Melsler, 2004, esp. Chapter 10). In fact, for the fictionalist, a sentence like “Mark believes that the number 73 goes to Oxford Street” *means* roughly what the Cartesian takes it to mean: taken literally, it means that Mark has a sentence inside his head with this content. It is just that, according to the fictionalist, we are only pretending that this is the case; we are not asserting it. We pretend that this is the case in order to say something about Mark's behaviour that we couldn't have said otherwise.

If we reject Cartesianism (in both its dualist and materialist forms) and behaviourism, where does this leave us? Recall that Ryle himself looks to answer this question by distinguishing between two different senses of “exist”: the mind may be said to exist, he tells us, but in a different sense to that in which the body can be said to exist. The fictionalist response is rather different. For the fictionalist, the question “do minds exist?” is rather like the question “do angry clouds exist?” In both cases, the answer is: in one sense “yes” and in another sense “no”. Are there clouds that are literally angry (or happy or miserable)? No. Are there clouds that are quite properly called angry (or happy or miserable) when we're indulging in a particular sort of pretence? Yes. The same lesson applies to talk about mental states. Are there really little notebooks inside people's heads that contain their memories? No. Are there real patterns in people's behaviour that are

picked out when we invoke this pretence?³ Yes. In a sense, then, the fictionalist is what Ryle calls a “Reductionist”: strictly speaking, the mind is “nothing but” patterns in our behaviour. And yet, she insists, we can only discern the relevant patterns by pretending they are the outward manifestation of a ghostly presence within.

5 Is it only a story?

Unlike Ryle, the fictionalist regards the notion of the mind as an inner world as part of our ordinary talk about the mind. She also concedes that the story is rather useful, giving us a powerful means of making sense of people and their behaviour. At this point, the Cartesian will ask why we don’t take it more seriously. Why think it’s only a story, rather than an attempt to describe the true nature of the mind? In fact, given how useful it is for explaining and predicting people’s behaviour, why not take it to be a largely *successful* attempt to describe the mind? There are at least two different questions here. The first is descriptive and the second is normative. First, *do* we regard the idea of the mind as an inner world as merely a useful story? Second, *should* we do so?

Let us start with the descriptive question. If the ghost in the machine is part of folk talk about the mind, how do the folk treat such talk? Do they see it merely as a useful story? Or do they believe it? I have suggested that folk talk is metaphorical and that we should analyse these metaphors in terms of pretence. At first glance, this might seem incredible. After all, we do not ordinarily feel as if we speaking metaphorically or pretending when we say what people think or want or feel. The issues here are more subtle than they first appear, however. For often we do not realise that we are speaking metaphorically until we are asked to reflect more closely upon what we have said. More generally, engaging in some domain of discourse is a different matter to offering an analysis of it. Ryle compares the ordinary speaker to a villager who knows his local area. The villager might know

³ For the notion of “real patterns”, see Dennett 1991.

perfectly well how to get from place to place and yet still struggle if asked to produce a map of the area. By contrast, the philosopher's task, like that of the cartographer, is precisely to produce such a map of the territory. It should not be too surprising if the ordinary speaker's workaday knowledge and the philosopher's more abstract knowledge do not immediately tally with one another: although they might both understand a particular area of discourse, their understanding is of a different sort.

How can we know whether our philosophical map is accurate or not? One way to check its accuracy, Ryle tells us, is to see whether it leads us to absurdity. This is Ryle's central objection to Cartesianism, of course. If we mistakenly place mental conduct terms in the categories of non-physical "thing" or "stuff", he claims, then we are led to ask seemingly unanswerable questions, like "how can the mind interact with the body?" Many of Ryle's more specific objections to Cartesianism take a similar form. Poking fun at epistemologists' descriptions of intellectual activity in terms of supposed inner cognitive acts, he writes:

Such expressions are employed, not indeed by the laity but by theorists, as if with their aid, and not easily without it, correct descriptions can be given of what has at a particular moment been occupying a particular person; as if, for example, John Doe could and should sometimes be described as having woken up and started to do some judging, conceiving, subsuming, or abstracting; as spending more than three seconds in entertaining a proposition, or in moving from premises to a conclusion; or as sitting on a fence, alternately whistling and deducing; or as having an intuition of something a moment before he coughed.

Probably most people feel vaguely that there is a tinge of unreality attaching to such recommended biographical anecdotes. John Doe's own stories about himself are not expressed in such terms, or in terms easily translatable into them. How many cognitive acts did he perform before breakfast, and what did it feel like to do them? Were they tiring? Did he enjoy his passage from his premises to his conclusion, and did he make it cautiously or recklessly?

Did the breakfast bell make him stop short halfway between his premisses and conclusion?" (1949, p. 276)

The fictionalist will dispute Ryle's claim that this conception of intellectual activity is entirely foreign to ordinary language. While some of his examples are largely technical terms (e.g. "subsuming"), others are surely used in this manner outside of epistemology (e.g. "reasoning", "intuition", "concept"). The fictionalist can agree, however, that we feel a "tinge of unreality" when describing our inner life in these terms. We would feel a similar tinge of unreality if asked about the causes of the cloud's anger or what breed of sleeping dog we were letting lie. In each case, the reason such questions baffle us is that we are speaking metaphorically. We often describe intellectual activity *as if* it involved assenting to inner sentences, moving through the steps of an inner argument, holding a silent debate between competing views in our heads, and so on. Most of the time we pay little attention to the question of what it is that we are doing when we talk in this way. It is only when someone excessively literal-minded, like a philosopher, demonstrates the absurdity of taking such talk too seriously that we realise that this is not how we normally use these terms.

This leads us directly to our second, normative question. Even if the folk do not intend such talk to be taken literally, *should* we do so? Perhaps the fictionalist is right to say that the folk are not speculating about what goes on inside our heads. But suppose that we find ourselves feeling more curious and do want to find out about our inner workings. Why not take the folk story as providing a promising set of hypotheses about this inner world? Even Ryle acknowledges that myths can be useful (1949, p. 24). Indeed, science often proceeds by taking a more familiar domain (e.g. water waves, billiard balls) as a model for another, less familiar one (e.g. transmission of sound, the behaviour of gases). Perhaps this is how we ought to understand the transfer of talk from the domain of mechanical systems (e.g. pushes, pulls, tensions etc.) and external representations (e.g. meaning, truth, inference) to the mind? Indeed, this is exactly how some have suggested that we make sense of such talk (e.g. Sellars 1956, Bloor 1976, Crane 2015). Moreover, given how productive such models have proved to be within psychology and cognitive science,

don't we have reason for thinking that the folk story is largely true?

Unfortunately, I think the folk story cannot be true since, if it is taken literally, it is simply incoherent. Consider metaphors again. Some metaphors can be understood literally. People can literally, as well as metaphorically, find themselves standing at a crossroads or carrying a heavy burden. Many metaphors do not make sense if taken literally, however (Walton 2000, p. 96). It is hard to know what it would mean to say that clouds are literally, rather than metaphorically, angry. Our ordinary concept of anger simply does not apply to objects like clouds. Similarly, there are parts of our folk story about the mind that simply do not make sense if taken literally, rather than metaphorically. Chief amongst these, I think, is folk talk about inner representations. Our ordinary notion of representation applies first and foremost to external, public forms of representation that represent through social conventions like spoken or written language, maps or diagrams. Talking as if the mind were an inner world containing such representations is an enormously productive for making sense of human behaviour. And yet, when we stop to think about it, we see that the idea that people could really have representations *like these* inside their heads makes little sense (cf. Sprevak 2013).

Notice, however, that this does not mean that talk of inner representations cannot be useful. For one thing, we might manage to come up with a new notion of representation that can be applied to inner states. In a sense, of course, this is precisely what proponents of the representational theory of mind have tried to do: they have tried to show how inner representations might gain their content in some other way—one that is not a matter of social convention—such as causal relations or evolutionary history (e.g. Stich and Warfield 1994). As things stand at the moment, this project has perhaps stalled somewhat, but we should not rule out the possibility that these efforts might eventually succeed. Moreover, even if this project is not ultimately successful and we are stuck with the folk notion of representation, this does not mean that talk about inner representations cannot do useful work—even in scientific contexts (cf. Bennett and Hacker 2003; Dennett 2007). Many scientific models involve assumptions that, if taken literally, seem to make little sense. If we model a bob bouncing on a spring as a simple harmonic oscillator, we

assume that the spring has no mass and that the bob is a point mass. Our model might be very useful and make accurate predictions. And yet, when we stop to think about it, can we really make sense of the idea of a spring without mass or an entity with mass but no extension? (Vaihinger 1924, p. 219; see also Appiah 2017).

6 A figment of the imagination?

It is now time to address the most serious challenge facing mental fictionalism. This is the charge that it is incoherent. One way to put the objection is as follows. Mental fictionalism argues that the inner states posited by folk psychology, like beliefs and desires, do not exist. According to the fictionalist, we ought to regard talk about such states merely as a useful fiction. And yet treating something as a fiction would itself seem to require certain sorts of mental states: we are asked to *imagine* what the fiction says is true or *make-believe* that such and such is the case. As a result, the critic argues, mental fictionalism is incoherent: it denies the existence of mental states while at the same time assuming the existence of at least some of these states, such as imagination or make-believe. In short, how can we even tell a *story* about the ghost in the machine (or anything else for that matter) if we don't have a real ghost (or at least its more substantial, material cousin) inside us?

Before considering how fictionalists might respond to this worry, there are two points that we need to bear in mind. The first is that, as we have seen already, it is misleading to say that fictionalism denies the existence of mental states. If mental states are presumed to be inner representations, then fictionalism does indeed deny that such things exist. Our talk about these inner representations is—and can be—only metaphorical. But the fictionalist does not deny that there are real patterns in our behaviour that are picked out by the use of such metaphors. In this sense, fictionalism does not deny the existence of mental states. Instead, it offers us an account of what mental states are and how they are picked out by folk talk about the mind. The second point to notice is that fictionalism does not deny the existence of representation or content across the board. In particular, the fictionalist grants

that there are external, public representations that are meaningful, such as written and spoken language. Fictionalism itself does not aim to offer a theory of how such representations gain their content. It is not a theory of meaning. Instead, fictionalism's aim is more modest: it seeks to show how the intentionality of the mental is derived from the intentionality of public representations. According to the fictionalist, this happens when the world of public representations is used as a metaphor for the mind as an inner realm.

With these points in mind, let us now turn to consider worries about the coherence of mental fictionalism. The most promising way to respond to these worries, I think, is to insist that the fictionalist's analysis of mental state attributions does not, in fact, make essential reference to mental states like imagination or make-believe. Instead, for the fictionalist's purposes, the core element of Walton's framework is the notion of public, rule-governed acts of *pretence*. At the heart of the fictionalist's analysis is the idea that in "pretending to say one thing, one may actually be saying, asserting, something else" (Walton 2000, p. 95). In the specific case of mental fictionalism, the core idea is that, in pretending to say that someone has a representation inside their head, we are actually saying something about their behaviour. This takes place within the context of folk psychology, which is a particular kind of rule-governed social practice. To avoid the charge of incoherence, the fictionalist must show that we can make sense of this notion of pretence without reference to mental states.

This might seem like a tall order. At first glance, no doubt many would be tempted to agree with Lynne Rudder-Baker when she writes that "it is difficult to see how even to understand pretense without presupposing belief. To pretend that this cup contains water seems to entail believing that it does not" (1994, 199). On closer inspection, however, the link between pretence and (dis-)belief is less straightforward than it might seem. For example, surely the children playing bears in the woods might well believe that bears growl and gnash their teeth, as well as pretend that they do so (Walton, 1990, p. 13; see also Ryle 1949, p. 244). Another natural thought is that, in order to count as pretence, an action must be accompanied by some private, inner act of imagination. As Ryle observes,

however, this claim is also implausible. Suppose that we come across a child who is padding around the woods, growling and gnashing his teeth (Ryle 1949, p. 243). To judge whether he is pretending to be a bear, we do not need to ask whether he is also undergoing some hidden, inner imaginative act of silent growling or gnashing of teeth. His public acts of growling and gnashing are enough.

Instead, it would appear that what is distinctive of acts of pretence is that they involve a particular form of *representation*. Given enough practice, a child's growl might sound indistinguishable from a bear's growl. And yet the two are very different sorts of act. The bear's growl is merely a growl but the child's growl is, at one and the same time, an act of representation (Ryle 1949, pp. 245-246). The child's growl only makes sense in light of the bear's, whereas the reverse is not true. Similarly, a quotation might be indistinguishable from the passage it quotes (*ibid.*). And yet the quotation remains a very different thing to the original: the original sentence might have been uttered in an impassioned speech while the quotation appeared in a dry newspaper report.

Understanding the difference between the two relies not upon close scrutiny of the particular sentence uttered or quoted, but on understanding public speaking and newspaper reporting as distinctive social activities with their own norms. In a sense, of course, this is precisely Walton's approach. The trouble is that his account characterises the relevant social practices and norms in terms of imagination and make-believe: "to pretend, in the sense in question, is to participate verbally in a game of make-believe" (1990, p. 221). The challenge for the mental fictionalist, it would seem, is to offer an alternative characterisation of pretence as a distinctive sort of rule-governed social activity, while retaining the central idea that such activities can be used to make genuine assertions about the world.

Conclusion

Some ghosts are thoroughly malevolent. They do nothing but cause trouble and it would be better for everyone if they could be laid to rest. Others are less straightforward characters. They can be unnerving, certainly. And yet somehow we feel that they are with us because they still have their own mysterious purpose to fulfil. The ghost in the machine is like that. It can undoubtedly lead us astray if we begin to take it too seriously. And yet we cannot stop telling our stories about it, for it provides our best way of understanding ourselves and our place in the world.

Acknowledgements

I would like to thank all the students who have taken my undergraduate course on *The Concept of Mind* over the past couple of years for many enjoyable and stimulating discussions of Ryle's ideas. Thanks also to John Dupré, Roman Frigg, Bill Lycan, Tom Roberts and Frédéric Vallée-Tourangeau for their helpful comments on a draft of this paper, and to Sonia Sedivy for kindly inviting me to contribute to this volume. Finally, and most importantly, I would like to thank Kendall Walton. I can still remember the excitement I felt when I first read *Mimesis as Make-Believe* back in 2004/5, when I was in the first year of writing a PhD on scientific modelling. It immediately transformed my approach to my doctoral project (not to mention my enjoyment of it) and engaging with Ken's ideas has been one of the great pleasures of my intellectual life ever since. For this—and for being so kind and supportive when I met him as a (rather nervous) young graduate student—I'll be forever in his debt.

References

- Appiah, K. A. (2017). *As If: Idealization and Ideals*. Cambridge, MA: Harvard University Press.
- Bennett, M. R., & Hacker, P. M. S. (2003). *Philosophical Foundations of Neuroscience*. Oxford: Blackwell.
- Bloor, D. (1970). 'Is the Official Theory of Mind Absurd?' *The British Journal for the Philosophy of Science*, 21(2), 167–183.
- Crane, T. (2015). 'The mental states of persons and their brains'. *Royal Institute of Philosophy Supplements*, 76, 253–270.
- Demeter, T. (Ed.) (2013a). *Mental Fictionalism* [Special Issue]. *The Monist*, 96(4).

- Demeter, T. (2013b). Mental Fictionalism: The Very Idea. *The Monist*, 96(4), 483–504.
- Dennett, D. C. (1991). ‘Real Patterns’. *The Journal of Philosophy*, 88(1), 27–51.
- Dennett, D. C. (2007). ‘Philosophy as Naïve Anthropology: Comment on Bennett and Hacker’. In M. Bennett, D. Dennett, P. Hacker and J. Searle (Eds.) *Neuroscience and Philosophy: Brain, Mind, and Language*. New York: Columbia University Press.
- Draaisma, D. (2000). *Metaphors of Memory: A History of Ideas about the Mind*. Cambridge: Cambridge University Press.
- Hampshire, S. (1950). ‘Review of *The Concept of Mind*’. *Mind*, 59(234), 237–255.
- Lycan, B. (forthcoming). ‘A Rylean Mental Fictionalism’, to appear in T. Demeter, T. Parent & A. Toon (Eds.) *Mental Fictionalism*. London: Routledge.
- Melser, D. (2004). *The Act of Thinking*. Cambridge, MA: MIT Press.
- Rudder-Baker, L. (1994). ‘Attitudes as nonentities’. *Philosophical Studies*, 76(2–3), 175–203.
- Ryle, G. (1949). *The Concept of Mind*. London: Hutchinson.
- Ryle, G. (1979). *On Thinking*. New Jersey: Rowman and Littlefield. (Edited by K. Kolenda)
- Sellars, W. (1956). ‘Empiricism and the Philosophy of Mind’. In H. Feigl & M. Scriven (Eds.) *The Foundations of Science and the Concepts of Psychology and Psychoanalysis. Minnesota Studies in the Philosophy of Science, Vol. 1*. Minneapolis: University of Minnesota Press.
- Sprevak, M. (2013). Fictionalism About Neural Representations. *The Monist*, 96(4), 539–560.
- Stich, S. & Warfield, T. (eds.) (1994). *Mental Representation: A Reader*. Oxford: Blackwell.
- Tanney, J. (2009). ‘Rethinking Ryle: A Critical Discussion of the Concept of Mind’. In

- Ryle, G. (1949/2009) *The Concept of Mind: 60th Anniversary Edition*. London: Routledge. pp. ix-lix.
- Toon A (2010). 'The ontology of theoretical modelling: Models as make-believe'. *Synthese*, 172(2), 301-315.
- Toon A (2011). 'Playing with molecules'. *Studies in History and Philosophy of Science*, 42(4), 580-589.
- Toon, A. (2012). *Models as Make-Believe: Imagination, Fiction and Scientific Representation*. Basingstoke: Palgrave Macmillan.
- Toon, A. (2016). 'Fictionalism and the folk'. *The Monist*, 99: 280-295.
- Vaihinger, H. (1924). *The Philosophy of 'As If': A System of the Theoretical, Practical, and Religious Fictions of Mankind*. Abingdon, Oxford: Routledge.
- Wallace, M. (2007). 'Mental Fictionalism'. Unpublished manuscript.
- Wallace, M. (2016). 'Saving Mental Fictionalism from Cognitive Collapse'. *Res Philosophica*, 93(2), 405-424
- Walton, K. (1990). *Mimesis as Make-Believe: On the Foundations of the Representational Arts*. Harvard: Harvard University Press.
- Walton, K. (1993). 'Metaphor and Prop Oriented Make-Believe'. *European Journal of Philosophy*, 1(1), 39–57.
- Walton, K. L. (2000). 'Existence as metaphor?' In Everett, A. & Hofweber, T. (Eds.) *Empty Names, Fiction, and the Puzzles of Non-Existence*. CSLI Publications (pp. 69-94). Reprinted in Walton, K., *In Other Shoes: Music, Metaphor, Empathy, Existence*. New York: Oxford University Press (pp. 89-113). (Page numbers cited in text refer to reprinted version.)
- Yablo, S. (1998). Does Ontology Rest on a Mistake? *Proceedings of the Aristotelian Society, Supplementary Volume*, 72, 229–261.