Debate on the nature of representation in cognitive systems tends to oscillate between robustly realist views and various anti-realist options. I defend an alternative view, deflationary realism, which sees cognitive representation as an offshoot of the extended application to cognitive systems of an explanatory model whose primary domain is public representation use. This extended application, justified by a common explanatory target, embodies idealisations, partial mismatches between model and reality. By seeing representation as part of an idealised model, deflationary realism avoids the problems with robust realist views, while keeping allegiance to realism.

KEYWORDS

cognitive representation, deflationary realism, idealisation, indeterminacy of content, naturalising intentionality, scientific models

1 | INTRODUCTION

A guiding principle that has marked most of cognitive science since its birth about 70 years ago is the idea that cognition is made possible, at least to a large extent, by internal computational processes that manipulate internal states that bear representational content. Much philosophical work has been dedicated to exploring the notions of representation and computation, with an eye especially to accounting for them in purely naturalistic, scientifically acceptable ways. This computational-representational framework is no longer the “only game in town”, with
interesting alternatives being developed by proponents of embodied and enacted cognition that often downsize, and sometimes eliminate, appeal to representation and/or computation in explaining cognitive capacities. At any rate, that framework plays to this day an important role in informing contemporary research in the cognitive sciences, and in light of this continuing relevance, philosophical interest in the nature of cognitive representation and computation goes on.

In this paper, I focus on the notion of representation in cognitive science. Philosophical work on representation has blossomed in the 1980s and 1990s, but since the turn of the century interest seems to have waned (Godfrey-Smith, 2006). The main candidates for a naturalistic theory of representation—informational semantics, inferential role semantics, and teleosemantics—all seem to face crucial difficulties; and the solutions put forward have failed to convince all, resulting in a lack of consensus and diminishing trust on the prospects of success of the naturalisation project. In the past years, however, hope has returned, and there has been a renewed interest in shedding light on the notion of representation. Some provide fresh developments of teleo-based theories (Millikan, 2017; Neander, 2017; Shea, 2018); while others have aimed at exploring alternative views of the nature and role of representation in the cognitive sciences (Egan, 2014; Godfrey-Smith, 2006; Sprevak, 2013).

Here I will follow the latter path, suggesting a different way of understanding cognitive representations in light of recent philosophical work on the nature and role of models and idealisations in science. In a nutshell, I will argue that cognitive representations are part of an idealised explanatory model, the representational model, whose primary domain of application is the practices surrounding the use of external representations. Despite containing idealisations in its application to cognitive phenomena, I hold that the representational model captures actual features of cognitive systems: features that are relevant for explaining certain patterns of behaviour that call for representational explanation, which I call representational patterns. These considerations invite a realist view of representation, although one in which troublesome commitments of traditional theories are seen as idealisations, thereby deflating their metaphysical import. I call the resulting view “deflationary realism”.

Here is how I will proceed in what follows. In Section 2, I provide a rough taxonomy of the main views on the ontological status of representation. I explore the features of the representational model in Section 3, arguing that representational explanation has its origins in the social practices sustaining public representation use, and that it has a privileged explanatory target, that is, representational patterns. In Section 4, I analyse the limits of the representational model when applied to cognitive systems. I put forward my alternative view of representation, deflationary realism, in Section 5, and I argue that it avoids the problems with traditional views of representation. Finally, I contrast it with other recent deflationary approaches, content pragmatism and fictionalism, as well as with robust realist views, showing that deflationary realism is a distinct view (Section 6).

2 | REPRESENTATION IN COGNITION

A promising approach to representations is to see them as theoretical posits that play an explanatory role in the cognitive sciences (Cummins, 1989). In the cognitive sciences, representations are typically seen as internal, often sub-personal states that are about entities and processes in the world, in a way not too dissimilar to how pictures are about what is depicted, words are about what they refer to, and maps are about the landscapes they map. Representational
contents place conditions on how the world is or should be, if the representation is, in a loose sense, to be appropriate. By being about things in the world, representations stand in for what they represent, and by operating over their own representational states, cognitive systems are able to produce behaviour that is appropriate to the circumstances the organism finds itself in. The physical realisations of representational states are representational vehicles: the physical states that bear representational content. In the case of the brain, representational vehicles are ultimately states and processes going on in neurons and/or in populations of neurons (and possibly in glia), although this does not mean that explanations in representational terms need be cashed out in neurophysiological terms.1

A crucial contrast that sets apart different approaches to representation in the cognitive sciences is between views that insist that cognitive representations are theoretical posits that refer to real, objective entities in cognitive systems; and views that deny it. I will refer to the former as realist views, while I will call the latter anti-realist views.2

Cognitive representations, by the realist’s lights, really exist within cognitive systems. The notion is typically posited in the course of attempts to explain cognitive phenomena and intelligent behaviour, and its explanatory value is supposed to hinge on its corresponding to features that cognitive systems possess, and would possess even in the absence of any sentient being having cognitive states about those features, and independently of the explanatory purposes and practices in which those notions are employed. For anti-realists, in contrast, the notion essentially depends on conceptual apparatus that we as cognitive scientists or users of folk psychology impose on the world, and which is not committed to their being faithful to its nature and organisation.

Some of the most influential realist theories of representation belong to a family of views that I call robust realism. Robust realists aim at naturalising representation, explaining the notion in purely naturalistic and non-representational terms. They take some inner states of cognitive systems to have representational content due to their standing in some special natural relation(s) to what they represent, which bestows fairly determinate contents on them. Representations are taken to be structures in the cognitive system whose boundaries are relatively clearly definable, stable, repeatable (i.e., participate in different cognitive processes while preserving the same content), and often composable (i.e., that can be systematically combined with other representations to yield more complex representations). Robust realists are also committed to there being identifiable subsystems whose role it is to use cognitive representations: representation consumers. Informational and teleoinformational semantics (Dretske, 1981; Neander, 2017), teleosemantics (Millikan, 2017), inferential role semantics (Block, 1986), and structural representation (Ramsey, 2007; Shea, 2018) are examples of robust realist theories.

Robust realist theories of representation face several objections, and despite decades of work dedicated to tackling them, it is still unclear whether they can be satisfactorily addressed. Perhaps the most damning are the various objections from indeterminacy of content. Briefly, the worry goes that robust realist theories yield representations with wildly indeterminate, non-unique contents. Representations would thereby have several simultaneous contents, many of which are unrelated to or of a different kind than the content that would seem as the most intuitive and/or explanatorily apt. Indeterminacy of content is an unfortunate result for several reasons. First, it seems strongly counterintuitive when applied to conscious representational states, as they strike us as having

---

1For recent debate and defence of the explanatory autonomy of representational, and cognitive explanation more generally, see the papers collected in Kaplan (2017).

2For a similar carving of the conceptual space, see Joyce (2013).
determinate contents. More importantly, indeterminacy of content jeopardises the explanatory role representations are supposed to play, risking to make representational explanation vacuous or uninformative, and misrepresentation hard to accommodate, or completely impossible. A less discussed difficulty for robust realist theories is their commitment to identifiable and relatively stable representational vehicles and representation consumers. Issues with vehicle and consumer individuation surface especially in teleologically based approaches, which currently appear to be the most promising family of robust realist views (Cao, 2012; Godfrey-Smith, 2006).

On the anti-realist side, a variety of positions withhold ontological commitment to representations on often different grounds. Some believe that representations do not exist, or if they do, that they play only a minor explanatory role in the cognitive sciences, being fruitfully replaced by alternative theoretical posits coming from dynamical systems theory and neurophysiology (Van Gelder, 1995), and/or from alternative, embodied approaches to cognition (Hutto & Myin, 2013; Thompson, 2007). I remain neutral on the prospects of success of anti-representationalist views. My aim in the foregoing is to try and provide a more satisfactory version of representationalism: deflationary realism. My view will then have to earn its keep, theoretically and empirically, in comparison to anti-representationalist alternatives—although, as I point out later (Section 7), deflationary realism removes some of the grounds for seeing the two camps as stark competitors.

Other forms of anti-realism about representation reject ontological commitment to representations—denying their existence or remaining agnostic about it—but hold that nonetheless the notion plays central and/or ineliminable roles in the cognitive sciences. My deflationary realism bears interesting relationships to this family of positions, whose most prominent candidates are content pragmatism, and representational fictionalism.

According to content pragmatism, cognitive representations are not posits of cognitive theories proper, but are rather part of the extra-theoretical apparatus that allows us to connect purely computational explanations of cognition to the externally-individuated tasks we are interested in explaining in light of heuristic and pragmatic considerations (Coelho Mollo, 2020; Egan, 2014). On the other hand, according to representational fictionalism talk of cognitive representation is a useful fiction (Sprevak, 2013), or a form of pretence (Toon, 2016): When making claims about cognitive representations, scientists engage in fictional discourse—discourse that works as a useful device for understanding behaviour, but that does not commit us to the reality of cognitive representations. A central difficulty for both pragmatism and fictionalism is to justify the claim that representational talk does not ontologically commit to cognitive representations, despite their being part and parcel of our best explanations in cognitive science. In other words, they seem to flout the principle of inference to the best explanation (Sprevak, 2013; Section 6 below). I believe that deflationary realism accommodates some of the insights that motivate pragmatism and fictionalism, while keeping at bay their problems by sticking to a less-than-robust form of realism. I compare the three views in Section 6.

---

1Determinacy of content is not the opposite of vagueness. Determinate contents may include vagueness, for instance if they involve vague concepts, or if the kinds they are about are fuzzy (Millikan, 2017).
2Three indeterminacy problems have been most discussed, the disjunction problem (Fodor, 1987), the distality problem (Dretske, 1986), and the functional indeterminacy problem (Fodor, 1990). See Ryder (2009) for an overview, and Neander (2017, chapter 7) for other types of indeterminacy problems.
3This is particularly problematic for consumer-based approaches to teleosemantics, which see representation consumers as helping determine content. In these accounts, difficulties individuating consumers infect content determination, being an additional source of indeterminacy.
4Egan (2018) has recently called her view “deflationary”. I will though refer to it throughout as “content pragmatism” to avoid confusion with the “deflationary realism” I propose here.
A particularly fruitful way of understanding the explanatory appeal of the notion of representation in cognitive science is by recognising its close ties to a model of explanation that has its original application in the social practices sustaining the use of public representations (Godfrey-Smith, 2006). This “representational model” is originally targeted at capturing some of the central features that characterise the use of public symbols by agents in social interactions, and in everyday life more generally. Clear examples include linguistic symbols, traffic signs, and city maps.

Representations, on this model, are public objects that agents use in guiding thinking and behaviour towards something else. Representations act as stand-ins thanks to some property they have that makes them into adequate guides, for specific purposes, and by agents with the adequate interpretive capacities, to thinking and behaving towards what is represented. Some typical features of public representations characterise the representational model: Representations are distinct from the agents that use them, agents intentionally use them as stand-ins for specific entities in specific circumstances, and there is some relation between the representation and what it is used to represent that, when correctly interpreted, endows them with determinate contents, allowing this usage to work adequately enough.

Take a map of Berlin. If it is accurate enough, I can use it to navigate Berlin successfully, even if I had never been to the city before. The map allows me to move around Berlin in line with my varying interests and needs during the trip. I may check the map so as to get to the Pergamon Museum, which stands in my list of attractions to visit. The map allows me to reach the museum from a variety of different starting points: a handy property for travellers like me, who enjoy wandering aimlessly around a new city. It also allows me to reach the museum regardless of variations in other, non-spatial aspects of my context, such as meteorological conditions, means of transportation (if it is raining, I may prefer taking public transport rather than walking), and even my desires and needs (I may want to go to the museum not for the artworks, but rather because I am meeting a friend there, or because I have been told that the cafe is excellent and I am hungry, etc.).

It so happens that I am actually quite familiar with Berlin, and I can reach the Pergamon Museum successfully and reliably across many different contexts without the need of a map, or a list of directions, or any other kind of external representation. A natural move then is to explain this capacity of mine in terms of an internal representation, perhaps even one that is to some extent similar to a map—for example, a “cognitive map” in entorhinal cortex (Moser, Kropff & Moser, 2008)—that cognitive subsystems use to guide my behaviour.

This is an illustration of the representational model being applied to explain a cognitive capacity. The approach consists, roughly, in seeing (some) cognitive states as guides to states and events in the world, like external representations—with intentional agents replaced by non-intentional cognitive processes and subsystems that use them in the appropriate ways (Godfrey-Smith, 2006). Application of the representational model to understanding cognition leads to attempts to individuate the corresponding parts of the model in cognitive systems. We try, that is, to find in cognitive systems states and processes that play roles analogous to those of external representations, and of the agents that interpret and use them. It becomes thus vital to explain how the representational model can work without making reference to intentional agents and social interpretative practices. This is the core challenge that realist, naturalistic theories of representation try to meet: explaining by non-intentional and non-semantic means how cognitive states acquire content, and in such a
way that they can be used as stand-ins by other cognitive processes (for which appeal to social practices of interpretation is out-of-bounds).

Importantly, the distinctive explanatory target of representational explanation, and its explanatory role in the cognitive sciences more generally, constrains the ways in which the representational model is applied to cognitive systems. Representation comes to the fore when the targets of explanation are certain types of patterns of behaviour, which we tend to see as intelligent and for which appeal to representation is distinctively explanatory, that is, for which appeal to representation leads to better explanations than would be possible otherwise (Rupert, 2018; Shea, 2018). I call these explanatory targets “representational patterns”. In order to individuate more precisely the domain of application of the representational model in the cognitive sciences, we have to look at the features that its explanatory targets possess in its primary domain of application, that is to say, the features of the distinctive behavioural patterns that the social practices tied to the use of public representations make possible.

As we have seen, public representations are objects used as guides for thinking or behaving towards something else. A city map allows its user to navigate a city, reaching desired destinations from a variety of different starting points, and regardless of variations in non-spatial aspects of the situation. The map thereby enables reliable behaviour, as it allows one to reach desired locations, and it does so robustly and flexibly, inasmuch as one can reach them from different points in the city, for different reasons, and under different environmental conditions. Use of the map leads to reliable and robust behavioural outcomes.7 This, I believe, is one of the characteristic features of public representations. They enable reliable, robust and flexible interaction between agents and aspects of the world across a variety of circumstances. They do so in complicated ways, as the complexity of cartographic representations—which contain approximate spatial mappings, icons, conventional symbols, and so on—illustrates (Camp, 2007; Rescorla, 2009).

If one knows a city well, one can navigate the city reliably across many different contexts without the need of a map, or a list of directions, or any other kind of external representation. The explanandum in this case, that is, a certain pattern of (navigational) behaviour, is sufficiently analogous to the one in the external map case to warrant the application of the representational model, positing internal states and processes that function analogously to external maps, namely cognitive maps.8

Representations help explain patterns of robust, reliable, flexible behaviour due to their standing in for what they represent, allowing “surrogative reasoning” (Swoyer, 1991). Representations also help explain why in some cases behaviour reliably fails by appeal to the falsity or inaccuracy of the representations used. If a map of Berlin wrongly shows the Pergamon Museum to be in the Tiergarten park, rather than on the Museum Island, it will reliably and robustly lead its users to fail to reach the museum if that is their desired destination (if they do not know better, they will rather robustly and reliably reach the Tiergarten park)—analogously if the inaccurate map is some sort of internal representation “consulted” by navigation systems. Correct representation helps explain

---

7The degree and kind of reliability, success, and robustness varies depending on the type of representation and its level of accuracy. A list of written directions is reliable only from one point of departure, so it is not much robust (although it is robust across non-spatial contextual variation).

8It may be argued that the similarity between the two cases stems from both involving the use of a cognitive map, in one case prompted or assisted by an external map, and in the other without such an external prop. This worry is unwarranted, as the two cases are importantly different. The internal representations built while using an external map are partial and parasitic on the use of the external prop, and behaviourally useless without it—an instance of cognitive offloading. In contrast, when navigating a well-known city without the aid of an external map, successful behaviour is enabled by rich internal representations that play an analogous guiding role.
successful behaviour—that enters in a representational pattern—and incorrect representation helps explain unsuccessful behaviour, that patterns in the characteristically robust and reliable way. This explanatory role is common to both public and cognitive representations.

Representational patterns consist in the behavioural regularities underscored by robust, flexible and reliable interactions between complex organisms and their ever-changing environments across time. Representational explanation is particularly adequate to capturing such behavioural regularities, even in cases in which explanation in terms of use of external representations is not possible. These distinctive patterns of behaviour invite the application of the representational model in so far as they bear a striking similarity to the feats that we, as intentional agents, can accomplish by means of using public representations. Internal representations can then be appealed to in explaining those behavioural patterns, in analogy to public representations and the behavioural patterns they make possible.9

These considerations make clear that, while the label “representational pattern” may suggest that this characterisation is circular or uninformative, this is not the case. The distinctive properties of representational patterns—that is, behavioural robustness, flexibility, and reliability—do not presuppose representations, but rather suggest representations as the suitable theoretical posits able to explain how those properties come about. In other words, representational patterns are types of pattern of behaviour that, when taken as explananda, call for positing representations as their explanantia. (Compare: Many diseases exhibit patterns of symptoms and patterns of transmission that are best explained by appeal to the action of microbes such as bacteria and viruses, but surely the germ theory of disease was neither circular nor uninformative even before microbes could be observed.)

More schematically, the main point of this section can be roughly summarised thus (see also Godfrey-Smith, 2006):

1. Agents display patterns of behaviour marked by a high degree of robustness, reliability, and flexibility.
2. Some of these patterns of behaviour are best explained by the use of external representations by intentional agents; such explanations have specific features that can be seen as composing a (representational) model of explanation.
3. Other similar patterns of behaviour cannot be explained by the use of external representations.
4. Given their similarity to the patterns of behaviour explained by external representations, it is plausible that the adequate model for explaining them is also the representational model.
5. Therefore, we have good reason to apply the representational model to cognitive systems by positing internal representations whose nature and function is analogous to that of external representations.

4 LIMITS OF THE REPRESENTATIONAL MODEL

As I argued in the previous section, once our aim is to explain certain kinds of patterns of behaviour—which I have been calling representational patterns—application of the representational model of explanation seems to be a promising way to go. Applying the representational

---

9This is not to say that representational explanation cannot be applied to other kinds of behavioural pattern, such as context-insensitive stereotypical behaviour. The claim is rather that when applied to behaviours other than those that make up representational patterns, representational explanation may be less adequate, and it may well play only a heuristic or pragmatic role. See Section 5.
model to help explain those patterns where public representations are not in question, as in many cases of human and non-human cognition, suggests positing states and processes internal to cognitive systems that play roles analogous to those of public representations and their users: cognitive representations and representation consumers. Even though recent work in embodied cognition offers alternative, often non-representational models for at least some cognitive capacities, application of the representational model to cognition has proved to be scientifically fruitful, spawning several fields of research that have been making progress to this day (see Kriegeskorte & Douglas, 2018 for a recent review).

There are however features of public representations that are difficult to square with how states and processes in cognitive systems work. These are the places in which the representational model fails to be faithful to its extended domain of application; where its limitations when applied to the cognitive sciences come to the fore. These are also the places that invite caution before turning elements of the model into requirements on the metaphysics of cognitive representation. I do not purport here to give anything close to a theory of public representations, or of the social practices that underlie their use. For our purposes it suffices to identify some of their central features so as to see whether and to what extent they can be shared by cognitive representations. I hope that the features I will mention are relatively uncontroversial, requiring the adoption of no specific theory of public representational practices.

First, public representations depend at least partly on intentional agents that take them to be representations, use them as representations, and are able to interpret them appropriately. Public representations are also typically intended to function as representations by their creators. If one wants to stay within a naturalistic approach, appeal to agents’ intentional states in explaining cognitive representation is off the cards. Second, public representations rely on the existence of social practices of interpretation, which allow members of a culture to extract the intended meaning from representations. Since I do not read Japanese, I cannot adequately interpret Japanese linguistic symbols. There are also less obvious cases: “reading” some kinds of maps, graphs, mathematical notations, not to mention pieces of art, requires interpretive skills that rely on specific practices maintained and transmitted by communities. This is another aspect that cognitive representations cannot share with their public counterparts.

The traditional solution to these disanalogies is to transform the agents and interpretive practices that underlie public representation use into very simple input–output cognitive mechanisms following (typically) computational rules (Dennett, 1978; Godfrey-Smith, 2006). Such simple mechanisms “use and interpret” only in an extremely watered down sense: There is no intentionality involved, just the automatic workings of (computational) mechanisms. While this “homuncular” strategy seems to accommodate the agential and interpretive elements of the representational model in its application to cognitive systems, there are reasons to be less confident about the details.

Public representations typically have relatively clear boundaries in space and/or time, at least given appropriate interpretive practices. A sentence has a clear beginning and end (and clear parts), a city map can fit in one’s pocket, and pictures can be framed and affixed to the wall. Public representations are typically spatially and temporally delimitable, often portable,
and clearly separable from the agent that makes use of them. In the case of biological cognitive systems, it is doubtful that there are any neat boundaries to representational vehicles. Candidate vehicles are features of neural activity in cortical regions, whose representational “code” is still unclear (deCharms & Zador, 2000), as are their boundaries: Where the representation begins and ends is a question with a much less straightforward answer in the case of cognitive systems, given the complexity of their connectivity structure, and the ubiquity of feedback loops (Artiga, 2016b; Cao, 2012). This is not to say that it cannot be answered; and in some sensorimotor areas there are populations of neurons that allow some degree of delimitation as representational vehicles, such as cortical columns in V1.

However, even in their case the situation is rather complicated (Cao, 2019). Such populations have preferred stimuli to which they respond, but they also respond to other stimuli, although less often and less strongly. Moreover, inhibitory connections to other populations of neurons play a role in their workings. Should we include populations of neurons that respond to a certain type of stimulus non-preferentially as part of the representation? Or what about the populations involved in inhibitory activity? Or the populations that block or allow such inhibitory activity to go through? And this without even mentioning neural plasticity, degeneracy and pluripotentiality, and network reorganisation, which add further layers of complexity and instability to vehicle individuation (Anderson, 2014). In other words, individuating representations and their parts in cognitive systems is rather difficult, and there is no reason to believe that cognitive representational vehicles can be bounded, even approximately, as public representations can.

Relatedly, the clear separation between representation and user that characterises public representations is hardly applicable to cognitive systems. Although the assumption of separability plays a central role especially in teleosemantic theories, where the distinction between representation producer and consumer helps make the approach get off the ground, it is unclear whether there is any support for such distinction in cognitive systems (Cao, 2012; Godfrey-Smith, 2006). For reasons analogous to the above, not only are the boundaries of representational vehicles far from clear, but so are also the boundaries of the putative subsystems that use them as representations.

As we have seen, part of what helps fix the content of public representations are the interpretive practices socially produced and transmitted in a community. These practices enable users correctly to interpret the content of public representations, using them in ways appropriate to what they represent, even in cases where the relationship between representation and what is represented is very convoluted and opaque, as in natural language. Public representations typically have determinate, or fairly determinate content because interpretive practices allow representation users appropriately to grasp the intended content of the representation.

In the case of cognitive representations social interpretive (intentional) practices cannot be appealed to in helping fix their content, if we are to keep allegiance to naturalism. It is to this problem that robust realist theories of representation have dedicated most of their attention. They propose natural substitutes to the interpretive practices that sustain public representation, in an attempt to show that natural relations and processes can play a similar role in helping fix the content of cognitive representations. Most of the debate has focused on which combinations of natural relations and processes are able to bestow fairly determinate content on representations, as interpretive practices do in the case of public representations. The failures of existing proposals in this regard are illustrated by the many indeterminacy problems that have surfaced in the literature.

In brief, application of the representational model to cognitive systems has important limitations, if taken literally. There are at least three central posits of the representational model that do not match known properties of cognitive systems or that are difficult to account for:
bounded representational vehicles, the separability of representation and consumer, and determinacy of content. These mismatches underlie some of the crucial objections moved against robust realist theories of representation. Robust realist theories take the representational model seriously, and claim that it is at least in principle, if not in (future) practice, possible to map its posits into actual entities and processes in cognitive systems. The failure of such mapping, however, need not lead us to anti-realism about representation. In the rest of this paper, I explore a view of representation at odds with both robust realism and anti-realism: deflationary realism.

5 | DEFLATIONARY REALISM

Deflationary realism contends that the representational model, when applied to cognition, is an idealised scientific model. In so far as it embodies idealisations, the model incorporates falsities. But it does so in order to bring forth the central causal factors that make the phenomena I have been calling representational patterns possible. Cognitive representation, as a theoretical posit of the model, approximates real properties of cognitive systems, despite the idealisations it involves.

Contemporary philosophy of science increasingly recognises the significance of models and idealisations in science, and there is a rich ongoing debate about their nature and role (Morrison, 2015; Potochnik, 2017; Teller, 2012; Weisberg, 2013). My understanding of scientific idealisation is close to Potochnik’s (2017): Idealisations are distorted representations of aspects of a target system, which by means of their partial falsity bring forward focal causal features of the system relevant to explaining a phenomenon of interest. For instance, an evolutionary model may assume infinite population size (an idealisation) to bring into focus the specific contributions made by natural selection to the evolution of traits within a finite population, disregarding in this way other causal influences such as genetic drift. Idealisations are, for Potochnik, inevitable in science, given the complexity of the world, our explanatory aims, and the limitations of our cognitive capacities and epistemic reach.

Idealised models permeate science: They are essential for achieving central scientific epistemic aims, such as explanation, prediction, and understanding (Bokulich, 2011; Potochnik, 2017). Idealised models are explanatory (at least) in so far as they truly represent, thanks to the simplifications and distortions they embody, focal patterns of (causal) dependency that are relevant for explaining a phenomenon of interest. Idealisations are partly false, so that they can be partly true: The falsities they embody are instrumental for bringing forth actual causal structures of particular explanatory interest, which would be indiscernible otherwise, lost in the enormously intricate web of causal dependencies and factors underlying most real phenomena.⁵

Deflationary realism holds that the cognitive representational model is an idealised model: A partially distorted and simplified picture of the causal features that contribute to bringing about the characteristic patterns of behaviour that it aims to explain, namely representational patterns. There are at least three grounds for taking the cognitive representational model to

---

⁵Some theorists have argued that we can more fruitfully consider the relationship between model and target in terms of degrees of similarity, with some models saying more true things than others about the modelled systems (Morrison, 2015; Weisberg, 2013). However, similarity measures are notoriously difficult to determine in principled ways. Although not much hinges on this for my purposes, I will follow Potochnik (2017) in taking idealised models to be true of the causal patterns they capture, while being false of phenomena as a whole, due to the idealisations they contain.
embody idealisations. First, as we have seen, it is a model that has its origins and primary motivation elsewhere. Application of the representational model to the cognitive sciences features limitations and shortcomings due to the fact that it is an extension of the application of concepts that belong to one domain—representational social practices—to a rather different one, that of cognitive states and processes. Second, cognitive systems are highly complex, and an indeterminate number of interacting causal factors are responsible for bringing about cognition and intelligent behaviour. Complex systems can be modelled only partially, and idealisation becomes crucial to make models tractable, understandable, and epistemically useful by simplifying some factors and ignoring others. Third, and relatedly, several of the relevant causal factors for explaining intelligent behaviour—for example, fine-grained neural workings, evolutionary history—are either currently or in principle epistemically inaccessible to us. In consequence, we can do little else than to idealise them in our models, in light of our best guesses, or else idealise or abstract them away.

According to deflationary realism, insofar as the representational model embodies idealisations, a degree of mismatch between its posits and the actual properties of cognitive systems is inevitable. I take this mismatch to underlie the philosophical conundrums that have plagued robustly realist theories of representation. These arise not because the theories are incoherent or false, but rather because they fail to recognise the idealisations they embody. When these conundrums are revealed as what they are, namely unsurprising offshoots of an idealised model, their metaphysical gravitas gives way to a healthier, epistemic modesty. In other words, deflationary realism suggests that the puzzles that have plagued the naturalistic project about representation are not, as assumed by most parts in the debate, deep metaphysical problems that threaten the naturalistic credentials of cognitive representations, and thereby their employment in explanations in the cognitive sciences. They are rather due to epistemic and modelling limitations.

Seeing representational vehicles as bounded is a simplifying assumption, one that idealises away many of the intricate causal influences that parts of the cognitive system effect on each other. This allows us to bring into focus in a tractable way the crucial causal contributions that specific parts of the system make to the behaviour of interest. Similarly, the separability of representations and their consumers is an idealisation: It ignores the close interdependencies between most parts of the system, simplifying its organisation so as to bring to the fore a useful distinction between different functional roles; distinction that does not exist in such a neat way in the actual system, but that makes it more comprehensible and tractable. This also applies to types of explanation that focus on coarse-grained functional and computational levels of description of cognitive systems, which beside idealisation also make use of abstraction to bring into focus the relevant explanantia. These types of explanation tend to remain non-committal about the vehicles that realise the posited functional or computational structure, abstracting away from fine-grained physical details insofar as they are taken to be of little or no relevance to their explanatory aims. In short, they abstract away the details that impede appreciation of the explanatory causal pattern for the phenomena they seek to explain, positing neatly distinguishable and functionally/computationally bounded units and processes (Weiskopf, 2017).

These aspects of the representational model work as idealisations. Even if it should turn out that such aspects of the representational model never strictly apply to cognitive systems, that is, that representational vehicles never have clear boundaries, and can never be fully separated from the subsystems that use them; they would still capture explanatorily relevant causal features of cognitive systems. Idealised (and abstract) models can be explanatory, as they reveal and bring into focus the subset of causal factors most relevant for explaining a phenomenon of interest.
These two idealisations embodied in the cognitive representational model are mostly due to the causal complexity of cognitive systems. Other idealisations may also be partly motivated by our epistemic limitations. This is the case, I take, of determinacy of content. Indeterminacy of content is widely taken to be the unavoidable rock that theories of representation need to navigate around, on pain of foundering. Deflationary realism suggests that (a) indeterminacy issues are to be expected for any theory of representation, given their partiality and idealised nature; and (b) that the explanatory and predictive success of the content ascriptions they yield is a sign of their capturing an important part of the truth. Let me briefly expand on these points, starting with the former.

A first source of indeterminacy may be a misapplication of the representational model to systems that do not really instantiate representational patterns of behaviour. Imposing the requirement that contents be fairly determinate on relatively simple systems, such as the widely discussed toy example of frogs’ tongue-snatching mechanism, can be misleading. There may be no grounds for insisting that the mechanism represents either “fly”, “frog food”, or “moving black dot”—and similarly for other philosophical toy cases—inasmuch as (toy) frogs do not really instantiate the flexibility, robustness, and reliability that characterise representational patterns. For instance, in the foregoing toy case, frogs do not change their behaviour in a different environment (labs, instead of their natural habitat), and they do not cease to snatch their tongues at non-nutritious items even after considerable (unsuccessful) experience with them. These are cases in which theorists have tended to commit to elements of the representational model that do not approximate much anything in the systems under analysis: In such cases, application of the representational model may be of only heuristic value, given the absence of the proper kind of explanandum. According to deflationary realism, the “fly”, “frog food”, or “moving black dot” debate is a spurious one, since such cases do not count as candidates for proper representational explanation, and therefore are not relevant to the naturalistic project about representation.13

Another source of indeterminacy—this one non-spurious—is due to the often used idealisation that contents consist in linguistically expressible propositions. Even if there were reasons to believe that simple systems have representations with determinate contents, there are no grounds for holding that such contents can be captured in linguistically expressible propositions. The ascription of propositional contents to non-human animals (and human subpersonal states), especially, is often an idealisation, given that plausibly their cognitive systems employ representational formats that are non-propositional in nature (Beck, 2013; Camp, 2009). The content of map-like representations, for instance, cannot be properly captured by a linguistic expression, but it can be nonetheless determinate (Beck, 2013; Camp, 2018; Rescorla, 2009). That several different ascriptions of linguistically expressible contents in such cases seem equally adequate is not thereby diagnostic of indeterminacy of content, but rather of the limits of our idealised ways of capturing those contents.

Moreover, the fact that the frog toy case has been widely discussed in the philosophical literature is interesting in itself. We have here a radically idealised model that leaves out most of the details about anuran perception that play, and played, a role in the selection of the relevant mechanisms, their workings, and their ties to behavioural control.14 That philosophical

---

13Thus, the foregoing account is at odds with very liberal accounts of representation, such as Artiga’s (2016a).
14For a detailed and empirically-informed treatment of anuran perception in connection to the debate on representation, see Neander (2017, chapter 5). Neander purports to defend a modest robust realist theory in light of this treatment, but one that admittedly relies on questionable assumptions about the proper explanatory targets of representational explanation.
discussion has largely taken an idealised model as a bona fide case study, one supposed to
decide between rival theories, is more informative of the confusions that follow from ignoring
the nature and roles of abstraction and idealisation in such simplified models, than of the the-
ory of cognitive representation to be preferred.

Other modelling and epistemic limitations can also be sources of apparent indeterminacy of
content. The naturalistic substitutes to the content-fixing interpretation practices of public rep-
resentations are unlikely to be fully satisfactory replacements. The complexity of the factors that
underlie representational patterns is plausibly such that no theory can fully account for them.15
Many of the potentially relevant factors are moreover epistemically opaque to us—we have very
little epistemic access to the details of the evolutionary history of cognitive systems, and we
have little idea of what entities in cognitive systems are candidates for standing in the relevant
naturalistic relations to things in the world. Given our epistemic limitations, our theories can
only capture those factors in a rather coarse-grained way, giving us at most an idealised, partial
grasp of the full story, which is though very often good enough for explanation and prediction.

According to deflationary realism, indeterminacy of content stems from the attempt to fit a
complex system into a relatively simple model. The partial grasp on the relevant cognitive
workings that existing theories of representation provide, their limited focus on one or some
aspects that are relevant in bringing about representational patterns result in content indetermi-
nacies. The variety and complexity of the factors relevant for fixing content—occurrent and
historical—and the epistemic opacity of many of them, makes it so that all our extant theories
of representation can only capture a subset of them, and moreover only with considerable sim-
plification and approximation. As a consequence, representational ascriptions involve ideal-
isations. Different theories employ different idealisations and emphasise different focal causal
patterns, leading to different content ascriptions.

Attention to the details left out or idealised away by any specific theory may invite criticism
of their preferred content ascriptions, as they inevitably do not include all the relevant details
for fixing content. Different content ascriptions, and criticisms based on missing details, have
tended to invite often despairing metaphysical conclusions about indeterminacy of content, or
at any rate, series of objections and counter-objections on the part of proponents of different
theories. In contrast, deflationary realism stresses that, when it comes to domains as complex
and epistemically opaque as this, idealisation, abstraction, and thereby partiality and multiplicity
of models is to be expected.

In brief, indeterminacy of content is the result of the limitations of our models and of our
knowledge; limitations that are unlikely to ever go away, given the complexity of the systems
modelled, the epistemic opacity of the relevant factors, and the consequent partiality of our
models. That indeterminacy of content marks all our attempts at giving a realist theory of repre-
sentation is not the expression of a metaphysical problem, but rather a consequence of the
adoption of an idealised model to explain certain phenomena. Incompleteness and opacity not-
withstanding, such partial grasps over the workings of cognitive systems furnish helpful approx-
imations of the core aspects of what makes representational patterns possible.

Regarding (b), deflationary realism holds that the partiality and multiplicity of representa-
tional models is no justification for anti-realist or quietist views of representation. Such models
can be compatible, each capturing a subset of the relevant causal patterns underlying the phe-
nomenon of interest (Chakravartty, 2010; Morrison, 2011). The caveats above notwithstanding,

15This is plausibly true of public representations as well. Much that calls for explanation is hidden in the superficially
straightforward notion of “interpretive practice”.
let us take once again, for simplicity's sake, the toy case of the frog. Informational theories typically ascribe to frogs the content “moving black dot” or similar, while consumer-based teleosemantics ascribes the content “frog food” or similar. Despite their apparent divergence, both theories provide strikingly similar explanations and predictions in the cases of most interest, namely when frogs are in their natural habitats, where moving black dots tend very often to be frog food. Both content ascriptions adequately explain and predict frogs’ behaviour in their habitats. What differentiates the two families of theory, motivating their different content ascriptions, is partly their different focus: Informational theories emphasise the sensory capacities of organisms, while teleosemantics mostly stresses the evolutionary rationale of the relevant mechanisms. This different focus underlies their divergences, especially when it comes to their categorisation of successful and unsuccessful instances of representational use.

By deflationary realism’s lights these divergences do not reveal any fundamental—and thus metaphysically worrying—incompatibility, but rather their complementary, both kinds of theory capturing different, albeit largely overlapping, parts of the story. The most sophisticated theories of representation on offer mostly agree on the basic ingredients for a theory of representation: natural information; evolutionary and learning processes; functions; and, to a lesser extent, structural correspondence. These converging stories never really meet at a convergence point, but the fact itself that they share so much suggests that they are all, in somewhat different ways, tapping into something true about how cognitive systems work, something that we should ontologically commit to. They are tapping into the fact that these systems quite often make use of cognitive representations in guiding thinking and intelligent behaviour, even though we can provide only partial and idealised models of their nature, implementation, and origins.

The partiality and multiplicity of representational models, and the idealisations they embody, do not hurt the explanatory value of representational models in explaining representational patterns of behaviour. Ascribing determinate representational contents to internal states of cognitive systems helps explain certain patterns of behaviour, even if those contents are partial and idealised. And even if the notion of representation only approximately captures the properties of the cognitive states and processes responsible for those patterns, it identifies relevant causal dependencies for explaining the latter. Representational posits pick out subsets of actual causal dependencies between entities and processes in cognitive systems and their environments—ones focal to explaining and predicting representational patterns—although without providing anything close to a true complete depiction.

In sum, according to deflationary realism the representational model (or models) largely works in explanations in the cognitive sciences because it approximates many important features of actual cognitive systems, without though being a perfect fit. We should thereby be realists about cognitive representation, but in a way that concedes that what we are being realists about is only approximately what our models posit. Relatedly, deflationary realism argues that the conundrums that have exercised philosophers interested in naturalising representation are epistemic, rather than metaphysical. Once a proper appreciation of the epistemic challenges involved in understanding and modelling cognitive systems is in place, the apparent lack of consensus among theories of representation is not only to be expected, but their considerable agreement is brought more clearly to the fore. Together with the explanatory and predictive success of representational explanation in the cognitive sciences, this invites a modest realist take towards cognitive representations: one that recognises that we are on to something, and even have a pretty good idea of what that something looks like, while also recognising that our models have limitations, and that many of their aspects cannot be precisely applied to the systems under investigation.
6 | DEFLATIONARY REALISM AS A DISTINCT POSITION

Where does this picture leave us for what regards the ontological nature of representation? At first glance, deflationary realism may look like a form of pragmatism or fictionalism. We apply the representational model to cognitive systems because it is a useful tool to explain certain interesting phenomena. But this does not require or suggest that we should thereby commit to the actual existence of cognitive representations. They may be only fictions in a model, or part of a gloss which, albeit inevitable, does not justify ontological commitment to its posits. What makes this deflationary model-based approach any different from those anti-realist (or at best quietist) positions?

Content pragmatism sees appeal to representation as part of an intentional gloss that, while helpful in so far as it allows theorists to track and grasp the relevance of internal computational states and process to the explanation of externalistically individuated tasks, is not part of the theory proper. The latter has recourse only to computational states and processes, representational content, or what Egan calls “cognitive content”, being ascribed purely for heuristic and pragmatic reasons. Interpreted in the intentional gloss as representational states, these computational states and processes, which the account is realist about, are then seen as representational vehicles.16

In contrast to content pragmatism, deflationary realism follows much contemporary philosophy of science in seeing models and idealisations as part and parcel of scientific theorising: Science is permeated with idealisation, which is in most cases essential to theories, explanations, and predictions. As parts of an idealised model explanatory of certain specific patterns of behaviour, representations are part of cognitive theories proper, and not a gloss external to them. Therefore, deflationary realism shares content pragmatism’s ontological commitment to computational vehicles. But it departs from content pragmatism in insisting that computational states and processes are also, at least often, representational states and processes, and that this fact is part of the theory proper, rather than of an extra-theoretical gloss.17

The version of fictionalism closer to deflationary realism is Sprevak’s (2013). In the picture he presents, representations are useful fictions, and scientific statements about representation are not aimed at truth, but rather, at best, at truth-in-the-fiction (i.e., in the idealised model). He even briefly considers understanding the pertinent kind of fiction as scientific idealisation (p. 556). Inasmuch as idealised models are not literally true of what they model, the representational model, when applied to cognitive systems, may be seen as a fiction (Frigg, 2010). It may therefore be argued that deflationary realism is not really a form of realism, since what it invites ontological commitment to are not cognitive representations per se—as they are part of an idealised model—but rather different kinds of entity with which the latter share some properties. Cognitive representations would thus be useful fictions: fictions because they do not really exist, given their status as idealisations, and useful because they approximate actual states and processes well enough to allow prediction and intervention.

Deflationary realism resists this conclusion. It admits that some features of the representational model are idealisations, distortions of reality. But a proper understanding of the nature of scientific idealisation suggests that this does not preclude the representational model from capturing actual causal features of the system most essential to explaining the phenomenon of

---

16 I thank an anonymous referee for prompting this clarification.
17 For a defence of the essential role of representations in cognitive (computational) theories, see Burge (2010) and Rescorla (2012, 2017).
interest (Bokulich, 2011; Potochnik, 2017). This distinguishes deflationary realism from anti-realist approaches: It ties idealisation to truth sufficiently strongly to resist the anti-realism (or quietism) about representations that fictionalism and pragmatism recommend. Idealised models can be explanatory and reveal true causal features of modelled systems.

Importantly, as Sprevak (2013) points out, fictionalism can arguably also help capture true causal features of systems: Fictional representational contents can label neural states and allow us to track their causal relations in ways that are relevant for predicting behaviour (this line of defence is also available to the content pragmatist). Deflationary realism, however, argues for a stronger claim. Representational ascriptions are much more substantive than mere labels—they capture, albeit with some imprecision and approximation, features that cognitive systems really possess.

Moreover, the “labelling” strategy is rather shaky grounds on which to defend fictionalism. Labels are typically arbitrary, and their conditions of application extremely liberal. This does not seem to capture the role that representational ascription plays in cognitive science, where representational ascriptions are assessed for their fit with various sources of data, and divergences between ascriptions are seen as substantive issues, rather than as a matter of terminological preference. An illustrative example is the discussion about the representational capacities of the fusiform gyrus in the inferior temporal cortex. Initially hypothesised to be involved only in the processing of faces, it became known as the fusiform face area (FFA). Further evidence supported the idea that the FFA is rather involved in the processing of complex stimuli for which subjects have a degree of expertise, such as houses, birds, and faces (Gauthier, Skudlarski, Gore & Anderson, 2000). The discussion is not about how best to label the FFA. After all, the label already stuck, and it keeps being used even though the face–processing hypothesis has lost support. The discussion is substantive: Does the FFA contribute to representational capacities involving only faces and face-like stimuli, or are its representational contributions broader?

If, on the other hand, fictionalists would insist that representational labels are substantially constrained by explanatory and predictive considerations, as the role representations play in the cognitive sciences suggests, it becomes difficult to motivate withholding ontological commitment to representations. This leads us to the main shortcoming of views of representation that accept the explanatorily essential role played by representations in the cognitive sciences, while denying or withholding ontological commitment to them—views like fictionalism and pragmatism. As Sprevak (2013) argues, such views have trouble with the principle of inference to the best explanation (IBE), whose importance throughout science, and especially in the special sciences, is hard to understate.

In the case of fictionalism, the problem can be condensed in the following question: If the representational fiction is so explanatorily and predictively powerful, what grounds are there for thinking it is a fiction? In the case of pragmatism, analogously, the question is: If ascription of representational content (“cognitive content”) is ineliminable from explanations of cognitive capacities, what grounds are there to claim that they are extra-theoretic, mind-dependent and pragmatically-motivated, rather than capturing real features of cognitive systems?

The underlying motivation for such views is scepticism about the prospects of naturalising representation, in light of the difficulties that the naturalistic project has faced. Deflationary realism argues that this scepticism is misplaced: There has been a general tendency to draw

---

18 Alternatively, we can see idealised models as approximating truth well enough to justify ontological commitment to their posits (Weisberg, 2013).
metaphysical conclusions from epistemic limitations, and quite natural and expectable ones to boot. Fictionalism and pragmatism admittedly want to reap the benefits of realism without paying its costs. But this balancing act founders when confronted with IBE (Sprevak, 2013). Deflationary realism, in contrast, does not suffer from this problem. By keeping allegiance to realism, it makes clear why appeal to representation is so important, and possibly ineliminable, from the cognitive sciences. The representational model captures those real causal features of cognitive systems that are directly causally relevant to explaining certain phenomena in the world: representational patterns. The causal and explanatory role of cognitive representations remains thereby unscathed. Consequently, there is no challenge to scientific explanatory principles: The close tie between explanation and ontological commitment that the principle of inference to the best explanation dictates is not put in jeopardy.

Finally, as suggested in Section 5, deflationary realism takes seriously the insights of robust realist theories of representation. But it insists that these theories provide partial representational models, divergences between them being differences on which aspects of the causal structure of cognitive systems, and their history, are given more emphasis. Under this deflationary light, the preferred natural relations that different robust realist theories appeal to are some of the relevant factors underscoring the existence of representational patterns, and consequently the value of representational explanation.19

It is unlikely that this relatively modest view is what many robust realists are pursuing. The centrality given to indeterminacy of content problems, and the high stakes ascribed to these debates as saving or sinking the prospects of naturalising representation and thereby of a naturalistic cognitive science, suggest that the aims of these theories are more metaphysically ambitious. Moreover, debates between robust realists have mostly taken the form of a fight between competitors, while deflationary realism sees these theories as mutually compatible. Modest versions of existing realist views along the lines above are therefore not at odds with deflationary realism, but are rather instances of it.

7 | CONCLUDING REMARKS

Before concluding, I would like to come back briefly to the dispute between representationalist and anti-representationalist views of cognition adumbrated in Section 2. The view I have been proposing has it that the representational model is an idealised, partial model particularly suited to explaining certain patterns of behaviour. Given the complexity that marks cognitive systems and their interactions with their material and social environments, it is inevitable that different research aims and methods lead to a diversity of partial models, each bringing into focus one or a few explanatorily relevant factors, contributing a part of the story of how cognition and intelligent behaviour are brought about. Embodied and enactive approaches to cognition offer distinct models, in light of often different explanatory aims. Representationalism,

---

19This is in the spirit of Dennett’s (1991) professedly mild realist view. However, it differs from it in at least three central aspects. First, Dennett’s target are propositional attitudes, that is, personal level states such as beliefs and desires, while the deflationary realism I propose targets mainly subpersonal states and processes relevant to the cognitive sciences. Second, Dennett relies on claims about the nature and features of what he calls real patterns, which need not be causal. In contrast, deflationary realism relies on the epistemic practices of modelling and idealising in science, which are aimed at revealing causal explanatory patterns and structures. Finally, he seems to suggest scepticism about internal representational and computational vehicles. While deflationary realism remains non-committal about many of the fine-grained features possessed by the states and processes that representations, as parts of idealised models, capture or approximate, it does not embrace scepticism about representational and computational vehicles.
once seen under the light that I recommend, is not incompatible with such approaches, which
tend to downsize or eliminate appeal to representations. Rather, we have a rich variety of par-
tial models, embodying different idealising assumptions and theoretical abstractions, all of
which can be legitimately explanatory. Representationalism and anti-representationalism are
thus not competitors, but candidates for pluralistic integration (Mitchell, 2003).20

ACKNOWLEDGEMENTS
I am indebted to Nicholas Shea, Michael Pauen, Mark Sprevak, Margherita Arcangeli, and the
audience of the workshop on Representation in Cognitive Science, Bochum 2020, for helpful
comments on previous versions of this paper.

ORCID
Dimitri Coelho Mollo https://orcid.org/0000-0002-0464-3535

REFERENCES
Anderson, M. L. (2014). After phrenology: Neural reuse and the interactive brain. Cambridge, MA: MIT Press.
Artiga, M. (2016a). Liberal representationalism: A deflationist defense. *Dialectica*, 70(3), 407–430.
Artiga, M. (2016b). Teleosemantic modeling of cognitive representations. *Biology and Philosophy*, 31, 483–505.
Beck, J. (2013). Why we can’t say what animals think. *Philosophical Psychology*, 26(4), 520–546.
Block, N. (1986). Advertisement for a semantics for psychology. *Midwest Studies in Philosophy*, 10(1), 615–678.
Bokulich, A. (2011). How scientific models can explain. *Synthese*, 180, 33–45.
Burke, T. (2010). Origins of objectivity. Oxford: Oxford University Press.
Camp, E. (2007). Thinking with maps. *Philosophical Perspectives*, 21, 145–182.
Camp, E. (2009). A language of baboon thought. In R. W. Lurz (Ed.), *The philosophy of animal minds*. Cambridge: Cambridge University Press.
Camp, E. (2018). Why maps are not propositional. In A. Grzankowski & M. Montague (Eds.), *Non-propositional intentionality*. Oxford: Oxford University Press.
Cao, R. (2012). A teleosemantic approach to information in the brain. *Biology and Philosophy*, 27, 49–71.
Cao, R. (2019). Computational explanations and neural coding. In M. Sprevak & M. Colombo (Eds.), *The Routledge handbook of the computational mind* (pp. 283–296). London; New York, NY: Routledge.
Cash, M. (2009). Normativity is the mother of intention: Wittgenstein, normative practices and neurological rep-
resentations. *New Ideas in Psychology*, 27, 133–147.
Chakravartty, A. (2010). Perspectivism, inconsistent models and contrastive explanation. *Studies in History and Philosophy of Science Part A*, 41(4), 405–412.
Chirimuuta, M. (2020). Charting the Heraclitean brain: Perspectivism and simplification in models of the motor
cortex. In M. Massimi & C. D. McCoy (Eds.), *Understanding perspectivism: Scientific challenges and methodo-
logical prospects*. London; New York, NY: Routledge.
Coelho Mollo, D. (2020). Content pragmatism defended. *Topoi*, 39, 103–113.
Cummins, R. C. (1989). *Meaning and mental representation*. Cambridge, MA: MIT Press.
de Charms, R. C. & Zador, A. (2000). Neural representation and the cortical code. *Annual Review of Neuroscience*, 23, 613–647.
Dennett, D. C. (1978). *Brainstorms: Philosophical essays on mind and psychology*. Cambridge: Bradford Books.
Dennett, D. C. (1991). Real patterns. *The Journal of Philosophy*, 88(1), 27–51.
Dretske, F. (1981). *Knowledge and the flow of information*. Oxford: Basil Blackwell.
Dretske, F. (1986). Misrepresentation. In R. Bogdan (Ed.), *Belief: Form, content, and function* (pp. 17–36). Oxford: Oxford University Press.
Egan, F. (2014). How to think about mental content. *Philosophical Studies*, 170, 115–135.

20Though see Chirimuuta (2020) for a less optimistic point of view.
