Directive Content

Patrick Butlin

Forthcoming in *Pacific Philosophical Quarterly*; please cite published version

Abstract: Representations may have descriptive content, directive content, or both, but little explicit attention has been given to the problem of distinguishing representations of these three kinds. We do not know, for instance, what determines whether a given representation is a directive instructing its consumer to perform some action, or has descriptive content to the effect that the action in question has a certain value. This paper considers what it takes for a representation to have directive content. The first part of the paper presents the Liberal View, which might be taken to be the default position on this issue. The Liberal View has some attractions, but as the second part shows, these are less conclusive than they might at first appear, and there is much to be said for an alternative, the Strict View.

1. Introduction

The most basic representations, in the simplest, most fundamental representational systems, have either descriptive content, directive content, or both. Representations with descriptive content purport to say how things are, and can be true or false, accurate or inaccurate. Representations with directive content tell the systems that consume them what to do, and can be satisfied or unsatisfied.[1] According to Ruth Millikan, whose theory of representation has been exceptionally influential in shaping the modern debate, what it takes for something to be a representation is for it to have either the kind of function that suffices for and characterises descriptive content, or that which is similarly associated with directive content, or both (1984, 2004); and in the simplest cases representations have both descriptive and directive content – they are 'pushmi-pullyu' representations (1995).

An obvious question to ask about descriptive and directive content is which representations have which kinds of content, and why. For example, consider the much-discussed case of the honeybee's waggle dance. When foraging bees return to the hive

[1] Descriptive and directive content are also known as 'indicative' and 'imperative' content, respectively.

after finding sources of nectar, they sometimes perform dances with features which correspond to the location of the source relative to the hive. The orientation of the dance corresponds to the direction to the source of nectar from the hive, relative to the position of the sun, and the number of waggles corresponds to the distance. Sometimes watching bees respond to the dance by flying to the location it indicates. Does the dance have descriptive content, saying something like: *There is a nectar source at x distance in y direction*? Does it have directive content, saying something like: *Forage at the place at x distance in y direction*? Or does it have both kinds of content, saying something that amounts to the amalgamation or conjunction of these two messages? If we take there to be a real distinction between descriptive and directive representations we should know how to answer these questions, and justify our answers.

In this paper I will argue that we do not currently have a well-established theory of what distinguishes representations with directive content. In the first part of the paper (sections 2 and 3) I will present an account of directive content which might be taken to be the default position. I will describe some of its advantages, and mention support for it in the literature. Then in the second part (sections 4 and 5) I will present an alternative account. I will argue that this alternative account has some advantages over the first one, and that some of the apparent advantages of the first one are less conclusive than they may initially seem. I will also discuss similarities and differences between the second account and existing views of directive content. My aim in the paper overall is to exhibit the attractions of the alternative account, and hence to show that the nature and prevalence of directive content is a topic worthy of more explicit discussion than it has so far received.

Before I turn to the first account, I need to make three preliminary, clarificatory comments.

First, to keep things relatively simple I focus on what might be called 'biological' representations, such as sub-personal mental representations, bodily sensations, and animal signals. If one of the accounts I consider here is the correct way to characterise biological representations with directive content, I would expect analogous accounts to be correct of directive content in representations of other kinds. But it is doubtful whether my two accounts could be applied to the case of sentences in human languages (for instance) without significant modifications, so I do not take my comments to have any immediate consequences beyond the realm of biological representation.

Second, for the sake of argument, I will assume that both descriptive and directive representations have states of affairs as their contents. Thus if the waggle dance has descriptive content, its content *qua* descriptive is the state of affairs: *there being a nectar source at x distance in y direction from the hive*. As a descriptive representation, it says that this is the case. If the waggle dance has directive content, its content *qua* directive is the state of affairs: *the watching bee's foraging at x distance in y direction from the hive, in the near future*. As a directive representation, it tells the watching bees to make this the case. To have descriptive content, I assume, a representation must stand in a complex, perhaps normative or historical relation of the right kind to some state of affairs, which will then be its content *qua* descriptive. If a representation does not stand in such a relation to any state of affairs, it lacks descriptive content. The same holds in the case of directive content, except that the complex relation is different. I therefore take it that it could be a consequence of some theory of descriptive and directive content that a representation could not have both kinds of content, but that in the abstract this is not ruled out: a representation could simultaneously stand in the descriptive-relation to one state of affairs, and the directive-relation to another. For ease of expression, I will also assume that representations are events that may recur; this allows me to talk of representations as occurring, and potentially on multiple occasions. I do not expect these metaphysical assumptions to influence the substance of my discussion.

Third, I must emphasise that my focus is on the issue of the prevalence of directive content, as opposed to the relative merits of the conceptual and theoretical frameworks which are used to theorise about content more broadly, including the descriptive/directive distinction. I will discuss directive content in the context of two such frameworks: one is teleosemantics (Millikan 1984, 2004; Papineau 1993, 1998; Price 2001; Shea 2007, 2018), and the other combines tools from game theory and the mathematical theory of information in analysing representation (Lewis 1969; Skyrms 2010; Huttegger 2007; Zollman 2011; Martínez & Klein 2016; Shea et al. 2016). However, I will not argue for one of these two frameworks over the other, and at least one of the two views of directive content that I will consider can be expressed equally well in either framework. I take the issue that I am addressing to be orthogonal to any controversy between these two alternatives. Both frameworks take representations to be used for co-ordinating the behaviour of producer and consumer systems (Godfrey-Smith 2013), and this idea will be central to my discussion.

2. The Liberal View

The first account of directive content I will consider has been explicitly advocated in various forms by David Papineau and Nicholas Shea, and has been widely attributed to Millikan. So it is arguably the canonical account of directive content in the teleosemantic tradition. I will call it *the Liberal View*, because compared to the alternative account which I will introduce in section 4, it entails that many more representations have directive content.

Here is a statement of the Liberal View:

> *Liberal View*: A representation has directive content if and only if it has the function of causing its consumer to behave in a particular way.

This way of expressing the Liberal View adopts some elements of Millikan's teleosemantic theory of representation. In Millikan's theory, all representations have consumers, which are entities with biological functions that involve behaving in different ways, depending on whether the representation occurs. The functions of a representation are, at a first pass, whatever effects representations of that kind have caused in the past, which have systematically contributed to the survival and reproduction of the organism's ancestors. So the Liberal View says that a representation has directive content if and only if there is some particular way that it has caused its consumer to behave – such as causing a particular outcome to be brought about – which has systematically contributed to survival and reproduction.

An immediate worry about the Liberal View might be that it appears to entail that every representation has directive content, since every representation has a consumer, and the consumer's function always involves doing something different when the representation occurs (this objection is elaborated by Artiga 2014). To avoid this conclusion, the advocate of the Liberal View should point out that many representations will cause their consumers to behave in a range of different ways, in accordance with the consumer's function, since the consumer may also be sensitive to other representations which it also consumes, or other features of the situation more generally. Such representations are like beliefs: a given belief can contribute to causing many different actions, depending on the agent's desires. The advocate of the Liberal View will also need a principled reason to deny that such representations have disjunctive contents, but

such a reason is likely to be available, especially since theorists in the teleosemantic tradition often appeal to explanatory considerations in identifying functions.

For the Liberal View's condition on directive content to be satisfied, it is not necessary that the behaviours by the consumer that fulfill the representation's function are all identical. Instead, they must merely have some property in common, which would be mentioned in a correct account of the representation's function, besides being appropriately performed when a particular state of affairs obtains (e.g. in the environment; the reason for this exception is that causing behaviours with a property of this kind in common is what characterises descriptive content, according to canonical teleosemantics). For instance, a directive representation could have the function of causing the consumer to get water, even if it took different means to this end on different occasions.

To see the Liberal View in action, consider again the case of the honeybee waggle dance. It is not the case that, as we might suppose, watching bees almost always forage immediately in the area indicated by the dance. Instead, the subsequent actions of watching bees are affected by the scents carried by dancing bees and their own foraging experience, as well as by the form of the dance (Grüter & Ratnieks 2011). So it appears that watching bees assess the likely quality of the foraging site discovered by the dancing bee, relative to that of the sites with which they are already familiar. Nonetheless, according to the Liberal View, waggle dances have directive content. This is because the function of each possible dance is to cause observer bees to fly to the corresponding location. It is by causing such flights that waggle dances have contributed to bees' survival and reproduction. The other behaviours that observer bees perform after watching dances (such as foraging in other locations) may contribute to the success of the hive, but it is not part of the function of the dance to cause these behaviours, because they could equally well have been performed even if there had been no dance.

In her 'Pushmi-Pullyu Representations' (1995), which focuses on descriptive and directive representations, Millikan adopts the Liberal View. She writes (p. 189):

> A representation is directive when it has the proper function to guide the mechanisms that use it so that they produce its satisfaction condition.

As I will explain in section 4, her position in *Language, Thought and Other Biological Categories* (1984) is less clear. But philosophers responding to Millikan's work have also tended to

attribute the Liberal View to her. For example, Artiga (2014, p. 552) writes that Millikan and other teleosemantic theorists are committed to the claim that:

> Necessarily, if a representation R is supposed to make the consumer system C produce an effect E, R imperatively represents E.

A similar attribution can be found in Price (2001, p. 137).

The Liberal View is also prominent in other versions of teleosemantics. Papineau's teleosemantics aims to provide a theory of content for beliefs and desires, understood as states defined by their functional roles. He argues that desires have satisfaction conditions (i.e. directive content) because they have the biological function of bringing about specific effects (1993, 1998, 2003). For example, he writes that 'we can analyse the satisfaction condition for any desire as that specific effect it is biologically designed to bring about' (2003, p. 111). Papineau's position counts as a version of the Liberal View even though he does not refer to the consumer, because this and the other differences between his theory and canonical teleosemantics do not directly concern the distinction between descriptive and directive content. This point also applies to Shea's (2018) theory of representation: the theory is broadly teleosemantic, although with important differences from the canonical formulation, but Shea adopts the Liberal View in his account of the descriptive/directive distinction. Shea's view is, roughly, that a representation has directive content if and only if its causing a condition C to obtain is part of what explains the organism's ability to perform a task (2018, ch. 7).

3. Advantages of the Liberal View

We have seen that the Liberal View is both easy to formulate in the teleosemantic framework, and popular among leading teleosemantic theorists. It also gives a fairly intuitive verdict about the waggle dance case, and this case illustrates that it leaves open the possibility of representations with directive content which the consumer may appropriately 'choose to ignore'. To put this less metaphorically: it leaves open the possibility that the consumer of a directive representation, acting in accordance with its function, may do something other than bring about the state of affairs that is the representation's content. This is also an apparent advantage of the Liberal View, because examples from language suggest that it is not an essential feature of directives that they should always be acted upon. In cases of presupposition failure, for instance, or cases in

which a person is given two incompatible instructions, it may be impossible for the hearer to do as they are told. So it is intuitive that there may sometimes be subpersonal or biological directives that the consumer is not always obliged to follow.

However, I suspect that for many philosophers the most significant advantage of the Liberal View will be that it coheres well with established ideas about *direction of fit*, understood as a property of beliefs and desires. So in the remainder of this section I will describe this coherence in some detail. In section 5 I will argue that this advantage is actually rather limited, and does not speak conclusively in favour of the Liberal View over the stricter alternative that I will shortly introduce, but for now my aim is just to show why the Liberal View is attractive.

The origin of the notion of direction of fit is often traced to Anscombe (1957), but the main development of the theory takes place in works by Platts (1979), Searle (1983), Smith (1987) and Humberstone (1992). The key idea is that beliefs and desires have opposite directions of fit, most often called the *mind-to-world* and *world-to-mind* directions of fit, respectively. Very roughly, according to the theory, beliefs have the mind-to-world direction of fit because they 'succeed' when they are true – that is, when the belief, which is in the mind, fits how things are in the world. Desires have the world-to-mind direction of fit because they 'succeed' when they are satisfied – that is, when the world fits (or comes to fit) the desire. The idea that there are opposite directions of fit is further supported by the point that when there is a mismatch between one's beliefs and the world the appropriate thing to do is usually to change one's mind, whereas when there is a mismatch between one's desires and the world it is typically appropriate to act, changing the world rather than the mind.

The phenomenon of direction of fit is evidently a close cousin of the descriptive/directive distinction, and indeed they are sometimes assimilated. On one hand, Smith (2011, p. 154) writes that to have a desire is to have within oneself a 'representation of how the world is to be' – which sounds like a representation with directive content. On the other hand, writers in the teleosemantic tradition sometimes use the terminology of direction of fit when discussing descriptives and directives. However, I will assume that the idea of direction of fit applies only to personal-level mental states. This means that the relationship between direction of fit and the descriptive/directive distinction as it applies to the representations in my remit is uncertain, because the relationship between personal-level and subpersonal-level mental states is a matter of debate. Despite this, there is a noteworthy coherence between the

Liberal View (which applies to subpersonal states) and philosophical common sense about direction of fit.

The first aspect of this coherence is that the Liberal View entails that subpersonal representations which play functional roles like those associated with desire have directive content. This point may be particularly important in motivating philosophers to adopt the Liberal View, because the idea that desires have the world-to-mind direction of fit is found so compelling.

The representations I have in mind are implicated in the theory of goal-directed control, which has been developed in behavioural psychology and cognitive neuroscience (Dickinson & Balleine 1994, Morris et al. 2018). Goal-directed control is a form of action selection used by humans and many other animals, which is distinguished from other forms by its sensitivity to information both about causal relationships between actions and outcomes, and about the values of outcomes. From behavioural evidence of this dual sensitivity, psychologists infer the use of representations of action-outcome contingencies and of the values of outcomes in selecting actions. For example, Adams and Dickinson (1981) showed that rats that had been trained to press a lever to receive a specific foodstuff would cease responding after that food had been devalued for them, even when the devaluation procedure took place away from the lever, and the test was carried out without delivering the devalued food. This is taken as evidence that the rats represented that pressing the lever would lead to the delivery of the specific food, and drew on both this representation, and a further representation of the value of the food, in deciding whether to press it. Humans are also capable of goal-directed control, and there is evidence that, in humans, some episodes of activity in the orbitofrontal cortex (OFC) represent the values of outcomes for the purposes of goal-directed control (Balleine & O'Doherty 2010, Rushworth et al. 2011). For the sake of a neutral term, I will call these episodes *OFC representations*.

This example supports the Liberal View, because OFC representations have a functional role much like that of desire (Heyes & Dickinson 1990), and the Liberal View entails that they have directive content. OFC representations and desires are similar in that they pick out possible outcomes as worthy of pursuit, and cause action in combination with what appear to be instrumental beliefs.

The reason why the Liberal View entails that OFC representations have directive content is worth going through more slowly. Suppose that at a given time an animal has OFC representations concerning the values of strawberries and orange pieces, and also

represents relationships between actions which are available to it and the outcomes of receiving each of these two foods (experiments in action selection often investigate situations of this form). The theory of goal-directed control takes action selection to be determined by the expected values of actions, in the sense familiar from decision theory (Rangel & Hare 2010). So what the animal will do in this situation depends on how valuable they take each of the possible outcomes to be, and on how likely they take it to be that each of the available actions will lead to each outcome. We can also assume that a certain psychological subsystem is the consumer of the various representations, calculates the expected values of actions, and causes the one with the highest expected value. Having an OFC representation concerning the value of strawberries therefore may or may not cause the animal to get strawberries, even if this subsystem works according to its biological function.

However, this OFC representation does have the function of causing the animal to get strawberries, because it is by having this effect that the representation can systematically contribute to the animal's survival and reproduction. In particular, if the animal gets some other beneficial outcome from the situation, like getting orange pieces, this will not have been caused by the OFC representation in question, because it could equally well have been achieved in the absence of the strawberry-representation.[2] So the Liberal View tells us that this representation has directive content: *get strawberries!*

The Liberal View also coheres well with how philosophers think about direction of fit in a second way. This is that it is analogous to one natural formulation of what is involved in the world-to-mind direction of fit. As Platts puts it, desires have the world-to-mind direction of fit because they 'aim at realisation' (1979, p. 257). On this way of putting it, it seems to be desires themselves that succeed when their objects are realised. What gives desires the world-to-mind direction of fit is what they themselves are for. The Liberal View appears to provide a less metaphorical version of the same idea, by claiming that representations with directive content have biological functions, which are fulfilled when they cause specific states of affairs to be brought about.

[2] There is a complication here. If the strawberry-representation was replaced by one that attributed a higher value to getting strawberries, the animal would have sought strawberries instead of orange pieces. So there is a case to be made that the particular strawberry-representation in this case did contribute to causing the animal's getting orange pieces. There is a benefit in having OFC representations which attribute the right values to particular outcomes that goes beyond that of bringing about those outcomes. However, the only specific beneficial outcome to which the strawberry-representation *systematically* causally contributes is getting strawberries.

Given that the Liberal View has these advantages, perhaps the reason why there is relatively little explicit discussion of directive content is that there is no problem to be solved: the obvious position – the Liberal View – gets things right, and is deservedly popular.

4. The Strict View

There are a range of possible alternatives to the Liberal View, but I will focus on just one, which I will call *the Strict View*. In this section I first introduce the Strict View and show how it differs from the Liberal View, then present some of its attractions (in 4.1), then finally examine accounts of directive content which adopt either the Strict View or closely related positions (in 4.2).

The Strict View is superficially similar to the Liberal View:

> *Strict View*: A representation has directive content if and only if its consumer has the function of behaving in a particular way, whenever the representation occurs.

The difference is that the Liberal View requires that the representation itself has a function of the right kind in order to have directive content, whereas the Strict View requires that the consumer has a function of the right kind with respect to the representation.[3]

To see how this difference plays out, we can return to the example of the waggle dance. Recall that when a bee watches a waggle dance, it does not always fly to the location indicated by the dance. Instead, there are other apparently adaptive subsequent behaviours which watching bees perform, such as foraging in other locations which are known to be good sources of nectar (especially when the scents carried by the dancing bee suggest that the plants in these locations may be flowering). This indicates that the watching bee's brain does not have the function of causing it to fly to the location specified by the dance, whenever a dance occurs. Its function involves causing different behaviours in response to dances, depending on other features of the situation, such as

[3] One minor complication for the Strict View is that sometimes representations are produced but not 'noticed' by their consumers. In Millikan's terms the consumer is not 'adapted' by the representation (1984, ch. 2). In these situations I take it that the consumers' function is to behave as though the representation had not occurred, so really the Strict View should claim that consumers of directives have the function of behaving in a particular way whenever they are adapted by the representation. I set aside this complication so as to minimise my use of potentially unfamiliar terminology.

the quality of other known foraging sites. Since there is no one way that the watching bee should behave, the Strict View entails that the waggle dance lacks directive content.

The Liberal View takes the waggle dance to be a directive – something like an instruction – to forage in a certain location, that the watching bees sometimes rightly choose to disregard. The Strict View says that the waggle dance has only descriptive content. It informs watching bees about a source of nectar, and 'leaves up to them' what to do with this information.

The reason why the Liberal View is more liberal than the Strict View, then, is that it is possible for a representation to have the function of causing its consumer to behave in a particular way, even though the consumer does not have the function of behaving in this way whenever the representation occurs. One scenario in which this is the case is when the consumer is also responsive to another input or combination of inputs which take precedence over the representation in question, and which cause the same output whether or not this representation occurs. As far as I can tell, it is not possible for a representation to have directive content on the Strict View, but not on the Liberal View – if the consumer has the function of behaving in a particular way when a representation occurs, that representation must have the function of causing this or a related behaviour.[4]

More abstractly, the difference between the Liberal View and the Strict View arises because, at least in the teleosemantic framework, there are two natural ways in which a representation may stand in a privileged relation to a state of affairs which it causes its consumer to bring about. A fundamental characteristic of directive representations is that their content is such a state of affairs, so to give a theory we need to identify a relation between representations and such states of affairs which is important in some way, and which holds in some cases of representation but not others. As we have seen, the teleosemantic framework gives us two ways to do this: either in terms of the representation's function, yielding the Liberal View, or in terms of the consumer's function, yielding the Strict View. Another way to look at the difference is in terms of the situations in which the consumer's behaviour must be invariant, for the representation in question to have directive content. According to the Strict View, the consumer's behaviour must be invariant (at least in one important respect) whenever the representation occurs. According to the Liberal View, the consumer's behaviour must be

[4] A reader of an earlier draft suggested to me that there might be counterexamples in cases of partial common interest – i.e. cases where the interests of the producer and consumer are not wholly aligned, despite being sufficiently closely related to make communication possible. However, I have been unable to come up with a fleshed-out case of the required form.

invariant across all occasions on which the occurrence of the representation contributes in its typical way to survival and reproduction. It may be possible to specify further theories in this way, by describing different conditions under which invariance is required, but it is difficult to imagine other, comparably natural candidates.[5]

One might be immediately inclined to dismiss the Strict View on the grounds that it is so strict as to entail that almost no biological representations have directive content. However, although the Strict View is restrictive, it does not seem to me to be absurdly so. One point to note here is that as in the Liberal View, 'behaving in a particular way' should be understood relatively broadly, in a way which makes it consistent with taking different means to an end. What it means for the consumer to 'have the function of behaving in a particular way' when a representation occurs is for all of the behaviours it might perform in this situation, which are consistent with its function, to have a property in common that would be mentioned in a correct account of the consumer's function. It is not as though possible responses to directive representations must be identical in every respect. This point is important because some theories in the teleosemantic tradition require that all consumers must have some 'flexibility of response' to representations (e.g. Cao 2012). This feature ensures that the Strict View is not incompatible with such theories.

I will also briefly present two cases in which biological representations do have directive content, even according to the Strict View. First, Schuchmann (1989) reports that hummingbird chicks open their mouths to allow the parent to feed them in response to two different stimuli at different stages of development. In the first few days of life, they gape when the parent taps them just behind their eye-bulges with its beak. Then at a somewhat later stage, they do so in response to air movements caused by the hovering parent's beating wings. In the earlier stage, the parent's taps seem to function as representations, and Schuchmann reports that in his studies they 'always' caused the chicks to gape. So this appears to be a directive representation by the standards of the Strict View.

[5] Thanks to an anonymous referee for suggesting this way of putting things. A different way to form alternative theories of directive content is by adding conditions to either the Strict View or the Liberal View. For example, Price (2001, ch. 6) adds to the Liberal View the condition that for a representation to have directive content, its consumer must take a variety of means on different occasions to bring about the state of affairs that the representation specifies. This doesn't change the range of situations in which the consumer's behaviour must be invariant; instead it gives a more detailed specification of the nature of the invariance required.

Second, *motor intentions*, as described by Pacherie (2006), appear to have directive content by the standards of the Strict View. Motor intentions are subpersonal mental states that initiate, sustain and guide bodily movements, directly controlling the operation of motor control systems. They do not specify every detail of sequences of muscle contractions, but instead represent intended outcomes at various levels of generality (Brozzo 2017). For example, a motor intention might represent that an apple is to be grasped, or grasped using a whole-hand grip with the right hand, with the motor control system filling in the details in a dynamic process as the action develops. For our purposes the crucial feature of motor intentions, in contrast to more distal forms of intention, is that the function of the motor control systems that consume them is always to bring about an action of the kind that they specify. This is because motor intentions are formed downstream of both processes of practical reasoning, and perceptual processes in which the affordances of the environment and the condition of the body are assessed. Motor intentions can be prevented from bringing about actions only by failures of functioning of the motor control systems or the body itself (Jeannerod 2006, pp. 3-4), or perhaps by being replaced and superseded by later contrary motor intentions. So motor intentions have directive content because they suffice to determine their consumer's functions.

Finally, one might also be inclined to dismiss the Strict View for the reason that it entails that OFC representations lack directive content, despite their functional similarity to desires. This is because the consumer of OFC representations has the function of behaving in different ways when a particular OFC representation occurs, depending on the content of the other representations which it also consumes at the same time. In personal-level terms: what one should do when one has a particular desire depends on what else one desires at the time, and on one's instrumental beliefs. In section 5 I will argue that this is a much less serious weakness of the view than it might appear.

4.1 Advantages of the Strict View

The Strict View has two main advantages, which I will describe in this section. Both are related to the game-theoretic approach to representation developed by David Lewis (1969) and Brian Skyrms (2010).

Lewis and Skyrms ask us to consider variations on the following basic set-up. There are two agents, a sender and a receiver. The sender can observe states, and the receiver can perform actions, and pay-offs to both sender and receiver depend on which actions

the receiver performs in which states. If the sender can produce signals[6] which are more readily discriminable by the receiver than the states themselves, and if the sender and receiver are capable of changing their behaviour so as to receive greater pay-offs, their behaviour may approach a mutually beneficial equilibrium. In this equilibrium the sender will send signals in response to states, and the receiver will condition its actions on these signals. Lewis and Skyrms call this a *signalling system*, and suggest that the evolution of signalling systems is a great part of the evolution of meaning.

I will explain shortly how this idea is used in analysing real-world representations, but the first advantage of the Strict View can best be seen by reflecting on the basic set-up just described.

It is clear that in this situation, one way in which signals can work is by being sufficiently reliably correlated with states that make a difference to the pay-offs returned by the receiver's actions. It is natural to think that signals that work in this way will therefore have descriptive content: they will say that the states in question obtain. For example, suppose that you work in a windowless room, and my office (which is next-door) has a window, and we have some mutual interest in your performing a certain action when it rains. We might develop a system in which I blow a whistle whenever I notice that it's raining – since I've got a window, I'm likely to be a more reliable observer of this than you. You perform the action when you hear the whistle. Leaving aside for now the question of whether the whistle-signal has directive content, it does seem to have descriptive content. It tells you that *it's raining*.

I suggest that representations with directive content must work in *some other way*, different to the way in which descriptives characteristically work. However, it should also be possible for representations to work in both ways at the same time, since it's very plausible that representations in the simplest signalling systems (including our system of the whistle) have both descriptive and directive content. So we face the question: what other way of working is there, besides being sufficiently reliably correlated with states that make a difference?

One possibility seems to be: by being such that, according to the established practice of the signalling system, the receiver should always respond by performing a particular action. Representations can work this way when the sender is capable of identifying, sufficiently reliably, circumstances in which it will be beneficial for both parties for the

[6] Lewis and Skyrms talk of 'signals' rather than 'representations'. I use the two terms interchangeably in this section.

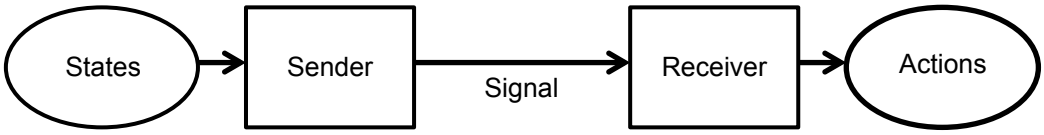
receiver to perform this action. Such representations are naturally described as having directive content, because their occurrence suffices to tell the receiver what to do. They work in virtue of their connection to a particular action, and describing them as having directive content means that by giving their content we will specify this action.

Furthermore, some such representations also work as descriptives do, while others do not. In the case with the whistle, according to the established practice, the signal is produced in response to a state with makes a difference to pay-offs from the receiver's actions, and is *also* such that there is a particular action which should always be performed when it is heard. This arrangement makes sense because things are easy – there is just one state that matters for one action. However, it is also possible to imagine signalling systems in which the sender may send the signal in any one of a range of possible states, which have nothing in common except entailing that a particular action by the receiver will yield a high pay-off. The possibility of this kind of case shows that the two ways of working are genuinely distinct, because here the signal does not work by being sufficiently reliably correlated with a state of any description, except insofar as this is trivially implied by its relation to the action it calls for. Such signals, I suggest, have directive but not descriptive content.[7]

The attractiveness of this proposal favours the Strict View, because both the proposal and the Strict View claim that for a representation to have directive content, the system that responds to the representation should be required to behave in a particular way whenever it occurs. The consumer/receiver has this requirement either according to its function, or according to the established practice of the signalling system, and in biological cases these seem to come to the same thing. So the Strict View appears to capture a fundamental distinction between two ways in which representations can work.

The second advantage of the Strict View is that it yields an appealing symmetry between the conditions for descriptive and directive content. This is most easily seen with the help of a diagram:

[7] This line of thought is inspired by Lewis's account of the distinction between descriptives and directives (he calls them 'indicative' and 'imperative'), which I discuss further in 4.2. My suggestion also allows for representations with descriptive but not directive content, because it is often the case that what the receiver should do when a certain state obtains also depends on some further feature of the situation, but it is nonetheless worthwhile to use a signal indicate this state.



The Strict View claims that, for a representation to have directive content, its consumer must have the function of behaving in a particular way, whenever the representation occurs. In other words – looking now at the diagram – whenever a directive signal is employed, the receiver is required to make it the case that a particular corresponding action is performed. So if all goes well, we get an 'outward' entailment on the diagram from the signal to the action; if the signal occurs, the action occurs. For descriptive content, it is plausible that the sender must have the function of producing the signal only when the corresponding state obtains (i.e. only when it will be true).[8] This condition means that when all goes well, we get a symmetrical 'outward' entailment from the signal to the state; if the signal occurs, the state obtains.

This symmetry can also be described in other ways. For instance: if a representation may correctly be produced under two or more different conditions, which are not united by the right kind of higher-level description, it cannot have descriptive content; and if a representation may be correctly responded to by performing two or more different actions, which are not united by the right kind of higher-level description, it cannot have directive content. This follows from the Strict View alone.

4.2 The Strict View in the Literature

The attractions I have just presented give us one kind of reason to take the Strict View seriously; while the view may be austere, its underlying features take the form that

[8] What about the point that it is often adaptive for producers of representations with descriptive content to be tolerant of false positives (e.g. when they indicate danger; Godfrey-Smith 1992) – doesn't this show that descriptives often *should* be produced when their content doesn't obtain? My thought is that in these cases the producer only fulfills its function when it produces true representations, but that the particular way in which it does so is also adaptive in virtue of producing false positives much more readily than false negatives. Similarly, the consumer of a directive representation may not always succeed in behaving the way that it is supposed to (e.g. it may not manage to bring about a particular result). In this situation, too, there are different ways to fail, and the consumer may be adaptive in that the way in which it behaves is much more likely to yield benign failures that life-threatening ones.

we should hope for in a theory of directive content. Another reason to take it seriously is that the Strict View and closely-related theories have been advocated in the literature, especially under the game-theoretic approach to representation.

Perhaps surprisingly, despite the way in which her view is usually interpreted, Millikan's original statement of teleosemantics appears to adopt the Strict View. In *Language, Thought and Other Biological Categories* (1984, p. 99) she writes that, 'In the case of imperative intentional icons, it is the proper function of the interpreter device, as adapted to the icon, to produce something onto which the icon will map in accordance with a specific mapping function…'. In our terms, 'imperative intentional icons' are representations with directive content, and 'interpreter devices' are consumers. For a consumer to be 'adapted to' a representation is for the representation to be so situated with respect to the consumer that the consumer's function is to produce something that relates to the representation in a particular way. Millikan's claim appears to be that the occurrence of directive representations entails that their consumers have the function of behaving in particular ways – of bringing about outcomes that map onto those representations.

I will not attempt to establish here which view Millikan actually prefers, or whether her position has changed over time. But the ambiguity in Millikan's position is worth noting, because the fact that she is usually interpreted in a way that apparently contradicts her most famous work illustrates the lack of focused debate on directive content. The Liberal View and the Strict View sound similar, but are in fact very different.

In the game-theoretic tradition, the descriptive/directive distinction is discussed by Lewis (1969), Huttegger (2007), Zollman (2011) and Martínez and Klein (2016). In *Convention*, Lewis proposed a theory of descriptive and directive content in signalling conventions used by intelligent agents. His proposal thus concerns representations of a different kind from those which are the focus here, but remains instructive. Lewis claims that three kinds of signalling conventions are possible (1969, pp. 144-146), which he says involve *neutral signals*, *indicative signals*, and *imperative signals*:

> *Neutral signals* are used when the signalling convention specifies both the conditions under which the signal should be produced, and the action which is to be taken when it is recognised;

Indicative signals are used when the convention specifies only the condition of production, and leaves the action open to the discretion of the receiver;

Imperative signals are used when the convention specifies only the action, leaving it to the discretion of the sender to decide when to produce it.

If we translate 'indicative' and 'imperative' as 'descriptive' and 'directive', and assume that neutral signals have both descriptive and directive content, Lewis's proposal means that a representation has directive content if and only if the convention governing it specifies the action to be taken when it is received, as opposed to leaving this action to the discretion of the agent.

This proposal evidently shares the spirit of the Strict View. Like the Strict View, it is motivated by the idea that descriptive and directive content should involve symmetrical constraints on the users of representations, and works by requiring that, for directive content, a representation must set a correctness condition of some kind on the behaviour of the system or agent that consumes it. Also, I shall shortly argue that Huttegger's account, which is an attempt to adapt Lewis's theory to cases in which representations are used by relatively simple systems, is a version of the Strict View.

However, the relationship between Lewis's proposal and the Strict View is not straightforward. The problem is that Lewis does not discuss how we should treat cases in which a representation does not suffice to determine what its consumer should do, according to the convention that governs it, but in which this is not left to the consumer's discretion either. For example, suppose that a luge racer must begin her descent when and only when she receives three 'all-clear' signals reporting on different parts of the course. Under this convention, none of the three signals individually entails that she should begin, but she is not given the opportunity to use her discretion. So on Lewis's theory it is not clear whether these signals are 'neutral', or have only descriptive content. If they have directive content, this must be conditional – perhaps something like: *if the other signals are clear, go!* The potential divergence from the Strict View comes if Lewis's claim is that discretion on the part of the consumer is necessary for a signal to lack directive content. On this interpretation, directive content will be much more widespread on Lewis's theory than on the Strict View.

Philosophers now use the game-theoretic framework pioneered by Lewis to analyse content in real-world biological systems by reference to the properties of idealised

models of those systems. The process of constructing such models begins by describing the situations in which representations are employed as signalling games, which involves identifying senders, receivers, states, actions and signals. We then find combinations of senders' and receivers' rules which are among the Nash equilibria for these games, and which also give plausible idealised descriptions of the real-world behaviour of the systems identified as senders and receivers. The contents of real-world representations are derived from the sender's and receiver's rules that govern the use of the corresponding signals in the model.

Huttegger (2007) investigated the possibility of using something like Lewis's proposal to distinguish between descriptives and directives in this framework. He considers the possibility that senders or receivers may be capable of 'deliberating' about whether to send signals, or how to act when in receipt of signals. What deliberating amounts to, for Huttegger, is the sender's gathering and taking into account information about the world other than that a particular state holds in determining whether to send a signal; or the receiver's gathering and taking into account information about the world other than that encapsulated in a particular signal in determining how to act (p. 417). He then suggests that signals have only directive content if the sender deliberates before sending them, but the receiver does not deliberate, and only descriptive content if only the receiver deliberates. So signals – that is, representations – have directive content unless the receiver deliberates, which means acting in different ways depending on other features of the situation.[9] Huttegger's suggestion can therefore be expressed as follows:

> *Descriptive content (Huttegger)*: A representation has descriptive content if and only if, in the model, the sender's rule requires that it produce the representation when and only when a particular state obtains (i.e. the rule requires that the sender does not deliberate).

> *Directive content (Huttegger)*: A representation has directive content if and only if, in the model, the receiver's rule requires that it perform a particular action whenever the representation occurs (i.e. the rule requires that the receiver does not deliberate).

This appears to be a version of the Strict View. Since game-theoretic models of real-world systems are intended to be plausible idealised descriptions of those systems, the

[9] Huttegger's account is not always interpreted in this way; see Zollman 2011, sect. 3.

patterns of behaviour of senders and receivers in these models are comparable to the patterns of behaviour that are the biological functions of real-world producers and consumers. So in particular, the claim that the receiver's rule in some model requires it to perform a particular action whenever a representation occurs is close to equivalent to the claim that the corresponding real-world consumer has the function of behaving in a particular way whenever the representation occurs. Even if there are cases where these come apart, it would remain the case that Huttegger and the advocate of the teleosemantic version of the Strict View take the same approach to the problem of distinguishing descriptives and directives, with any differences arising from their adoption of alternative frameworks for the analysis of content more generally.

Having seen that the Strict View can be expressed in the game-theoretic framework, it is worth pausing briefly to note that this also appears to be true of the Liberal View. Consider the following condition on directive content:

> *Liberal View (game-theoretic version):* A representation has directive content if and only if, in the model, the action required by the receiver's rule is invariant across all cases in which the representation makes a difference.

This condition expresses something close to the teleosemantic version of the Liberal View, provided that we understand 'cases in which the representation makes a difference' as those such that, if the input to the receiver differed only in that the representation was absent, the receiver's rule would require a different action. It captures in game-theoretic form the suggestion above that the Liberal View requires invariance across all cases in which the representation makes a contribution to fitness.

Finally, returning to our survey of the literature, Zollman (2011) and Martínez and Klein (2016) have combined the game-theoretic framework with an appeal to information theory in analysing descriptive and directive content. I will focus on Martínez and Klein's account, since it is more explicit, but both rely on the same idea. This is that representations have both descriptive and directive content when, in models constructed in the way I have just described, those representations carry equal amounts of information about states and actions. They have only descriptive content if they carry substantially more information about states than actions (in Martínez and Klein's terms, they are 'predominantly indicative'), and only directive content ('predominantly imperative') in the opposite case.

Martínez and Klein calculate the *mutual information* between representations, on the one hand, and states or actions on the other, and use these values in determining whether representations have descriptive content, directive content, or both. Mutual information in this case is the extent to which knowing whether or not the representation occurred reduces uncertainty about which state obtains or which action will be performed, and is equal to the difference between the unconditional entropy of states/actions and the entropy of states/actions conditional on the occurrence of the representation. Formally, if the possible states are S_1-S_n, the unconditional entropy of states (measured in bits) is:

$$H(S) = -\sum_{i=1}^{n} Pr(S_i) \log_2 \Pr(S_i)$$

The entropy of states conditional on the occurrence of a representation R is:

$$H(S|M) = -[\Pr(R)\, H(S|M = R) + \Pr(\neg R)\, H(S|M = \neg R)]$$

(here M is a variable over signals, which can take the values R or ¬R). The entropy of actions can be defined in just the same way. Mutual information between the representation and states is:

$$I(S; M) = H(S) - H(S|M)$$

and mutual information between representations and actions is:

$$I(A; M) = H(A) - H(A|M)$$

So on Martínez and Klein's account, a representation has only descriptive content if and only if:

$$I(S; M) \gg I(A; M)$$

has only directive content if and only if:

$$I(A; M) \gg I(S; M)$$

and has both descriptive and directive content if and only if:

$$I(S;M) \approx I(A;M)$$

I will illustrate the relationship between Martínez and Klein's account and the Strict View by examining two examples.

Martínez and Klein's account agrees with the Strict View about cases with the general form of the waggle dance, which are one important class of cases on which the Liberal View and the Strict View diverge. To illustrate this, we can consider a much-simplified version of the waggle dance scenario.

Consider the following set-up, which might be used for an experimental study of bee behaviour. On day 1, a single bee, *the watcher*, is released into an arena with a hive at the centre. In half of the trials a flowering rosebush will be placed at the north-east corner; in the other half there will be no nectar source in the arena. On day 2, a second bee, *the dancer*, is introduced to the arena, and a flowering rosebush may or may not be placed at the south-west corner – again in half of trials, independently of the presence or absence of the bush in the north-east. The dancer is given the opportunity to forage on the bush at the south-west, and then to return to the hive. When the dancer returns, the scent of roses is introduced to the hive if a rosebush is present in either corner of the arena.

We can suppose that the bees in this set-up will behave in a way that can be captured in a deterministic idealised model with the following features:

- If there is a rosebush in the north-east, the watcher will find it, and remember its location. If this is the case and the scent of roses is later introduced to the hive, the watcher will forage in the north-east whatever else happens.
- If there is a rosebush in the south-west, the dancer will find it and perform a corresponding waggle dance when it returns to the hive. If the watcher observes this dance and does not remember finding a bush in the north-east, the watcher will forage in the south-west.

The states, signals and actions in this scenario can therefore be summarised in the following table. The watcher's memories of foraging in the north-east are included as signals, because these are also inputs to the watcher's choice mechanism:

States	**Signals**	**Actions**

Rosebush only in NE	Memory	Forage in NE
Rosebushes in both NE and SW	Memory + Waggle Dance	Forage in NE
Rosebush only in SW	Waggle Dance	Forage in SW
No rosebushes in arena	No signal	Remain in hive

If we apply the formulae stated above to this scenario, bearing in mind that each row on the table is equiprobable, we get the following results. The unconditional entropy of states is 2 bits, and the unconditional entropy of actions is 1.5 bits (there are more possible states than actions, so the uncertainty associated with states is greater). The entropy of states conditional on the occurrence of the waggle dance is 1 bit, because if the waggle dance occurs there are two equally probable distinct states, and the same is true on the condition that the waggle dance does not occur. For just the same reason, the entropy of actions conditional on the occurrence of the waggle dance is also 1 bit. So the mutual information between the waggle dance and states is 2 – 1 = 1 bit, and the mutual information between the waggle dance and actions is 1.5 – 1 = 0.5 bits. The waggle dance carries more information about states, so the account tells us that it has only descriptive content.

This pattern will be repeated in many cases in which the Liberal View identifies some representation as having directive content, and the Strict View denies this. This is because the Strict View requires that there should be one particular way in which the consumer behaves when a directive representation occurs, which will tend to imply that mutual information between this representation and actions is high. The Liberal View allows that the consumer may behave in more than one way when a directive representation occurs, and that one or more of these ways of behaving may also be prompted by other representations. So on the Liberal View mutual information between directives and actions will, on the whole, be lower. However, Martínez and Klein's acccount is not straightforwardly a version of the Strict View, because there are cases that lack directive content on the Strict View, but have directive content on Martínez and Klein's account. So with respect to these cases, the account is more liberal than the Strict View.

Cases of this kind arise when representations are capable of causing more than one action, in accordance with the receiver's rule, these actions do not belong to a common kind, and the actions are performed only when the representation in question occurs. For

example, consider an internal representation formed when a potential mate is present. We can imagine that this representation causes two possible actions: when a rival is also present, it causes aggressive behaviour towards the rival; and when no rival is present, it causes courtship of the potential mate.

States	**Signals**	**Actions**
Potential mate present, rival absent	R	Courtship
Potential mate present, rival present	R+S	Aggression
Potential mate absent, rival present	S	Other actions
Neither potential mate nor rival present	0	Other actions

In this situation, the mutual information between {R, not-R} and actions is equal to the mutual information between {R, not-R} and states. So according to Martínez and Klein's account, R has both descriptive and directive content. But the Strict View entails that R lacks directive content, because what should be done when R occurs depends on S. What's more, even the Liberal View does not attribute directive content to R here, because R contributes to causing both courtship and aggression, on different occasions. So while Martínez and Klein's account agrees with the Strict View about an important class of cases, the wider picture shows that its relationship with the Strict and Liberal Views is complex.

5. Reconsidering the Advantages of the Liberal View

In section 3, I presented two apparent advantages of the Liberal View, both connected with the idea of direction of fit. In this section, I argue that these features provide only limited support for the Liberal View over the Strict View.

The first advantage was that the Liberal View entails that OFC representations, which have a functional role much like that usually associated with desires, have directive content. The Strict View does not have this consequence, because the consumer of OFC representations does not have the function of behaving in a particular way whenever a particular OFC representation is produced. When an OFC representation occurs which

attributes a positive value to getting strawberries, the animal should seek strawberries only if this course of action has the highest expected value of those available to it – which also depends on how valuable it takes other possible outcomes to be, and on how probable it takes it to be that particular actions will lead to each possible outcome.

This difference between the Strict and Liberal Views is an advantage for the latter only if there are good reasons, apart from the plausibility of the Liberal View itself, to think that OFC representations must have directive content. The line of thought suggested in section 3 was that OFC representations are like desires, directive content is like the world-to-mind direction of fit, and it is a philosophical commonplace that desires have the world-to-mind direction of fit. However, this is not especially compelling, because our main criterion for testing theories of content for subpersonal representations should be their coherence with cognitive science. If the content that one of our theories identifies in a collection of subpersonal representations fits into our best algorithmic-level explanations of what those representations are for, that is a major mark in favour of the theory, and if it does not, it is a mark against (Shea 2018). Considerations like the one raised in section 3 are secondary at best.

In fact, neuroscientists who study action selection standardly describe the states that I have called 'OFC representations' as representing the values of outcomes (e.g. Daw & O'Doherty 2013, Padoa-Schioppa 2011, Schoenbaum et al. 2009). That is, these scientists take these representations to have descriptive content concerning value. Descriptive content is compatible with directive content (or at least, we have seen no evidence to the contrary), so this in no way contradicts the Liberal View's attribution of directive content. But it does show that there is an attractive way to understand what OFC representations are doing as representations which does not require us to identify them as directives: the goal-directed system keeps track of the values of outcomes and of action-outcome contingencies, and uses descriptive representations of these features to calculate the expected values of available actions; it then selects the action with the highest expected value. In computational neuroscience, the goal-directed system is interpreted as implementing model-based reinforcement learning, which requires learning, 'a representation, for each action, of the likelihood of producing each outcome, and also a representation of how pleasurable each outcome is' (Daw & O'Doherty 2013,

p. 399).[10] So the best way we have to test attributions of content to subpersonal states does not favour the Liberal View's verdict on this case.

The second advantage of the Liberal View identified in section 3 was that it nicely reflects one possible formulation of the world-to-mind direction of fit. The idea was that desires have the world-to-mind direction of fit because they succeed when they are satisfied – they aim at realisation – and that the Liberal View reflects this, because it claims that representations with directive content themselves have biological functions which are fulfilled when certain states of affairs are brought about. This apparent advantage may well evaporate under further scrutiny, however, for two reasons.

First, there is another possible account of the world-to-mind direction of fit. This is that mental states with the world-to-mind direction of fit set success or correctness conditions on entities other than those states themselves, such as our actions. When Searle (1983, p. 8) writes that, 'it is, so to speak, the fault of the world if it fails to match the intention or the desire,' he appears to have a view like this in mind. The Strict View neatly fits this account, because it entails that representations with directive content suffice to set correctness conditions on the subsequent behaviour of their consumers. So whether the Liberal View or the Strict View has the advantage on this point depends on which of these two ways of thinking about direction of fit, each suggested by one of the original disseminators of the concept (Platts and Searle, respectively), is to be preferred.

Second, the 'very idea' of direction of fit has recently come in for trenchant criticism by Kim Frost (2014; see also Milliken 2008), who argues that no theory of direction of fit is correct. If Frost is right, the Liberal View certainly does not enjoy this second advantage, and the reason given in section 3 for suggesting that it enjoys the first advantage is also undercut. Without undertaking a detailed investigation of direction of fit we cannot say whether there is anything to be gained for either of our two views from appeals to the parallel between the mind-to-world/world-to-mind and descriptive/directive distinctions. But we can be confident that there is no obviously conclusive advantage for the Liberal View in this area.

6. Conclusion

Of the two accounts I have considered, the Strict View has the more compelling attractions; the arguments of section 4.1 show that it neatly satisfies what seem to be

[10] The reference to pleasure here is just one way of construing the nature of the value represented in the goal-directed system, which is more frequently referred to as 'reward value'.

fundamental desiderata on a theory of directive content. The considerations in favour of the Liberal View, discussed in sections 3 and 5, seem to underlie its status as the apparent default position, but are far from conclusive. However, not least because there are a range of other possible theories of directive content, it would be premature to conclude that the Strict View is the correct account (see e.g. Price 2001, ch. 6 for just one further proposal). What the attractions of the Strict View do show is that we have quite some way to go to understand the nature and prevalence of directive content.

References

Adams, C. & A. Dickinson. (1981). 'Instrumental responding following reinforcer devaluation'. *The Quarterly Journal of Experimental Psychology Section B* 33 (2): 109-121.

Anscombe, G. E. M. (1957). *Intention*. Oxford: Blackwell.

Artiga, M. (2014). 'Teleosemantics and pushmi-pullyu representations'. *Erkenntnis* 79(3): 1-22.

Balleine, B. & J. O'Doherty. (2010). 'Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action'. *Neuropsychopharmacology* 35 (1): 48-69.

Brozzo, C. (2017). 'Motor intentions: How intentions and motor representations come together'. *Mind & Language* 32 (2): 231-256.

Cao, R. (2012). 'A teleosemantic approach to information in the brain'. *Biology and Philosophy* 27 (1): 49-71.

Daw, N. & J. O'Doherty. (2013). 'Multiple systems for value learning'. In Glimcher & Fehr, eds., *Neuroeconomics: Decision-Making and the Brain* (2nd edition). London: Academic Press, pp. 393-410.

Dickinson, A. & B. Balleine. (1994). 'Motivational control of goal-directed action'. *Animal Learning and Behavior* 22 (1): 1-18.

Frost, K. (2014). 'On the very idea of direction of fit'. *Philosophical Review* 123 (4): 429-484.

Godfrey-Smith, P. (1992). 'Indication and adaptation'. *Synthese* 92 (2): 283-312.

Godfrey-Smith, P. (2013). 'Signals, icons, and beliefs'. In Ryder, Kingsbury & Williford, eds., *Millikan and her Critics*. Oxford: Wiley-Blackwell, pp. 41-58.

Grüter, C. & F. Ratnieks. (2011). 'Honeybee foragers increase the use of waggle dance information when private information becomes unrewarding'. *Animal Behaviour* 81: 949-954.

Heyes C. & A. Dickinson. (1990). 'The intentionality of animal action'. *Mind & Language* 5 (1): 87-103.

Humberstone, I. L. (1992). 'Direction of fit'. *Mind* 101 (401): 59-84.

Huttegger, S. (2007). 'Evolutionary explanations of indicatives and imperatives'. *Erkenntnis* 66: 409-436.

Jeannerod, M. (2006). *Motor Cognition: What Actions Tell the Self*. Oxford: Oxford University Press.

Lewis, D. (1969). *Convention: A Philosophical Study*. Oxford: Blackwell.

Martínez, M. & C. Klein. (2016). 'Pain signals are predominantly imperative'. *Biology and Philosophy* 31 (2): 283-298.

Millikan, R. (1984). *Language, Thought and Other Biological Categories*. Cambridge, MA. MIT Press.

Millikan, R. (1995). 'Pushmi-pullyu representations'. *Philosophical Perspectives* 9: 185-200.

Millikan, R. (2004). *Varieties of Meaning: The 2002 Jean Nicod Lectures*. Cambridge, MA. MIT Press.

Milliken, J. (2008). 'In a fitter direction: Moving beyond the direction of fit picture of belief and desire'. *Ethical Theory and Moral Practice* 11 (5): 563-571.

Morris, R., A. Bornstein & A. Shenhav. (2018). *Goal-Directed Decision Making: Computations and Neural Circuits*. London: Academic Press.

Padoa-Schioppa, C. (2011). 'Neurobiology of economic choice: A good-based model'. *Annual Review of Neuroscience* 34: 333–59.
Papineau, D. (1993). *Philosophical Naturalism*. Oxford: Blackwell.
Papineau, D. (1998). 'Teleosemantics and indeterminacy'. *Australasian Journal of Philosophy* 76 (1): 1-14.
Papineau, D. (2003). 'Is representation rife?'. *Ratio* 16: 107-123.
Platts, M. (1979). *Ways of Meaning: An Introduction to a Philosophy of Language*. Cambridge, MA: MIT Press.
Price, C. (2001). *Functions in Mind: A Theory of Intentional Content*. Oxford: Oxford University Press.
Rangel, A. & T. Hare. (2010). 'Neural computations associated with goal-directed choice'. *Current Opinion in Neurobiology* 20 (2): 262-270.
Rushworth, M., M. Noonan, E. Boorman, M. Walton & T. Behrens. (2011). 'Frontal cortex and reward-guided learning and decision-making'. *Neuron* 70: 1054-69.
Schoenbaum, G., M. Roesch, T. Stalnaker & Y. Takahashi. (2009). 'A new perspective on the role of the orbitofrontal cortex in adaptive behaviour'. *Nature Reviews Neuroscience* 10 (12): 885–92.
Schuchmann, K.-L. (1989). Release of gaping in hummingbirds (Trochilidae). *The Wilson Bulletin* 101 (3): 477-481.
Searle, J. (1983). *Intentionality: An Essay in the Philosophy of Mind*. Cambridge, UK: Cambridge University Press.
Shea, N. (2007). 'Consumers need information: Supplementing teleosemantics with an input condition'. *Philosophy & Phenomenological Research* 75 (2): 404-435.
Shea, N. (2018). *Representation in Cognitive Science*. Oxford: Oxford University Press.
Skyrms, B. (2010). *Signals: Evolution, Learning and Information*. Oxford: Oxford University Press.
Smith, M. (1987). 'The Humean theory of motivation'. *Mind* 96 (381): 36-61.
Smith, M. (2010). 'Humeanism about motivation'. In O'Connor & Sandis, eds., *A Companion to the Philosophy of Action*. Oxford: Wiley-Blackwell, pp. 153-185.
Zollman, K. (2011). 'Separating directives and assertions using simple signalling games'. *Journal of Philosophy* 108: 158-169.