Constraint-evading surrogacy: the missing piece in Radical Embodied Cognition’s non-representationalist account of intentionality?

Andrew Robinson1 · Christopher Southgate1

Abstract
Radical Embodied Cognition (REC) is an anti-representationalist approach to the nature of basic cognition proposed by Daniel Hutto and Erik Myin. While endorsing REC’s arguments against a role for contentful representations in basic cognition we suggest that REC’s ‘teleosemiotic’ approach to intentional targeting results in a ‘grey area’ in which it is not clear what kind of causal-explanatory concept is involved. We propose the concept of constraint-evading surrogacy as a conceptual basis for REC’s account of intentional targeting. The argument is developed via a reconsideration of Peter Godfrey-Smith’s ‘environmental complexity thesis’. We show how constraint-evading surrogacy clarifies REC’s account of intentionality by specifying the role of external natural signs and certain internal states (REC’s ‘Local Indexical Guides’). We argue that the resulting view of intentionality can be extended to proto-biotic autocatalytic systems, offering support for the life-mind continuity thesis. We leave open the question of whether it is helpful to consider constraint-evading surrogacy to be a form of (non-contentful) representation, but we suggest that there is a prima facie case for such a view.

Keywords Intentionality · Representation · Semiotics · Radical embodied cognition · Environmental complexity thesis · Autopoiesis · Origin of life

1 The grey area in REC’s ‘teleosemiotic’ programme

Radical embodied cognition (REC) is an important but not (yet) mainstream approach to the philosophy of mind (Hutto & Myin, 2013, 2017). REC is aligned with other E (embodied, enactive, extended, embedded, and ecological) approaches
to cognition (Chemero, 2011; Wilson & Galonka, 2013; Gallagher, 2017: 43; Hutto & Myin, 2017: 1; Shapiro, 2019). To these approaches to embodied cognition (EC) REC adds its radical (R) anti-representationalism. The distinguishing characteristic of the representationalism targeted by REC is the notion of semantic ‘content’. The notion of content that REC denies as a feature of basic cognition is, as Hutto and Myin put it, “any notion of content that assumes the existence of some kind of specified correctness condition. To be in a contentful state of mind is to take (‘represent’, ‘claim’, ‘say’, ‘assert’) things to be a certain way such that they might not be so” (Hutto & Myin, 2017: 10). According to REC, the idea that cognition necessarily involves representational content is mistaken. REC-ers hold that basic cognition does not operate by using internal representations of the external world. Furthermore, they hold that such non-representation-using cognition is able to account for quite sophisticated cognitive achievements, including some aspects of memory (Hutto & Myin, 2017: 121) and, in line with Gibson’s ecological psychology, affordance perception (Myin, 2020; cf. Gibson, 1979).

In this paper we suggest that REC’s programme, which we largely endorse, leaves a ‘grey area’ which would benefit from some conceptual clarification. The clarification we offer relates to the conceptual apparatus that is involved, or ought to be involved, in REC’s account of intentional targeting; that is, in the aspect of the programme that REC calls ‘teleosemiotics’. In order to explain why we see this as a grey area in REC’s approach we suggest it is helpful to characterise what we see as four aspects of REC’s programme, namely, 1) a methodological strategy, 2) a default conceptual lens, 3) a deflationary manoeuvre, and 4) a terminological decision.

REC’s methodological strategy is to focus on refuting the prevailing consensus that cognition (even basic cognition) involves representational content. Specifically, REC chooses to target the problematic nature of representational content rather than addressing the nature of representation in general. REC’s methodological strategy comes out clearly in its response to Rowlands’ argument that there are so many different representational concepts at large that defeating just one kind of representationalism does not achieve any theoretical progress (Rowlands, 2017). We concur with REC that its methodological strategy circumvents Rowlands’ criticism, and that it does so because of the centrality of contentful representation to current cognitive science (Hutto & Myin, 2018). REC claims that divesting basic cognition of reference to representational content would amount to a revolution in cognitive science and we agree.

We suggest that in addition to this methodological strategy REC adopts a default conceptual lens. The conceptual lens in question is that REC sees no explanatory utility in the concept of non-contentful representation. For REC, representation is synonymous with content (e.g., Hutto & Myin, 2013, 12–13). REC’s default conceptual lens is not strictly necessary to its main methodological strategy, which, as noted above, is primarily a criticism of the view that basic cognition involves representational content. REC’s default lens is inherited from the cognitive science programme that it seeks to revolutionise. Standard cognitive science usually equates representation with representationalise. REC’s tendency is to do the same.
REC’s *deflationary manoeuvre* is well known: the deflation of ‘teleosemantics’ (Millikan, 1984) into a teleo-functional account of intentionality (Hutto & Myin, 2013: 78–79). This deflation dispenses with the ‘content’ aspect of Brentano’s concept of intentionality but retains its targeting, ‘aiming’ aspect (cf. Brentano, 1973[1874]: Book Two, I, §5). Another way of expressing this is that REC denies that teleosemantics is able to deliver a naturalistic account of intentionality with an ‘s’ (semantic reference) but accepts that it offers a basis for intentionality with a ‘t’, that is, intentional ‘directedness’ (Hutto & Myin, 2013:79). Central to REC’s account of intentional directedness is the way in which organisms make behavioural responses to ‘natural signs’ in their environments (Hutto, 2008: 51). REC regards such responsiveness to natural signs as the basis of a basic kind of intentionality: ‘Ur-intentionality’ (Hutto & Myin, 2017: 114–120). REC calls its account of adaptive responsiveness to natural signs ‘teleosemiotics’ (Hutto, 2008: 41–63; Hutto & Myin, 2013: 78–82).

REC does not think ‘natural signs’ involve representation, holding them instead to be non-content-bearing features of the world which enable the organism to ‘intentionally target’ aspects of its environment (Hutto & Satne, 2015: 530; Hutto & Myin, 2017: 52, 92). REC’s choice of terms such as ‘teleosemiotics’ and ‘natural signs’ to label what remains of intentionality after any dependence on the concept of contentful representation has been removed (through its deflationary manoeuvre) is what we refer to as its *terminological decision*. This decision creates a tension with semiotic theory, the original philosophical home of terms deriving from the Greek *semeion* (sign). In the field of semiotics natural signs (indexical signs, in semiotic terminology) are taken to be a specific instances of the general category of ‘representamen’ (see, for example, Short, 2007: 53–56). In other words, where REC’s use of terms such as ‘teleosemiotics’ and ‘signs’ is not intended to imply the operation of any form of representation (contentful or otherwise), in semiotic theory the use of such terminology implies representation of some kind (usually non-contentful). From the semioticians’ point of view, semiotics is not deflated teleosemantics; semiotic theory does not stipulate content as an inherent feature of representation and would see teleosemantics (and standard cognitive science) as peculiar in thinking that it is. When a semiotician hears the word ‘teleosemiotics’ used in the context of analysing the role of ‘natural signs’ in intentionality they will tend to hear this as implying some kind of notion of (non-contentful) representation (see Short, 2007: 289–316, especially p. 301).

REC is untroubled by this because all it means by a ‘natural sign’ is a ‘worldly offering’ (Hutto, 2008: 50) which, because of its historical correlation with certain other worldly states of affairs, has enabled a (non-representational) ‘indexical’ or ‘targeting’ relationship between the organism and its environment. Teleosemiotics is REC’s terminological way of signaling that it has deflated the teleosemantic programme while retaining some of the latter’s features; REC’s use of the term ‘teleosemiotics’ is not intended to indicate the retention of any dependence on representational concepts. This is in keeping with the ‘default conceptual lens’ identified above.

Of course REC is free to use terms as it pleases provided it is internally consistent in its usage. We suggest, however, that this creates a grey area in REC’s account. It is grey in two respects. First, REC has chosen a terminology which, in its original home (semiotic theory) would normally allow and imply a causal role for non-contentful representation (Short, 2007: 289–316, especially p. 301).
A. Robinson, C. Southgate

2007, 289–316). Second, and more important for our purposes here, by proposing (and naming) a new approach to intentional targeting (viz. 'teleosemiotics') REC implies that the 'aiming' aspect of Brentano’s intentionality has some specific conceptual basis even once the 'content' aspect of intentionality is removed. REC is clear that this residual conceptual base is not representational (either in a contentful or non-contentful sense) and is confident that its formulation of teleosemiotics nevertheless adds something to, or clarifies, the notion of 'intentional targeting'. (Note that phrase ‘conceptual basis’ as used here does not refer to any capacity for concept formation on the part of the intentional agent, but to concepts that may need to be used by a philosopher seeking to understand the nature of intentionality.)

We suggest that what REC’s teleosemiotics adds by way of explanation of the notion of intentional targeting is less clear than REC supposes. It is beyond the scope of the present paper to survey and evaluate REC’s arguments against any role for contentful representations in basic cognition. For the purposes of this paper we shall accept that REC’s anti-representational arguments in that regard are well-founded. Our interest is in the conceptual residuum that survives REC’s deflationary manoeuvre, to which REC’s terminological decision to use the language of signs and teleosemiotics applies. If this terminology identifies a distinctive kind of causal-explanatory category, but not a representational one, then what kind of explanation is involved? If the answer to this is “nothing distinctive” then why the need for a special term (teleosemiotics) to label it; indeed, why retain the language of ‘intentionality’ at all? If, on the other hand, intentionality does involve a specific kind of explanation (but not one involving representation), what exactly distinguishes this kind of explanation from common-or-garden mechanistic causation? These questions delineate what we call REC’s ‘grey area’.

Our proposal in this paper is that the distinctive causal-explanatory concept needed to clarify REC’s grey area, and to justify the continuing use of the language of ‘intentionality’ without representational content, is that of ‘constraint-evading surrogacy’. In the sections that follow we suggest that this concept has the necessary ‘job description’ for understanding the notion of intentional targeting in the context of responsiveness to ‘natural signs’ and the internal states correlated with such signs.

We leave open the question of whether constraint-evading surrogacy can helpfully be understood as a form of non-contentful representation. REC could rectify the problem of its grey area by acknowledging the role of constraint-evading surrogacy in intentional targeting, with or without regarding constraint-evading surrogates as (non-contentful) representations. Either option would strengthen REC’s defences against the re-encroachment of contentful representation into basic cognition.

2 The Watt governor, representation, and constraint-evading surrogacy

In this section we introduce the concept of ‘constraint-evading surrogacy’ via the example of the role of the Watt governor in regulating the operation of a steam engine.

The Watt governor entered the debate about the nature of cognition in van Gelder’s paper, What is Cognition if not Computation? (van Gelder, 1995). In this paper Van Gelder aims to show that there is a viable alternative to the ‘computational’ view of cognition (computationalism). He offers two strategies by which the speed of a steam
engine could, in principle, be regulated under varying conditions of power generation (by the engine) and load (from the task being performed). The first is a ‘computational’ strategy. This would involve measurement of the engine’s speed, calculation of the actual versus the desired speed, calculation of the adjustment needed to the throttle valve to adjust the speed, and application of the necessary adjustment to the valve. This algorithmic approach would require a number of individual elements that can produce, manipulate, and use symbols. In that sense the computational solution would be analogous to the representationalism of standard cognitive science (cf. Fodor, 1975).

Van Gelder contrasts this computational solution with that actually devised by James Watt, whose ‘governor’ involves a spindle geared to the engine’s main flywheel. Two balls are attached to the spindle by levers. As the speed of the flywheel (and therefore the spindle) increases, centrifugal force causes the balls to move upwards and outwards, altering the angle of the levers. The angle of the levers is used to control the throttle valve, so that as the speed of the engine increases the valve is progressively closed (and vice versa), thus producing fine control of the flywheel speed under varying conditions.

Van Gelder points out that (as Maxwell originally showed) the behaviour of the engine and governor can be described using linked differential equations that specify, on the one hand, the angle of the levers in relation to the speed and acceleration of the flywheel and, on the other, the effect of the angle of the levers on the state of the throttle valve. In other words, the behaviour of the system can be fully described in non-computational, dynamical terms (van Gelder, 1995: 358). Van Gelder therefore offers the Watt governor as a model of a kind of system that might appear to be operating computationally but is actually operating through (non-representational) dynamic processes.

Van Gelder’s argument, namely that a dynamic systems approach is a viable alternative to computational views of cognition, anticipated the REC project. However, the example of the Watt governor remains contentious (Chemero, 2011: 68–73). Hutto and Myin remark that although all parties to the debate agree that the behaviour of the Watt governor can be described by means of coupled differential equations, there are differing opinions on whether the governor serves a representational function in any sense (Hutto & Myin, 2013: 59–62). Gallagher (2017: 98) argues that the debate about the Watt governor shows that if one is committed to representationalism one can describe any process in terms of representations. Ultimately, he suggests, representational concepts do no useful work in this context (or in the explanation of intentional actions in general) and any heuristic short-hand value that they may have is outweighed by the time that would be required “to distinguish any legitimate sense of representation from amongst the plethora of uses of that term” (Gallagher, 2017: 106).

The debate about the Watt governor illustrates how easily discussions of the legitimacy of invoking representational concepts (contentful or non-contentful) can slide into mere semantic or definitional disagreements. The approach we offer here aims to avoid such dead ends by clarifying an aspect of the character of the dynamic systems understanding of the Watt governor using the concept of ‘constraint-evading surrogacy’. We will argue that this concept has an explanatory role that adds to the dynamic systems account. We leave open the question of whether it is helpful to regard constraint-evading surrogacy as a (non-contentful) form of representation.
Recall that the role of the Watt governor is to regulate the speed of the engine’s flywheel. And note that the flywheel is itself an engineering solution to the problem of how to smooth out fluctuations in the power-output of an engine. There may be various reasons for needing to regulate the speed of the flywheel. In an extreme case the flywheel will destroy itself and the engine to which it is connected if it rotates too rapidly. To achieve the necessary regulation in a way that works from an engineering point of view, two conditions must be met. First, the regulatory mechanism (governor) must not use a significant amount of the overall power output of the engine, otherwise it would render the engine inefficient in performing its task. Second, of all the aspects of the operation of the engine that could be described (its temperature, the pressure of steam, the flow-rate of the steam, the power the engine is outputting, etc), only one of these parameters is directly relevant to the problem of regulation, namely the rotational speed of the flywheel.

We suggest, then, that although the behaviour of the system can be described by a set of differential equations, as van Gelder argued, these equations do not in themselves capture why the Watt governor is a good solution to the problem of regulating the speed of the flywheel. This involves taking what might be called an ‘engineering perspective’ on the role of the Watt governor, as outlined in the previous paragraph. (For the idea of taking an ‘engineering perspective’ in a different context see Bergstrom and Rosvall (2011)). The rotation of the spindle (or the angle of the lever-arms) involves a partial conversion of the flywheel’s energy into another physical form. In doing so the Watt governor serves the function of ‘abstracting’ something about the flywheel and enabling that abstraction to be used to control the throttle valve. This is an abstraction not merely in the (potentially trivial) sense that by virtue of the physical conversion the spindle’s rotational speed is causally correlated with the flywheel’s rotational speed, but in the non-trivial sense that the abstraction effected by the physical conversion is designed to serve a specific function, namely to dissociate one feature of the steam engine’s performance (the rotational speed of the flywheel) from all its other characteristics (especially the kinetic energy of the flywheel).

The way in which the Watt governor can be described in terms of its abstractive function, and the fact that such a description is complementary to, and does not compete with, the dynamical description, can be summarised by the term ‘constraint-evading surrogacy’. The rotational speed of the spindle of the governor is a ‘surrogate’ for the rotational speed of the flywheel. It is a ‘constraint-evading’ surrogate because it is designed to circumvent the problems that would arise if the regulatory mechanism operated at the same level of kinetic energy as the flywheel itself.

The remainder of this paper will further explore the explanatory potential of the concept of constraint-evading surrogacy. In the following section we will do so by relating the concept to Peter Godfrey-Smith’s ‘environmental complexity thesis’.

### 3 Constraint-evading surrogacy and the environmental complexity thesis

The relationship between constraint-evading surrogacy and evolved responsiveness to external natural signs can be introduced by considering the ‘environmental complexity thesis’ (ECT) proposed by Peter Godfrey-Smith (1996). Godfrey-Smith hypothesised
that, in the most general terms, the function of cognition is to cope with environmental complexity. By environmental complexity he means heterogeneity of the environment of a kind that is relevant to the organism (Godfrey-Smith, 1996: 24). Godfrey-Smith argues that the conditions under which cognition may be advantageous are where, to paraphrase his argument, (i) the environmental state is unpredictable and not directly observable; (ii) there is a feature of the environment that is reliably (but not invariably) associated with the state of the environment; (iii) the organism can ‘observe’ (detect) the condition of this ‘proximal’ feature, even though it cannot directly observe the ‘distal’ state of the environment; and (iv) the organism can make a phenotypic response to the observed proximal feature which increases the probability of it being in the appropriate phenotypic state with respect to the current state of the environment (Godfrey-Smith, 1996: pp. 118–123).

Godfrey-Smith went on to use mathematical analyses to explore quantitatively the relationships between environmental probabilities, phenotypic fitness differences in varying environments, the reliability of the relationship between ‘proximal’ and ‘distal’ environmental features, and the ‘costs’ to the organism of having such phenotypic flexibility (Godfrey-Smith, 1996: 207–290). Similar quantitative approaches to the benefits of responsiveness to environmental cues have been developed by others (Bergstrom & Lachmann, 2004; Donaldson-Matasci et al., 2010; Lean, 2014; McNamara & Dall, 2010; Rivoire & Leibler, 2011). A feature of these latter investigations (unlike Godfrey-Smith’s) is the connection made with the mathematical theory of information. These subsequent approaches have generally been undertaken in the context of ecological and evolutionary concerns, rather than as an approach to basic cognition.

Following Godfrey-Smith’s initial presentation of his thesis (Godfrey-Smith, 1996) subsequent discussions of his hypothesis about the relationship between environmental complexity and cognition became somewhat mired in debates about wider philosophical issues, such as the distinction between ‘externalist’ and ‘internalist’ views of the relationship between organism and environment (Godfrey-Smith, 1996: 30–65, 1997: 584–586). As a result, it appears that the relevance to cognition of phenotypic responsiveness to environmental cues, where the latter are understood as detectable proximal indicators of undetectable distal environmental states, may have somewhat dropped out of view in critical evaluations of the environmental complexity thesis, including in Godfrey-Smith’s own thinking. For example, this aspect does not merit a mention in his later summaries of the thesis (e.g., Godfrey-Smith, 2002: 23, 2017: 467–469).

One result of the downplaying of the role of the accessibility of certain environmental features but not others was that the environmental complexity thesis became vulnerable to what may be called the ‘triviality objection’. Karen Neander articulated this objection in an early review of Godfrey-Smith’s 1996 book (Neander, 1997). Neander suggested that the environmental complexity thesis potentially encompasses such a broad range of ways in which cognition could be adaptive (such as enabling food-foraging, mate selection, evading predators, etc.) that it would be difficult to falsify it. “If all the world were food we wouldn’t need to tell food from non-food; if all potential mates were the same, there would be nothing to choose between them; and if predators were all entirely predictable, one wouldn’t need to be so alert” (Neander, 1997: 574). In that sense, Neander argues, the thesis that cognition is a way of connecting environmental complexity with behavioural complexity is so uncontroversial as to be trivial.
We suggest that Neander’s ‘triviality’ objection (which Godfrey-Smith appears to be inclined to take seriously) can be countered by recognising what amounts to the role of constraint-evading surrogacy in the original formulation of the environmental complexity thesis. As noted above, it is intriguing how Godfrey-Smith’s earlier emphasis on the use of regularities (the relation between a cue and the state of the environment) to cope with irregularities (the unpredictable nature of the environment), and the link between this and the distinction between unobservable ‘distal’ features of the environment and observable ‘proximal’ features, tended to drop out of later discussions. By ceasing to present the environmental complexity thesis primarily in such terms, Godfrey-Smith thus underplayed a key link in the chain of argument that was elsewhere implied (but increasingly unstated) in his analysis. An enduring element of his thesis ought to have been an acknowledgement that not only are many aspects of the state of an organism’s environment unpredictable, they are also not directly detectable. More precisely, there are constraints imposed by the structure and organisation of the organism that would make the acquisition of a capacity to directly detect the state of the environment evolutionarily costly and perhaps, in effect, impossible. The complexity (unpredictability) of environments would not present a problem if the relevant environmental states could be easily determined by the organism. The use of environmental cues can be explained as an adaptive strategy for dealing with the unpredictability and undetectability (at least directly) of certain environmental states of importance to the organism. The cue is being used as a constraint-evading surrogate, enabling the organism to place itself in the appropriate phenotypic state for the current state of the environment.

Once that crucial but increasingly overlooked step in the environmental complexity thesis is made explicit, the thesis is no longer vulnerable to the triviality objection. It may be trivial to state that one purpose of behavioural complexity is to cope with environmental complexity, and that cognition is what enables organisms to make that connection. It is not trivial to hypothesise that many or all the specific examples of cognitive mediation between environmental and behavioural complexity (be it foraging for food, selecting a mate, evading a predator, etc.) have in common the strategy of using one or more constraint-evading surrogates (different kinds of surrogate for dealing with different kinds of constraint).

4 Constraint-evading surrogacy and REC: Adopted and adapted surrogates

We have seen that the environmental complexity thesis implicitly recognises what amounts to the role of surrogacy in basic cognition and, again implicitly, the role of the surrogate in evading a constraint, the constraint being the unknowability of the state of the environment. REC, in a complementary fashion, recognises that responsiveness to external cues (natural signs) is the basis of the most primitive kinds of intentionality (see Section 1). What REC misses is the fact that the external cue, the sign, is useful, and is the basis of ur-intentionality, by virtue of acting as a constraint-evading surrogate.
To illustrate this it is helpful to consider Hutto’s (2008) seminal formulation of ideas that were subsequently incorporated into the REC programme. Hutto’s original test case for the concept of teleosemiotic ur-intentionality was the phase-locked character of FM (frequency modulated) neurons in the auditory cortex of bats, the firing of which is triggered by a specific auditory stimulus, namely the sound of the beating wings of a particular insect (Hutto, 2008: 52). In keeping with the ‘terminological decision’ we identified in Section 1, Hutto is happy to call the relationship between the sound of the wing-beats and the presence of this particular insect a natural ‘sign’ (Hutto, 2008: 53). The sound is associated reliably enough, but not necessarily perfectly, with the presence of the insect (there might be other, inedible, insects which come to occupy the bats’ territory and which make the same sound in flight as the preferred prey). Conversely, there may be other natural signs of this species of insect (such as pheromones or other chemical traces), but the bat is adapted to detect the sound of the wing-beats.

Putting this in terms of the environmental complexity thesis, the environment is complex in that the prey species is sometimes present and sometimes not. Applying Godfrey-Smith’s approach, the bat is making use of a regularity (the connection between a specific kind of wing-beat sound and the presence of the prey species) in order to cope with an irregularity (the uncertainty about the presence or absence of the prey species). As we have seen, the non-triviality of the environmental complexity thesis rests on recognising that the advantage of responsiveness to an environmental cue is not simply that an irregularity in the distal environment is compensated for by a regularity in the proximal environment, but that responsiveness to the proximal feature renders accessible to the bat something otherwise inaccessible about the environment. In other words, responsiveness to the wing-beat sounds is a constraint-evading strategy. The wing-beat sound is (or, rather, can be used as) a surrogate for the presence of the insect; responsiveness to this surrogate evades the constraints imposed by the fact that the bat needs to detect the presence of the insect while still at some distance from it.

In Hutto’s account the emphasis is on the lack of ‘information’ or ‘content’ in either the wing-beat sounds or in the bat’s way of registering the sounds. As Hutto puts it,

in responding to telltale stimuli [e.g., the sound of wing-beats] the perceptual subsystems of organisms neither extract or decode it [information] for further use. When conditions are normal, creatures are hooked up to the world in the right sorts of ways and their actions succeed (mostly). This only requires that the appropriate correspondences do in fact hold: it does not require the manipulation of internal contents that say that they hold. There need be no cognitive registering of the fact that insects are present (Hutto, 2008: 53).

In terms of the bat example, the ‘correspondence’ that ‘does in fact hold’ is that between wing beat sounds and the presence of the insect. The FM neuron performs the function of enabling the bat to be ‘hooked up to the world’, but not in a way that involves the manipulation of internal ‘content’ that ‘says’ anything about the correspondence. In keeping with REC’s project of banishing representational content from accounts of basic cognition, Hutto introduces the term ‘local indexical guide’ (LIG) as a way of describing the role of FM neuron activity while avoiding any
A. Robinson, C. Southgate

notion of ‘cognitive registering’. Such internal neural states “are not representations – not even of a very basic variety.” Instead, the role of LIGs is to,

coordinate actions with respect to specific worldly offerings. They are like Sherpas in that they have the right local connections to safely direct others through specific domains. Yet, unlike their human counterparts, LIGs have no mediating practical knowledge of their own” (Hutto, 2008: 54).

Whereas Hutto seems comfortable about applying the term ‘sign’ to external worldly offerings, he avoids doing so when describing the function of FM neurons (to which the term Local Indexical Guide applies). This is understandable because even on full-blown representational accounts of cognition ‘natural signs’ such as wing beat sounds would not normally be regarded as having representational content. In contrast, internal neural states often are held to involve representational content. Ruth Millikan’s account of neural states serving as ‘pushmi pulyu’ representations is an example, and Hutto’s concept of LIGs is explicitly developed as an alternative to Millikan’s representationalism (Hutto, 2008, 54).

A consequence of REC’s different terminological treatment of external natural signs and internal LIGs (which are not referred to as ‘signs’) is that a formal similarity between their functional roles is, in our view, obscured. On our account the notion of constraint-evading surrogacy applies equally well (though in different ways) to the neuronal activity and to the sound that stimulates it, the external natural sign. The sound is a pattern of fluctuating air pressure. The behavioural response of the bat to the wing beat sound is mediated by electrical activity in peripheral nerves and the central nervous system of the bat, ultimately resulting in the muscular activity by means of which the bat flies towards the insect. Fluctuations in air pressure cannot themselves be used directly to trigger this neuromuscular activity, and this disparity between the nature of the stimulus and the nature of the required output constitutes a constraint. The constraint is circumvented by converting the fluctuations of pressure that move the bat’s tympanic membrane into electrical activity in (among others) the FM neurons. In the broadest terms, the activity of FM neurons is therefore performing the function of acting as a constraint-evading surrogate.

Other examples could be cited to illustrate the constraint-avoiding role of using activity in a particular neuronal class as a surrogate for a non-neuronal stimulus, or as a surrogate for some collective property of a group of neurons providing inputs to the neuron in question. Think for example of the neurons in the anuran (frog and toad) visual system which are preferentially active in response to a shape moving along its longitudinal axis (Neander, 2017: 105–115). The neuronal activity serves to make available to the rest of the nervous system a surrogate for the actual physical movement made by a variety of food-species (such as centipedes). Or consider simple cell activity in the mammalian visual cortex, which is correlated with the presentation to the retina of lines of a certain spatial orientation (Hubel & Wiesel, 1968). The inputs to simple cells are from parallel rows of neurons in the lateral geniculate nucleus (LGN). Simple cell activity is a surrogate for the simultaneous firing of rows of LGN neurons. By collating information from those neurons, simple cells solve the problem of making the impingement of certain spatial patterns of photons on the retina available for processing in the neuronal systems responsible for visual perception.
The examples could be multiplied. Crucially, nothing in this view of Local Indexical Guides as constraint-evading surrogates requires any notion of ‘internal content’ or ‘cognitive registering’. The important point is that if the role of external natural signs can be seen to be that of acting as constraint-evading surrogates, there is no reason in principle to deny the same status to Hutto’s Local Indexical Guides. That the wing beat sounds are external to the bat and the FM neurons are internal is no reason to deny that both (in different ways) can perform the function of constraint-evading surrogacy. The most important distinction is not between internal and external but between an element of the cognitive system that is adapted to perform as a constraint-evading surrogate and a ‘worldly offering’ (REC’s term) that has been adopted to act as a constraint-evading surrogate.

A further example employed by Hutto to illustrate the concept of a Local Indexical Guide, the waggle dance of the honey bee, happens to illustrate the point. Waggle dances serve as LIGs because although they “have the proper function of generating characteristic responses in cooperating consumer mechanisms – the watching bee or bees,” they do not, “say how things stand in order to guide action, they guide it directly” (Hutto, 2008: 54–55, emphasis in original). As Hutto points out, the features of the waggle dance correlate only with the distance and direction (relative to the sun) of the nectar-source from the hive, so the dance only works as a guide to the location of the nectar at the particular time of day when it is performed. We agree with Hutto that this is one of the reasons for holding that the waggle dance is a correlation-based guide rather than a contentful propositional statement. We would add, however, that the reason that the dance has evolved is that the location of the nectar-source is not directly perceivable by the bees in the hive, a constraint on the efficient gathering of nectar which the dance serves to overcome, albeit in a way that is useful only at the time of the dance.

That the dance enables the bees to respond to something they cannot directly perceive may seem so obvious as to be trivial. Nevertheless, it is our contention that such use of constraint-evading surrogates is a key feature of intentionality. It is what makes intentionality distinct from mere mechanistic causation and, moreover, it does not require any reference to representational content. As we have seen, Hutto ascribes the same broad functional role – that of local indexical guide – to the waggle dance, which is external to the observing bee’s nervous system, and to the FM neuron, which is internal to the bat’s nervous system. Similarly, we ascribe the role of constraint-evading surrogacy to both of these: constraint-evading surrogates can be internal or external to the agents who use them. Where we differ from Hutto is that we hold that external natural signs play the same broad functional role as the waggle dance and the FM neuron, in being surrogates selected to circumvent particular kinds of constraint. The difference is that natural signs (such as wing beat sounds) already exist in nature and have been adopted by the agent (bats) to serve the surrogacy function, whereas the bees’ waggle dance and the bats’ FM neurons have evolved to perform that function.

It might be objected that our argument that external natural signs, evolved external signals, and internal neural states hooked up to these signs and signals are all examples of constraint-evading surrogacy gives rise to an embarrassing and explanatorily self-defeating proliferation of constraint-evading surrogates. We acknowledge
that on our account various kinds of cognition of different degrees of sophistication will likely involve many different kinds of constraint-evading surrogate. The specific role of each such constraint-evading surrogate would merit analysis in its own right. Our focus here, however, is on a more general hypothesis, namely, that constraint-evading surrogacy is an indispensable component of a naturalistic account of intentionality, one currently missing (or at best only implicit) in REC’s account.

REC’s programme revolves around the claim “that there can be intentionally directed cognition and, even, perceptual experience without content” (Hutto & Myin, 2013, x). Jean-Michel Roy argues that the most important feature of Brentano’s notion of intentional mental states is that of ‘objectivation’, by which he means that such mental states, “make a world of objects, as opposed to a world of things, emerge and, consequently, make the subject/object opposition possible” (Roy, 2015, 96). He criticises REC for being silent on how, precisely, its account preserves this aspect of intentionality and argues that, as a result, REC’s position sometimes appears to amount to nothing more than affirming that organisms are, “responsive to certain elements of their environment” (Roy, 2015, 123). This is a reflection of what we identify as REC’s “grey area” (see Section 1). Hutto and Myin’s response that past evolutionary selection of the relevant responses has a normative element that accounts for objectivation (2017, 116) is unpersuasive on its own because not all selected traits have the property of intentional directedness. Our contention is that what is needed to preserve an account of basic intentionality as distinct from mere mechanistic causality is the concept of constraint-evading surrogacy. This claim is not undermined by the fact that various elaborations on the basic strategy of constraint-evading surrogacy will be involved in the many different circumstances and degrees of sophistication in which cognition operates.

5 Constraint-evading surrogacy and information

It is worth considering briefly the relationship between the concept of constraint-evading surrogacy and the mathematical theory of information. As noted in Section 3, Godfrey-Smith’s mathematical analysis of the environmental complexity thesis potentially lends itself to an information-theoretic approach (see references in Section 3). When the fitness value of environmental cues is investigated in informational terms some intriguing results emerge. For example, it can be shown that the maximum fitness value of an environmental cue is equal to the entropy of the environment (i.e., the degree of uncertainty about the state of the environment; Bergstrom & Lachmann, 2004). It is important to recognise, however, that information theory simply provides a measure of reduction in uncertainty. If, in the example discussed in Section 4, the bat is ‘uncertain’ whether or not an insect is nearby, this uncertainty can be expressed mathematically. And if the presence of wing-beat sounds reduces the uncertainty, that reduction in uncertainty can likewise be expressed in informational terms.

It is often (and rightly) said that information theory is not primarily concerned with the semantic aspects of communication (Smith, 2000, 181). An informational analysis of the relationship between wing-beat sounds (or odours, or any other natural sign) and the presence of the prey species can tell us, for example, how much
the predator’s uncertainty about the environment may be reduced if it has evolved a responsiveness to the environmental cue in question. Putting this in general terms, such an analysis can tell us by how much our uncertainty about X is reduced if we know the state of Y. Quantitatively we may obtain the same answer if X is the presence of a particular insect and Y is the presence of a particular pattern of air-pressure fluctuation, or if X is the presence of a line of a particular orientation and Y is firing of a simple cell in the mammalian visual cortex. In each of these cases Y may ‘transmit’ the same amount of information about X, but Y ‘means’ something quite different (presence of a prey insect or presence of a line of particular orientation).

None of this is new in debates about information theory and semantics. However, recognition of the role of constraint-evading surrogacy puts a further level of distance between mathematical information-measures and semantic concepts. In the above formulation, Y is a surrogate for X. In some circumstances, as we have seen, we can quantify, in terms of the theory of information, by how much knowledge of Y reduces the uncertainty about X. We have argued in this paper that, when used in a particular way by an agent, the surrogate Y can be understood as performing the function of circumventing a constraint. Let us say that Y is a surrogate for X, and that using Y as a surrogate for X evades a certain constraint, Z. In the case of the use of wing-beat sounds (Y) as a surrogate for the presence of a particular kind of insect (X), the constraint circumvented (Z) is the difficulty the bat has in locating food at a distance. In the case of the use of simple cell activity (Y) as a surrogate for the presence of a line of particular orientation in the optic field (X), the constraint circumvented (Z) is the need to present to the next stage of the visual processing apparatus a single input corresponding to the global activity of a set of topographically arranged cells in the lateral geniculate nucleus.

The gap between an information-theoretic measure of certain correlations, or reductions of uncertainty, and the ‘semantic’ or ‘semiotic’ dimension of those correlations is therefore even wider than the assertion that Shannon information is not concerned with semantics recognises. Shannon information is a measure of one aspect of the relation between X and Y, namely, the extent to which knowledge of Y reduces the uncertainty about X. Shannon information does not tell us anything about what sort of thing X is, or why Y has the potential to function as a surrogate for X. This is a question about how Y could stand in for X. Notoriously, at one level anything could stand for anything else. Nevertheless it is the case that surrogates can stand in for things in various ways, and it is the absence of this ‘how’ element from the mathematical analysis that constitutes what we might call the ‘semantic gap’ in Shannon information. The further gap, the one that is less well recognised, is the ‘constraint gap’. This is the gap between, on the one hand, the informational account of the relationship between X and Y, and, on the other, an account of why Y is a suitable surrogate for X. Suitable, that is, in the sense that the use of Y as a surrogate enables a certain constraint, Z, to be circumvented. This why question, which constitutes what one could call the ‘constraint gap’, is distinct from the how question that constitutes the ‘semantic gap’.

Analyses based on the mathematical theory of information have a powerful capacity to uncover correlations and possible functional connections between different structural
or dynamic elements of complex systems, including biological systems (e.g., Kim et al., 2015; Walker et al., 2016). All such analyses deal, in various ways and degrees of sophistication, with the reduction in uncertainty about one variable, or set of variables, \( X \), that is afforded by knowledge of another variable, or set of variables, \( Y \). However, informational analyses cannot, in themselves, provide an account of how a particular variable, \( Y \), comes to act as a surrogate for \( X \) (the semantic gap). Neither can an informational analysis give an account of why \( Y \) is able to circumvent a particular constraint, \( Z \) (the constraint gap). Both of these dimensions of explanation (the surrogacy relation of \( Y \) to \( X \), and the constraint-evading relation of \( Y \) to \( Z \)) are subject to contingencies that lie beyond the direct reach of mathematical description.

This observation should give pause to those who hold that understanding what life is (e.g., Nurse, 2008; Davies & Walker, 2016) or what consciousness is (e.g., Oizumi et al., 2014) can be reduced to understanding information ‘processing’ or information ‘flows’ as a unifying concept. Informational analyses of biological systems and networks may yield important insights into functional aspects of neural or genetic networks, but they will not be capable, in themselves, of fully delineating the specific functions of particular elements of those networks because at least one functional aspect, that of constraint-evading surrogacy, cannot be reduced to the mathematical theory of information.

6 Constraint-evading surrogacy and Ramsey’s job description for representations

We have argued that the additional conceptual apparatus required by REC to make sense of intentional targeting is the concept of constraint-evading surrogacy. We have shown that this concept clarifies the adaptive basis of responsiveness to external natural signs (connecting it with Godfrey-Smith’s environmental complexity thesis). We have also suggested that while such responsiveness can be analysed in information-theoretic terms the concept of constraint-evading surrogacy adds an additional element to the irreducibility of intentionality to mathematical measures of information. We have so far left open the question of whether constraint-evading surrogacy is helpfully regarded as a form of (non-contentful) representation. In this section we will suggest that there is a prima facie case for considering that it may be.

The case for constraint-evading surrogates as (mostly non-contentful) representations may be put in terms of William Ramsey’s argument that to be usefully considered a representation the thing in question must meet the requirements of a certain kind of ‘job description’ (Ramsey, 2009). Ramsey holds that representations are surrogates (Ramsey, 2009, 69–70, 87). This is uncontroversial. At a mere etymological level representations are re-presentations. In semiotics a sign is something that stands in for something else (\( \textit{aliquid stat pro aliquo} \), in the medieval formula). Ramsey adds to this the requirement that we can specify how the representation’s being a representation is necessary to the function it performs: “Principally, there needs to be some sort of account of just how the [representational] structure’s possession of intentional content is (in some way) relevant to what it does in the cognitive system” (Ramsey, 2009, 27). By ‘content’ Ramsey means what the surrogate “stands in for,”
rather than the stronger notion of content targeted by REC, namely, what a representation says, claims, or asserts (Ramsey, 2009, 87; cf. Hutto & Myin, 2017: 10). Ramsey’s inclusion of ‘content’ in his job description therefore need not distract us because the job description challenge will be met provided the surrogate’s acting as a surrogate for something else can be shown to be essential to the function it performs. In other words, for Ramsey, being employed as a surrogate for something else does not require a “notion of content that assumes the existence of some kind of specified correctness condition”, which is REC’s criterion for the kind of representation for which it rejects any role in basic cognition (Hutto & Myin, 2017: 10). Of course, REC also denies that there is any kind of representation that does not meet this criterion, but, as we argued in Section 1, that reflects REC’s ‘default conceptual lens’, which is distinct from its core programmatic project.

The requirements of Ramsey’s job description are most clearly evident in the two examples he develops of supposedly representational processes in the Classic Computational Theory of Cognition (CCTC) that, he argues, do legitimately deserve to be called representations. The context in which he presents these examples is his contention that the connection made in the standard interpretation of CCTC between computational symbols and propositional attitudes – beliefs, desires, and the like – fails to justify treating those symbols as representational entities or states (Ramsey, 2009, 39). According to Ramsey there are, nevertheless, ways in which CCTC can be regarded as involving states of the cognitive system that do pass the job description challenge. The first such legitimate ascription of representational function can be found at a global input-output level of analysis (Ramsey, 2009, 68–77). According to Ramsey, the inputs and outputs of cognitive processes are necessarily representational. For example, the cognitive process of addition requires a number (which is an abstract entity) to be encoded as a numeral (a representation of the abstract entity) before it can be cognitively manipulated. Likewise, the output of the calculation will be a numeral, which is a representation of the abstract entity that is the result of the calculation (Ramsey, 2009, 68–69). Similarly, in the cognitive task of face recognition, “The input to any cognitive system that recognises faces will not be actual faces, of course, but some sort of visual or perhaps tactile representation presented [to the cognitive system] by the sensory system” (Ramsey, 2009, 69). Beyond these examples of global input-output relations Ramsey further argues that the input-output elements of task-decompositional cognitive subsystems (construed according to the classical computational model of cognition) similarly meet the job description challenge (Ramsey, 2009, 72–74).

We agree with Ramsey’s construal of the input-output features of cognitive systems as inherently representational. (Like Ramsey, we are not thereby endorsing CCTC as the best model of cognition, but simply following his reasoning about what would legitimately count as a representation in such models.) We would simply add to Ramsey’s account the claim that the function the physical encoding of a numeral performs for a cognitive system undertaking addition or multiplication, or the function a certain sensory apparatus performs that enables a cognitive system to undertake facial recognition without processing “actual faces”, is that of constraint-evading surrogacy. Abstract entities and (often) external objects such as faces cannot themselves be manipulated, which is a constraint faced by the cognitive system. The
constraint is circumvented by providing a numeral as a surrogate for the abstract entity of number, or a neural state as a surrogate for an actual feature of the face. The surrogate circumvents the relevant constraint— and this is precisely the reason why Ramsey considers it to be a genuine representation— by being manipulable in a way that the abstract entities or external objects are not.

Ramsey’s second legitimate sense of representation, alongside global or sub-system input-output relations, involves internal cognitive states or structures that instantiate “simulation” or “structural” isomorphisms. Ramsey uses the analogy of Bob, a man who uses a pencil and notebook to sketch his family tree in diagrammatic form. This enables him to see relationships between people that he had not previously appreciated (Ramsey, 2009, 81). By analogy, a cognitive system could perform a similar task provided it had a way of encoding the people and their relationships. Significantly, Ramsey argues that the internal cognitive sub-system itself does not have to understand the meaning of the symbols, it just has to know the appropriate algorithm for manipulating them (Ramsey, 2009, 84). The ‘diagram’ (in whatever form it is instantiated) is representational because it accurately models real-world family trees (Ramsey, 2009, 85). If it did not do so the algorithmic manipulation of the symbols would be worthless. In our terms, the diagram of Bob’s family tree stands as a surrogate for his actual relatives and their relationships. Bob could have gathered all his family together and arranged them spatially in a room in a way that would have helped him understand how they are related. The impracticality of doing so constitutes a constraint that is circumvented by the diagram he draws (or by the diagram’s neurally or computationally-based equivalents).

Constraint-evading surrogacy therefore appears to specify the implicit (and quite stringent) requirements of Ramsey’s representational job description. It does so in a way that provides a coherent notion of non-contentful representation (if by ‘content’ one means something that claims, says, or asserts things according to some specified correctness condition). REC could therefore afford to accept the case for constraint-evading surrogates as representations (i.e., it could modify its ‘default conceptual lens’— see Section 1) without undermining its core commitment to banishing representational content from accounts of basic cognition.

7 Constraint-evading surrogacy and life-mind continuity

In this penultimate section we use protobiotic autocatalytic systems as a test case to further explore the role of constraint-evading surrogacy and to recap some key features of our argument for its role in basic intentionality.

Godfrey-Smith originally presented his environmental complexity thesis as what he called a ‘weak’ version of the life-mind continuity thesis. That is, he saw parallels between the characteristics that make living things alive and those that make cognising agents cognitive, without implying that being alive and being cognitive are co-extensive (Godfrey-Smith, 1996: 72–73). If, as we argued above (Section 3), the ‘triviality objection’ to the environmental complexity thesis can be addressed by reasserting the centrality of (what amounts to) constraint-evading surrogacy then it is possible to reconsider the relationship between the environmental complexity
thesis and life-mind continuity in a stronger form than that countenanced by Godfrey-Smith. As a test of the relevance of constraint-evading surrogacy to the life-mind continuity thesis we offer a thought experiment based on existing mathematical modelling of autocatalytic systems. Autocatalytic systems ought to be of interest to REC because they are an empirically tractable example of the principles of autopoiesis, in some conceptions of which REC finds a natural ally (Hutto & Myin, 2017: 75). It is therefore worth asking whether proto-biotic autocatalytic systems could, in principle, use constraint-evading surrogates as a way of adapting to environmental complexity.

The idea of chemical systems in which the synthesis of each constituent molecule is catalysed by molecules within the system was originally proposed and modelled mathematically in the 1970s (Kauffman, 1971, 1986, 1993). Since then several in vitro autocatalytic systems have been successfully constructed (Sievers & von Kiedrowski, 1994; Vaidya et al., 2012; Higgs & Lehman, 2015) and the evolvability of autocatalytic systems has been explored using mathematical modelling (Smith et al., 2014; Hordijk & Steel, 2017).

An important conclusion from mathematical modelling is that autocatalytic networks will not be evolvable (i.e., able to change by the accumulation of fitness-enhancing mutations) unless they possess certain features (Vasas et al., 2012: 2). Specifically, they will only be capable of evolution if the variants of the system are able to arise and compete with each other. Such variation turns out to be possible if the autocatalytic system as a whole is comprised of several hierarchically ordered sub-systems of autocatalysis (Vasas et al., 2012). These subsystems can emerge in different temporal orders and stabilise at different ‘attractor’ states of the total system, giving rise to variations between different ‘individual’ versions of the system. The state in which a particular system stabilises may depend on stochastic factors such as the order in which certain slow, uncatalysed reactions among the ‘food’ molecules happen to occur (Hordijk & Steel, 2017: 7; Vasas et al., 2012: 8).

Although there is no template-based replication of genetic information, these modelled systems can have heritable elements (an autocatalytic ‘core’ that can reconstitute itself from food molecules if just one of the catalysts comprising the core autocatalytic set is present) and non-heritable elements (peripheral reactions which depend on the activity of the core and which cannot reconstitute themselves in the absence of the core). This distinction between core and peripheral elements is analogous to the genotype-phenotype distinction in more evolved biotic systems (Vasas et al., 2012: 4). Importantly, modelling has been extended to include investigation of the population dynamics of spatially separated protocells, where each protocell contains a version (variant) of the autocatalytic system. The formation of such protocells can itself be modelled by including, among the subsystems of the network, the autocatalytic production of self-assembling constituents of a semipermeable membrane (Hordijk & Steel, 2017: 8).

In the context of this developing body of work on the evolvability of autocatalytic systems, we offer the following thought experiment. Imagine a population of membrane-enclosed autocatalytic networks undergoing evolution by natural selection based on non-genetic (i.e., non-nucleic acid based) variation. Suppose that there are two environmental conditions, ‘benign’ and ‘hostile’, depending on certain physical
or chemical parameters. Suppose further that these different environmental conditions would each favour a different reproductive strategy. In the benign environment it would be advantageous for an individual protocell to grow (by synthesis of catalysts and progressive extension of the autocatalytically-synthesised membrane) until the protocell had reached a certain maximum size, at which point, triggered by physical factors, it would divide into two or more daughter cells. In the hostile environment, in contrast, it would be advantageous to divide early and frequently, division taking place while the protocell was relatively small. This difference in reproductive strategy would correspond to the well known distinction in ecology between K-selection and r-selection (Pianka, 1970). Suppose further that two variant phenotypes are available, based on two different attractor states of the autocatalytic network. One of these states switches on an autocatalytic sub-set which produces a molecule that becomes incorporated into the membrane, reducing the membrane’s physical stability. This results in a tendency for protocell division to occur while the cell is still relatively small.

Suppose that a ‘problem’ (constraint) faced by the protocell is that it cannot directly detect the state of the environment, benign or hostile, in a way that would enable it to switch appropriately between the two phenotypic strategies (K and r). In other words, the hostile environment would not, simply by virtue of whatever makes it hostile, cause the protocells to switch from the K to the r strategy. However, protocells could become capable of such switching if they acquired a responsiveness to a constraint-evading surrogate. (In this regard it is relevant that existing mathematical models have included modelling the effect of introducing selective inducers of different autocatalytic subsets – Hordijk et al., 2018).

Suppose, say, the hostile environment was associated with a higher concentration of a certain chemical ‘signal’. If the protocell membrane was permeable to this chemical cue, and if the cue triggered a change in the relative activities of different sub-systems of the autocatalytic network, then the autocatalytic network as a whole could be switched from the K-selection attractor to the r-selection attractor. This would occur if the change in the attractor state of the network (the relative activities of the relevant subsystems) in response to the cue resulted in production of the molecule that destabilises the membrane and thereby caused protocell division to occur at smaller protocell sizes. The cue would then be acting as an external natural sign (an ‘adopted’ constraint-evading surrogate – see Section 4) which the autocatalytic system uses to switch itself into the most advantageous internal state for the prevailing (but otherwise undetectable) physico-chemical state of the external environment. The system would ‘adopt’ the cue to act as a surrogate for the state of the environment, avoiding the need to evolve a way of detecting the change from the benign to the hostile environment directly.

Let us review how the thought experiment about an environmentally-responsive protobiotic system reflects our argument about the role of constraint-evading surrogacy in a ‘radically embodied’ (REC) approach to basic cognition. The system’s capacity to change its own state between ‘K’ and ‘r’ reproductive strategies is a primitive form of intentionality: the internal change of state from one attractor state of the system another is ‘about’, is ‘aimed at’ the state of the environment, benign or hostile. This intentional state of the system does not have
‘content’ in a propositional, assertive, sense. Rather, the ‘aiming’ aspect of the system’s intentionality is a function of the system’s use of an external environmental cue. This cue, a ‘natural sign’, stands in as a surrogate for the state of the environment. The adoption of the environmental cue as a constraint-evading surrogate is the basis of the autocatalytic system becoming an intentional system. The cue is a \textit{potential} stand-in because it is correlated (not necessarily perfectly) with the state of the environment. It is a \textit{suitable} stand-in because it enables the system to circumvent the constraint that it (the system) is unable directly to detect the state of the environment. Neither the cue or any aspect of the internal state of the system have any representational ‘content’. On a particular understanding of the concept of representation – one that allows for the notion of non-contentful representation – the role of the cue as a constraint-evading surrogate might reasonably be regarded as representational, but this is not essential to the argument for constraint-evading surrogacy as necessary for a non-content-dependent account of basic intentionality.

\section{Conclusion}

We have suggested that, while REC successfully shows how basic cognition does not depend on contentful representation, its formulation of intentional targeting as ‘teleosemiotics’ amounts to a ‘grey area’ in its account. This is a grey area a) because REC denies the utility of non-contentful representational concepts whereas the field of semiotic theory, from which REC has borrowed its teleosemiotic terminology, would regard ‘natural signs’ as non-contentful representations and, more importantly, b) because having dismissed representation (whether contentful or non-contentful) as the basis of intentional targeting, it is not clear whether REC’s teleosemiotic account of intentional targeting involves a distinctive kind of causal-explanatory framework in addition to a merely mechanistic account, and if so what is the basis of this distinctiveness.

We have argued that the concept of constraint-evading surrogacy provides what REC’s grey area currently lacks, namely an explanation of what justifies treating intentional targeting as something that involves more than common-or-garden mechanistic causation but without implying or invoking any notion of contentful representation. We have explored this via a reconsideration of Godfrey-Smith’s environmental complexity thesis and have considered how the resulting view of the role of environmental cues relates to information theory. In doing so we have introduced a distinction between the ‘semantic gap’ and the ‘constraint gap’, both of which contribute to the irreducibility of intentionality to the mathematical theory of information. We have also shown how the concept of constraint-evading surrogacy can generate testable hypotheses about the evolution of protobiotic autocatalytic systems, thus potentially extending the applicability of REC’s insights to the question of life-mind continuity and the origin of life.

Since REC’s primary aim is to show that basic cognition does not depend on the employment of contentful representations, our attempt to clarify REC’s ‘teleosemiotic’ account of intentional targeting is bound to raise the question of the
relationship between the concept of constraint-evading surrogacy and representation. We emphasise that the argument we have presented does not depend on equating constraint-evading surrogacy with representation. Nevertheless, we suggest that there is a prima facie case for supposing that constraint-evading surrogacy might usefully be regarded as being consonant with Ramsey’s (2009) job description for representational function in a way that does not undermine REC’s rejection of any role for contentful representation in basic cognition. REC could rectify the problem of its grey area by acknowledging the role of constraint-evading surrogacy, with or without accepting (or wishing to investigate further) the prima facie case for regarding constraint-evading surrogacy as a specification of a job description for (non-contentful) representation.

Acknowledgements This research was generously funded by a grant from the John Templeton Foundation. However, the views presented here do not represent the views of the Foundation. We also thank Niles Lehman, Wim Hordijk and two anonymous reviewers for their helpful comments on drafts of the paper. This study did not generate any new data.

Funding This work was funded by JTF grant 60882, Cooperation and Interpretation in the Emergence of Life.

Declarations

Conflicts of interest/competing interests The authors have no conflicts of interest or competing interests to declare.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Bergstrom and Rosvall. (2011). The transmission sense of information. Biology and Philosophy, 26, 159–176.
Bergstrom, C.T., & Lachmann, M. (2004). ‘Shannon information and biological fitness’. Information Theory Workshop, pp. 50–54. https://doi.org/10.1109/ITW.2004.1405273.
Brentano, F. (1973 [1874]). Psychology from an empirical standpoint. Routledge and Kegan Paul.
Chemero, A. (2011). Radical embodied cognitive science. MIT Press.
Davies, P. C. W., & Walker, S. I. (2016). The hidden simplicity of biology. Reports on Progress in Physics, 79, 102601. https://doi.org/10.1088/0034-4885/79/10/102601
Donaldson-Matasci, M., Bergstrom, C. T., & Lachmann, M. (2010). The fitness value of information. Oikos, 119, 219–230.
Fodor, J. A. (1975). The language of thought. Thomas Y. Crowell.
Gallagher, S. (2017). Enactivist interventions: Rethinking the mind. Oxford University Press.
Gibson, J. J. (1979). The ecological approach to visual perception. Houghton Mifflin.
Godfrey-Smith, P. (1996). *Complexity and the function of mind in nature*. Cambridge University Press.
Godfrey-Smith, P. (1997). Replies to sober, sternly and Neander. *Biology and Philosophy*, 12, 581–590.
Godfrey-Smith, P. (2002). Environmental complexity and the evolution of cognition. In R. J. Sternberg & J. C. Kaufman (Eds.), *The evolution of intelligence*. Lawrence Erlbaum Associates.
Godfrey-Smith, P. (2017). Complexity revisited. *Biology and Philosophy*, 32, 467–479.
Higgs, P. G., & Lehman, N. (2015). The RNA world: Molecular cooperation at the origins of life. *Nature Reviews Genetics*, 16, 7–17.
Hordijk, W., & Steel, M. (2017). Chasing the tail: The emergence of autocatalytic networks. *BioSystems*, 152, 1–10.
Hordijk, W., Naylor, J., Krasnogor, N., & Fellermann, H. (2018). Population dynamics of autocatalytic sets in a compartmentalised spatial world. *Life*, 2018(8), 33. https://doi.org/10.3390/life8030033
Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, 195, 215–243.
Hutto, D. D. (2008). *Folk psychological narratives: The sociocultural basis of understanding reasons*. MIT Press.
Hutto, D. D., & Myin, E. (2013). *Radicalizing Enactivism: Basic minds without content*. MIT Press.
Hutto, D. D., & Myin, E. (2017). *Evolving Enactivism: Basic minds meet content*. MIT Press.
Hutto, D. D., & Myin, E. (2018). Much ado about nothing? Why going non-semantic is not merely semantics. *Philosophical Explorations*, 21(2), 187–203. https://doi.org/10.1080/13869795.2018.1478988
Hutto, D. D., & Satne, G. (2015). The natural origins of content. *Philosophia*, 43, 521–536. https://doi.org/10.1007/s11108-015-9644-0
Kauffman, S. A. (1971). Cellular homeostasis, Epignesis and replication in randomly generated macro-molecular systems. *Journal of Cybernetics.*, 1(1), 71–96.
Kauffman, S. A. (1986). Autocatalytic sets of proteins. *Journal of Theoretical Biology*, 119, 1–24.
Kauffman, S. A. (1993). *The origins of order*. Oxford University Press.
Kim, H., Davies, P., & Walker, S. (2015). New scaling relation for information transfer in biological networks. *Journal of the Royal Society Interface*, 12, DOWE. https://doi.org/10.1098/rsif.2015.0944
Lean, O. (2014). Getting the most out of Shannon information. *Biology and Philosophy*, 29, 395–413.
McNamara, J. M., & Dall, S. R. X. (2010). Information is a fitness enhancing resource. *Oikos*, 119, 231–236.
Millikan, R. G. (1984). *Language, thought, and other biological categories*. MIT Press.
Myin, E. (2020). On the importance of correctly locating content: Why and how REC can afford affordance perception. *Synthese*. https://doi.org/10.1007/s11229-020-02607-1
Neander, K. (1997). The function of cognition: Godfrey-Smith’s environmental complexity thesis. *Biology and Philosophy*, 12, 567–580.
Neander, K. (2017). A mark of the mental. In *Defence of informational Teleosemantics*. MIT Press.
Nurse, P. (2008). Life, logic and information. *Nature*, 454, 424–426.
Oizumi, M., Albantakis, L., & Tononi, G. (2014). From the phenomenology to the mechanisms of consciousness: Integrated information theory 3.0. *PLoS Computational Biology*, 10(5), e1003588. https://doi.org/10.1371/journal.pcbi.1003588
Pianka, E. R. (1970). On r- and K-selection. *The American Naturalist*, 104, 592–597.
Ramsey, W. M. (2009). *Representation reconsidered*. Cambridge University Press.
Rivoire, O., & Leibler, S. (2011). The value of information for populations in varying environments. *Journal of Statistical Physics*, 142, 1124–1166.
Rowlands, M. (2017). Arguing about representation. *Synthese*, 194(11), 4215–4232.
Roy, J.-M. (2015). Anti-Cartesianism and anti-Brentanism: The problem of anti-representationalist intentionalism. *The Southern Journal of Philosophy*, 53(Spindel Supplement), 90–125.
Shapiro, L. (2019). [first published 2010]. *Embodyed Cognition 2nd Edition*. Routledge.
Short, T. L. (2007). *Peirce’s theory of signs*. Cambridge University Press.
Sievers, D., & von Kiedrowski, G. (1994). Self-replication of complementary nucleotide-based oligomers. *Nature*, 369, 221–224.
Smith, J. M. (2000). The concept of information in biology. *Philosophy of Science*, 67, 177–194.
Smith, J. I., Steel, M., & Hordijk, W. (2014). Autocatalytic sets in a partitioned biochemical network. *Journal of Systems Chemistry*, 5, 2.
Vaidya, N., Manapat, M. L., Chen, I. A., Xulvi-Brunet, R., Hayden, E. J., & Lehman, N. (2012). Spontaneous network formation among cooperative RNA replicators. *Nature*, 491, 72–77.
van Gelder, T. (1995). What might cognition be, if not computation? *The Journal of Philosophy, 92,* 345–381.
Vasas, V., Fernando, C., Santos, M., Kauffman, S., & Szathmáry, E. (2012). Evolution before genes. *Biology Direct, 7,* 1–14.
Walker, S. I., Kim, H., & Davies, P. C. W. (2016). The informational architecture of the cell. *Philosophical Transactions of the Royal Society A, 374,* 201550057.
Wilson, A. D., & Galonka, S. (2013). Embodied cognition is not what you think it is. *Frontiers in Psychology, 4,* 58. [https://doi.org/10.3389/fpsyg.2013.00058](https://doi.org/10.3389/fpsyg.2013.00058)

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.