Anticipatory Representational Mechanisms in Animals

Some of animal behavior can be explained by appeal to their internal or mental representations. For example, it is usually agreed that rats are capable of path integration (even in complete darkness, and when immersed in a water maze) because they maintain a cognitive map of their environment. Exactly how and why neural states give rise to mental representations is a matter of an ongoing debate. The purpose of this paper is to present a general mechanistic framework to analyze representational claims made in behavioral science, and offer a way to distinguish genuinely representational explanations from the ones that invoke representational talk just for honorific purposes. In particular, it will be shown that anticipatory mechanisms involved in rats' cognitive maps do meet Ramsey's (2007) "job description challenge": it is clear in what way they are *representationally* relevant for explaining and predicting rats' behavior.

First, I introduce the idea of representational mechanisms, which is used to analyze the current research in ethology, cognitive science, and neuroscience. Then, I present a particular subspecies of representational mechanisms, the anticipatory ones. The framework is then applied to the example of cognitive maps in rats. Based on current neuroscientific evidence, it is suggested that the mechanisms involved in dead reckoning in rats are anticipatory, in contrast to simple ballistic tracking mechanisms present in invertebrates. In conclusion, I point out in which way the account of representational mechanisms meets Ramsey's challenge.

1. Representational mechanisms

The notion of representation has come recently under fervent attack from several camps. The proponents of the dynamic and embodied accounts of cognition suggest that the notion can be entirely eliminated from behavioral and cognitive sciences (Chemero 2009; Garzon 2008; Hutto 1999; Hutto and Myin 2013; Keijzer 2001). Some have argued that at least a large number of appeals to representation in cognitive sciences and neuroscience can be explained away in a deflationary manner; in particular, simple feature detectors and tracking mechanisms do not seem to warrant any genuinely representational talk (Ramsey 2007). In addition, proponents of dynamical explanations of cognition charged their opponents with the objection that their use of 'representation' is overly liberal, by allowing representation in such systems as Watt governor, which seems to imply that representations abound but at the same are trivialized as far as they are not in any way special to cognitive systems (Van Gelder 1995; Bechtel 1998; Stepp, Chemero, and Turvey 2011; Nielsen 2010). My strategy in answering the attacks will be to sketch a general mechanistic framework, whose purpose is to constrain the notion of representation without committing oneself to deciding the features of representational mechanisms that may or may not occur in cognitive systems. In other words, I will suggest that a large number of questions are not to be answered in this conceptual framework and that the particular answers should be given as based on empirical evidence. In other words, what I offer is a general specification, or a *schema*, of a representational mechanism, whose particulars are to be decided by referring to empirical evidence. The schema may be filled differently as based on extant theories and evidence in particular cases. In this paper, one such filling will be offered.

Also, I do not present a conceptual analysis of the notion of representation; any such analysis should offer a lexical definition, but the notion itself is used in many incompatible ways, resulting in vagueness, confusion, and ambiguity. Although I do think that my account can be treated as precising the current usage, the relationship between the representational mechanisms and different notions of representation is the same as between Millikan's notion of proper function and various notions of function in use (Millikan 2002). The term 'representational mechanism' is technical and whether it corresponds or not to what is called 'representation' in everyday speech (or folk psychology) is not my concern here.

Mechanism is now one of the most successful accounts of explanation in special sciences, and has been applied to analyze explanatory strategies in such areas as biology, neuroscience, and cognitive sciences (Machamer, Darden, and Craver 2000; Glennan 2002; Bechtel 2008; Bechtel 2011; Bechtel and Abrahamsen 2005; Craver 2007; Piccinini 2007; Miłkowski 2013). While the definitions of mechanisms offered by various authors accentuate different aspects, the main idea can be summarized in the following way: mechanisms are complex functional structures, involving organized components and processes (or activities) that interact with each other and contribute jointly to a function of the structure. The mechanistic explanation is a species of causal explanation, and interactions of components are framed in causal terms. For my purposes here, an important distinction is to be made between a complete description of the mechanism (sc. complete with regard to causally relevant components), which is required for its explanatory relevance, and incomplete descriptions. Craver distinguishes two kinds incomplete descriptions of mechanism: a *sketch*, which represents the mechanism with gaps (usually filled with generic filler terms), and a *schema*, which is a truncated representation that abstracts from some details of organization for presentation purposes. Importantly, these details are not to be filled with generic filler terms. What I offer below is supposed to be a mechanism schema: it will do without using generic terms such as 'inhibit,' 'encode,' 'represent' but it is still to be considered a template for more concrete representations. The present account should not be thought of as offering an abstract specification of organization of representational mechanisms, as this organization is not explanatorily illuminating without filling out the details. In other words, it is not an abstraction in the sense of Levy & Bechtel (2013).

As many authors argued, there are genuine advantages in using the mechanistic framework in analyzing theoretical entities posited in special sciences. One could justify the use of the framework just by pointing out that representation is a theoretical entity in behavioral, brain and cognitive sciences, so applying mechanism to analyze the account of this entity is just natural. But there are specific advantages of using the mechanistic framework. First, the mechanistic explanation requires one to specify the capacity (or capacities) of the mechanism (the explanandum phenomenon), which is then explained causally. By following this practice, we need to ensure both that we specify exactly the capacity of the mechanism – including the interest of the researcher who wants to explain it – and that the overall account is causal, which is naturally linked with the result of naturalizing representation. Second, mechanistic explanation is linked with the focus on organization of the system, in which the mechanism is posited. Instead of abstracting away from the exact role of the representation in an organism that peruses it, the mechanistic account requires that the role is specified. Representation cannot simply float freely without being a part of a complex system. This way, the mechanistic framework by itself requires that Ramsey's challenge be met.

While representational mechanisms may be used both to explain the work of larger mechanisms (also, as explanantia), I will focus them as explananda. However, depending on how the explanatory problem is posed, the capacity of a representational mechanism may be framed in various manners. Another way to spell out this point is to say that the function ascription is interest-dependent: depending on what the theorist wishes to explain about the system, the capacity of the representational mechanism may be carved this way or another. What is common in all such explanations is that the representational mechanism has a capacity to make some information available for the cognitive system (Miłkowski 2013). The information in this case becomes *semantic* as far as it modifies the readiness of the system to act this way or another (or, to use a slightly more precise phrase, the conditional probabilities of actions of the system are changed accordingly when the information changes, cf. (MacKay 1969)). Note that in my account of the mechanism, capacities of the mechanism are framed in functional terminology; depending on how robust one's notion of function is, the account of representational mechanism may be made more or less similar to the requirements commonly accepted in teleosemantics (Millikan 1984; Dretske 1986; Cao 2011), namely that the information in question has a function (in some non-trivial sense) to be semantic.

By framing the capacity of the representational mechanism as modifying the readiness to act as based on the information available to the cognitive system, this framework is committed to two theses. First, that representation is essentially action-oriented, albeit this orientation does not mean that all representation is linked with activation of effectors of the system, or that representation is there just to control the motor activity of the system. There might be content that is not exploited in action; what is changed is just the *readiness* to act. It seems that other action-oriented accounts of representation try to avoid the charge of introducing a too close linkage of representation with motor activity (see, e.g., (Clark 2001)) in a similar vein (Anderson and Rosenberg 2008). The notion of action is to be understood in a liberal way, so that it includes cognition as well. Second, it makes non-trivial use of the notion of information. Though there are various mathematical measures of information, they should not be confused with the notion of information, as MacKay (1969) stressed. All I need for my purposes here is that there is a physical medium with at least two degrees of freedom (or two levels of logical depth) that make difference to the behavior of the system; in other ways, the system reacts differently to at least two distinct states of the physical medium. The notion of information introduced informally here is equivalent to so-called structural information (called *logon* by MacKay); by introducing senders and receivers, a channel and the uncertainty of the receiver, we could use the Shannon's measures as well. For my purposes here, it is not needed, though these measures *are* applicable as long as one can talk of information being received by some subsystem. I leave these technical details however mostly aside in this paper.

The initial specification of the capacity of the representational mechanism is not to be confused with a full-blown theory of representation. The mechanism has other important capacities cited in explanatory texts. Usually, the notion of representation is introduced to talk of targets (or referents, or extension) to which the representation refers, and to talk of the characteristics of the targets (or intension). Also, the information cannot simply 'sit' there in the mechanism; the system has to care about it somehow. Thus, there are at least three other capacities of the mechanism in question:

(a) Referring to the target (if any) of the representation;

(b) Identifying the characteristics of the target;

(c) Evaluating the epistemic value of information about the target.

While the first two capacities bear close resemblance to traditional notions of extension and intension, which have been used in one way or another in theories of meaning at least since John Stuart Mill (Frege 1960; Carnap 1947; Mill 1875; Cummins and Roth 2012; Millikan 1984; Cummins 1996), the third one is supposed to link the representational mechanism with the work of the agent or system that peruses it. Note that even if the traditional claim that intension fully determines extension has been rightly criticized (Putnam 1975), it does not mean that the partial determination of extension cannot be at least sometimes successful if there are some referents (members of the extension). Realistically speaking, animals that are able to detect some typical features of food might err in non-standard environments (one does not need to talk of a Twin Earth here to notice the environmental dependence), but partial determination might be good enough for them to survive. By including the characteristics of the target as one of the possible (though not necessary) features of the representational mechanisms, I do not want to suggest that it is impossible for purely extensional representations to exist.

Some might object that including the characteristics of the target violates the principle of parsimony. For example, the proponents of the causal theory of reference seem to dispose of the remnants of the idea of intension altogether. However, even if their account of some neural states as simple indicators is cogent, distributed neural structures cannot be framed in terms of a mere hotchpotch of individual indicators. It was even argued that indicators are not full-fledged representations just because they do not seem to characterize targets (Cummins 1996; Ramsey 2007). Also, I am not defending Descriptivism, according to which representation always requires description of targets, and the characteristics of the target may be (though need not) understood in terms of 'mental files' (Récanati 2012). The account of representational mechanisms is ecumenical: whether a given system peruses a richly structured medium or not is a matter for empirical investigation, and not for armchair conceptual analysis. Note as well that referential opacity, which is arguably one of the marks of genuine representation (Dennett 1969), is easily explained by appeal to the characteristics of the target. Without at least some minimal access to the characteristics (such as knowing the label used to identify the target, or having mental files *sensu* Récanati), intensionality is hard to pin down.

The third capacity of the representational mechanism, namely evaluating the epistemic value of information about the target, may seem misplaced, and not directly related to representing at all. Granted, sometimes it may be beneficial to abstract away from such factors; for example, formal grammars are usually not related to any epistemic values (at least not directly, though in some theories, tractability of the computational account is an important issue). Nevertheless, in many psychological and behavioral theories, a complete story about representing is linked with a story about the value of the representation for the organism in question. For example, an ethological story about an animal representing a smell as indicator of a predator makes, at minimum, an implicit appeal to the adaptive value of predator avoidance. An important point here is that different values may lead to complex trade-offs, which mean that truth, accuracy or certainty are not to be treated as exclusive epistemic values. For a beaver, it may be beneficial to misrepresent some sound as indicating danger rather than to ignore it (Millikan 1989). In other words, behavioral, brain, and cognitive theories usually show why the system cares about having the representation in the first place; this is related to the representation's adaptive value. However, the adaptive value may be related in various ways with other values here, such as tractability, generality, precision, accuracy, truth, or certainty. Also, ignoring some information may be beneficial because of the irrelevance of the ignored information for action; for example, city maps, even if created from satellite photos, do not depict how cars were parked at the time when the photo was taken. Cars are noise for a cartographer. But for a traffic engineer or car park designer, the position of cars is essential, and buildings may be noise. In other words, with the same available input information, another representation can be created by processing different pieces of input, depending of its relevance for one's goals. For this reason, maximal accuracy in preserving all detail is almost never beneficial.

Although I introduced some examples of epistemic values, not all of them are evaluable (truth is not computationally evaluable owing to known incompleteness theorems, for example) or practically evaluable. For example, it would be too resource consuming for animal brains to evaluate tractability of some algorithms of information processing; a faster and more frugal way to deal with speed requirements would be simply to stop processing after some period (which may be behaviorally manifested as 'boredom' etc.).

All abovementioned representational capacities are closely interlinked. The semantic information that modifies the readiness for action of the cognitive system, which may be in the form of characteristics of the target, or simply indicate the target, is what is epistemically evaluable. It is time now to introduce the main idea of how the mechanism is organized to display these capacities in a stable way. As it will turn out, there are considerations that make epistemic evaluation especially important. To evaluate the information already present in the system, the mechanism needs to be able to compare two sets of characteristics of the target. In other words, epistemic evaluation requires more than negative feedback in the information-processing mechanism: this feedback simply modifies the system's input value. What is required instead is that the error is detected by the system. The idea that system-detectable error gives rise to genuine representationality is by no means new (for an extended argument, see Bickhard (1993; 1998)). Let me elaborate.

Simple negative feedback systems, such as the notorious Watt governor, are *not* representational mechanisms in the sense introduced in this paper. The Watt governor is a mechanical device for stabilizing the speed of the steam engine and its work is usually explained in terms of control theory; it is used as a metaphor for dynamical account of cognition (Van Gelder 1995). The governor includes two heavy balls on a frame that is driven by the engine, and the centrifugal force makes them go up or down, and then closes and opens the steam valve. In other words, there is negative feedback between the speed as detected by the centrifugal governor and the engine. Yet even if one analyzes the Watt governor in information-processing terms (as Bechtel (1998) or Nielsen (2010) do), the structure of the system does not contain any error-processing routines. Instead of *evaluating* and *modifying* its previous *output information* (the "signal" to open or close the valve), it simply changes its operation by having different *input* (balls move up or down). For this reason, Watt governor does not detect any error in the information: although in the eyes of the beholder, it might have misrepresented the state of the steam engine that it controls, the error is not system-detectable.

System-detectability of error may be realized in various ways; for example, when the system has two independent sources of information (Dretske 1986), or when previous input information is used to predict the future input state, and then is compared with that state. The second solution, though somewhat similar to negative feedback, is different in that it influences the processing of information on multiple levels. This is typical, for example, for contemporary predictive processing frameworks (Clark 2013; Friston and Kiebel 2011). The idea of such predictive processing was commonly used in early cybernetics (Boccignone and Cordeschi 2012), and has traces at least to biosemiotics of Uexküll (Uexküll 2001a; Uexküll 2001b), who used special diagrams to depict it. To wit, a model of the future state of the system is built, and then the model is compared with that state, when it occurs; from a technical point of view, the model becomes another source of information, though not statistically independent (this kind of solution is analyzed in the following section). To summarize, a complete description of the representational mechanism needs to detail the structure of the evaluation subsystem by specifying exactly in what way the mechanism detects discrepancy between two pieces of input information or between the model and the input information.

Just because representational mechanisms need to process information (for example, to evaluate it), some of their functioning can be couched in computational terms; therefore, representation requires computation – but it is not reducible to computation, as the notions such as 'action' are clearly not computational (for an extended argument about the relationship of computation and representation, see (Miłkowski 2013)).

Let me summarize. The framework of representational mechanisms specifies their capacities and describes constraints that various mechanisms must fulfill in order to qualify as representational. The constraints, though highly abstract, guarantee not only that the mechanisms store and process information that modifies their readiness to act, but also that they are (at least in some cases) able to refer to targets, identify them, and (necessarily) evaluate the value of information stored. This way, representational mechanisms are able to detect that they are in error (via evaluation of the epistemic value) and, at least in some cases, are prone to misidentification of targets because of the referential opacity. Both aspects, namely system-detectable error (highlighted by Bickhard and Anderson) and referential opacity (accentuated by Dennett), have been discussed as the basis for causal relevance of content as content. In other words, the Ramsey's challenge can be met: representation has a distinctive role in the representational mechanism. We will see that in detail in the example of cognitive maps in rats, introduced in section 3 of this paper.

2. Anticipatory mechanisms

Anticipatory representational mechanisms not only represent but do so in an anticipatory manner; in other words, they anticipate the future characteristics of the represented target. Such capacities are posited widely in current cognitive science (Pezzulo 2008; Pezzulo 2011), but the conception has roots at least in Helmholtz's idea of unconscious inferences as inherent in active movements of the eye: A motor signal from the central nervous system is sent both to the motor system and, as a copy, to an internal forward model (Holst and Mittelstaedt 1950; Meulders 2010, chap. 9). Anticipation as the role of representation was posited by diverse researchers in behavioral and cognitive sciences. For example, it was popular in American pragmatism since Dewey's (1896) criticism of reducing thought to the reflex arc; in Soviet Union, P. K. Anokhin and N. Bernstein conjectured about it in cybernetic models of thinking and motor action (Sudakov 1998; Egiazaryan and Sudakov 2007; Sporns, Edelman, and Meijer 1998; Bongaardt and Meijer 2000; N. Bernstein 1967); Neisser (1976) made it part of his account of perceptual cycle; and Rosen thought that anticipatory systems are the key to understanding life (Rosen 2012). Interestingly enough, the idea was more often than not reinvented without knowing that it had been already proposed in the past; no wonder that this led to ambiguities (Poli 2010). In this paper, I distinguish *anticipation* from *prediction*. The latter term may be used to refer to a process of inferring current events; for example, in the predictive coding framework, the task of the brain is usually to predict current sensations, given their causes, and only in generalized predictive coding, the task may cover both current and future sensations (Friston, Mattout, and Kilner 2011, 138). In the terminology adopted in this paper, only predicting future sensations is anticipatory.

To posit an anticipatory representational mechanism in a cognitive system is to reject two claims. First, it is to reject anti-representationalism. These anticipatory mechanisms meet Ramsey's challenge; by appeal to them, one can show in what way representation is both causally relevant and useful for cognitive systems. Second, it is to reject the simplistic idea that representing is nothing over and above having information about past and present sensory stimuli. But what is the positive claim?

Let me introduce the idea of modeling relationship as presented by Rosen (Rosen 1991; Rosen 2012), as this can be used to shed some light on anticipation. His theory is framed in terms of complex diagrams and explicated in category theory, but the idea is straightforward: the natural system is modeled by a formal one if and only if the causal process in the natural system is congruent with the three elements: (a) measurement of a value of the natural system; (b) inference on that value; (c) and the inferred value in terms of some physical quantity. All in all, this just means that there is a mapping (at least a homomorphism) between two entities: (1) measurement values encoded by the formal system, inferences on values, and inferred values, decoded in terms of physical quantities and (2) the natural system. Rosen makes some assumptions that are not exactly essential for modeling relationship; for example, one can model a natural system in a non-causal manner. The definition of modeling relationship should not decide whether non-causal modeling is really modeling or not; it simply needs an additional argument, and Rosen failed to offer one. But even opponents of the idea of causality in fundamental physics (Russell 1912; Ross and Spurrett 2007) need not reject Rosen's analysis of modeling if they simply substitute 'causation' with any term they think would best describe the basic structure of physical processes. Also, the notion of inference needs not be understood as entailment in a logical calculus (after all, not all formal systems are logical calculi, and not all computation in formal systems is simply derivation of steps in a proof, even if all formal systems may be described by logical calculi); all we need is that there is an algorithm described by the formal system, which encodes measured values, and decodes them in terms of physical quantities. Note: Non-encoded values, or values that were not measured, are abstracted away, and ignored in models, but this does not undermine the modeling relationship.

Having defined the modeling relationship, we can now define the anticipatory system as the one "containing a predictive model of itself and/or of its environment, which allows it to change state at an instant in accord with the model's predictions pertaining to a later instant" (Rosen 2012, 312). It becomes clear that based on values measured in the cognitive system itself or in its environment the model is fed with these values and used to predict their later states. Importantly, without the model, there would not be any changes of states in the anticipatory system (and it would not be an anticipatory system in Rosen's sense).

One important distinction between two kinds of anticipatory systems was introduced by Dubois (2003). Namely, *weak anticipation* (or *exo-anticipation*) uses externally produced data to internally model future states of the environment, while *strong anticipation* (or *endo-anticipation*) uses internally produced data to model future internal states (cf. also Collier (2008)). Some authors argue that strong anticipation is not necessarily related to internal models and point to phenomena of anticipating synchronization (Stepp and Turvey 2010). In this case, however, strong anticipation might be reducible to the weak one, at least explanatorily. And, especially in the case of coupled and synchronized systems, one might be tempted to eliminate the notion of representation altogether. However, in anticipatory representational mechanisms, the organization of the mechanism requires more than anticipating synchronization.

Let me elaborate. The anticipatory representational system has a capacity to derive its own future state, and that future state is evaluated epistemically, for example by comparing whether there is a discrepancy between expected sensory information (the future state) and actual sensory input. The discrepancy is then, typically, used to correct the error for future derivations of states. Basically, any anticipatory system that has such error correction mechanisms is already a representational mechanism in my sense. Anticipating synchronization in strongly coupled dynamical systems does not have any role for error correction, and hence it need not be representational at all.

3. Anticipation, error, and path integration

Cognitive maps in rats are one example of representations that have a long history in the debates in psychology (Tolman 1948). For a theory of mental representation, they are a specially interesting case: they seem to be structured, even compositional, but not reducible to language-like symbol media (Rescorla 2009), they constitute an instance of what Cummins (1996) called *S-Representation* without being simply picture-like representations.

At first, cognitive maps seemed to be inevitable in explaining navigational capacities of rats, but later on, the role of cognitive maps has been questioned; even if contemporary neuroscience vindicates an idea of cognitive maps by locating them in hippocampus (Redish 1999; O'Keefe and Nadel 1978; Derdikman and Moser 2010), it is still not universally accepted that rats peruse them in navigation, or 'path integration', as this ability is usually called. Tolman wasn't very clear when he introduced the term 'cognitive map' and gave no definition in his seminal paper; no wonder that the term is used to mean different things, and one may enumerate at least three meanings commonly occurring in the literature: (1) the *trivial* one, in which a cognitive map is whichever mechanism involved in spatial navigation; (2) the *loose* one, in which the map is simply any representation that models the geometric aspects of the environment; (3) the *strict* one, in which cognitive maps have a format typical for maps, so they not only represent geometric aspects but do that in a geometric manner (Rescorla 2009, 381). Now, obviously, cognitive maps in the trivial meaning are always present in all spatially navigating animals, but only trivially so. The debate about the existence of cognitive maps in rats is obviously over the loose and strict meaning. The question then is whether navigation of rats really depends on cognitive maps.

The reason given for skepticism is that even in the circumstances, in which the use of cognitive maps would be beneficial, rats seem to unable to reach their goals (Benhamou 1996; Whishaw 1991). Usually, they seem to prefer simple visual cues, and for that reason, there would be no reason to assume multimodal path integration in rats; their behavior would be fully explainable in terms of tracking the environmental cues, just like tropisms or taxis in simpler organisms. Invertebrates seem to navigate successfully without multimodal integration (Cruse and Wehner 2011).

The whole story about rats is however much more complex. Behavioral experiments confirm that at least in some cases, rats are able to return to their starting position, even if they were exploring the environment in complete darkness, devoid of smell (in a water platform or a water maze), and without whiskers to orient towards walls, simply by using their motor signals and vestibular system (Cheung et al. 2012). There are also fairly successful neurocomputational models of this ability (Conklin and Eliasmith 2005). The main problem with confirming the hypothesis about cognitive maps is that experiments need to be conducted in darkness, and one of the kind of neural cells responsible for navigation, head direction cells, become unstable after three minutes in darkness, while place and grid cells (other components of the neural system of navigation) seem to fire stably for another half an hour. Cheung et al. have shown that path integration beyond three minutes is theoretically implausible but landmarks alone (even if they are known perfectly) cannot suffice for a stable positional representation. Hence, a complex organization emerges, one that involves both path integration and cue tracking.

In other words, path integration in rats is not modular (or encapsulated in the sense of Fodor (1983)): it uses various kinds of information. In principle, it could be modular, just like in insects, but the hippocampus in rodents has a function of integration of various sources of information (Lisman and Redish 2011). More precisely, there are hypotheses (Conklin and Eliasmith 2005) that this integration involves using sensory information for error correction. For this reason, the architecture of the system underlying navigation in rats seems to correspond to what I have dubbed 'a representational mechanism.' In addition, this mechanism is anticipatory: the error correction concerns the prediction made by the path integration system, as its function is to predict the location of the rat as based on its previous location and his own motor commands. When the prediction is in conflict with data from memory (a landmark is found in an unexpected location), the system updates its predictions. A skeptic might say that what is predicted is not the future location of the rat but its current location; hence, the representational mechanism is merely predictive and not really anticipatory.

However, there is empirical evidence that landmark location is retrieved in an anticipatory fashion, leading to phenomena called "phase precession": experiments show that many spikes fired by place cells actually represent a position ahead of the rat (Lisman and Redish 2011, 261). These spikes do not occur when the animal is put on the running wheel. In other words, there is a specific way of modeling the future location in place cells (using phase precession that does not occur normally, so the rat can distinguish between anticipating and actually being located at the landmark). But these anticipations are only about locations less than a meter from the current position of the rat that the animal will usually arrive at in several seconds. Rats, however, according to some (Naqshbandi and Roberts 2006) seem to lack prospective memory, or the kind of memory involved in planning future action at larger timescales, even if there is some indirect evidence for it (Wilson and Crystal 2012). Whatever the case may be, the anticipation in representational mechanisms is not required to span larger timescales at all.

The ability to predict future location ahead of the rodent is just part of the complex mechanism of path integration, which is still poorly understood. But mere evidence of anticipatory encoding of location does not suffice to establish that the path integration mechanism is actually representational. As I claimed in the section above, what is needed as well is an evaluation subsystem. There is evidence that when people realize that they make mistakes (for example, when forced to solve tasks very quickly), there is specific error-related negativity in recordings of their electrical brain activity (P. S. Bernstein, Scheffers, and Coles 1995; Coles, Scheffers, and Fournier 1995; Wessel et al. 2012). Also, there is a specific kind of neural encoding of prediction error related to reward, particularly in the mesolimbic dopaminergic system (Schultz and Dickinson 2000), and the dopamine signal encodes for "prediction error" (Fiorillo, Tobler, and Schultz 2003). What is to be expected in rats is that some such evidence to be found, and there are studies that confirm exactly that for head direction neurons (Valerio and Taube 2012). So even if the overall structure of the mechanism is unknown, it seems fairly plausible that rodents have abilities that seriously outstrip mere reactive landmark tracking. Owing to the multimodal nature of information integrated in the hippocampus, the mechanism of path integration will comprise multiple error correction mechanisms for all sensory systems and for different kinds of neural cells involved in navigation.

Let me wrap up. I argued that it is plausible to hypothesize that anticipatory representational mechanism is involved in path integration in rats. Rats are able to detect errors in their representation, which is found at least for head direction cells and is plausible for other hippocampal cells, and that fulfills one of the criteria for representational systems. Also, misidentification of location is possible, and even if that is undetected by the rat, it is explainable by assuming that it misrepresented the location. In other words, referential opacity is possible for cognitive maps.

4. Conclusion

The account of representational mechanisms is sufficient to answer Ramsey's challenge. It points out that there are at least two ways in which representations are causally relevant: by being referentially opaque, they are indispensable in explaining behavior, especially in cases of misrepresentation; and when their being in error is detectable by the system, it shows that the system really treats them as representations. In other words, they are not just factors useful in predicting and explaining the system's behavior; they are specifically representational factors.

Anticipatory mechanisms are subspecies of the general category of representational mechanisms; one that is particularly important for biological reasons. Simply, ruminating about the past seems to be much less biologically valuable than planning ahead and avoiding future danger. This said, I do not want to suggest that other kinds of representational mechanisms cannot exist. Obviously, we humans can ruminate about our past. Also, predictive mechanisms, as posited for example in the Bayesian account of the brain, can be used to determine current sensory input. Knowing the current situation is just as biologically important. At the same time, a growing body of evidence shows that there are multiple mechanisms geared specifically towards anticipating the future.

In the case of path integration in rats, the account of anticipatory mechanism, along with the empirical evidence cited, vindicates representational talk about cognitive maps. Granted, current knowledge of neural mechanisms in rats is incomplete and the overall organization of navigational system in rodents may turn out to be different from what is hypothesized today. But for my purposes this is not so important. The point was to show that anticipatory representational mechanisms are entities that are explanatorily and predictively relevant. By assuming that they operate during navigation, we make robust predictions (for example, that there will be a specific kind of anticipatory spiking).

Representational talk about cognitive maps is not just a matter of convenience, and my argument for using it is not based on the assumption that more facts can be subsumed under representational generalizations than under neural generalizations (this is one of the most popular arguments for representational explanation, cf. (Pylyshyn 1984)). By positing representational mechanisms, we may have a similar level of generality as before. But generalization is not the most important factor in explanation: After all, generalization may occur at the price of abstracting from details, and the account of navigation in terms of simple taxis or tropisms could be extrapolated from invertebrates to rodents, yielding certainly a higher level of generalization. Yet by using the account of representational mechanisms we may discover the function of neural organization and understand why there are error correcting pathways at all. In other words, there is no real worry that the notion of representation may turn out to be eliminable here; the function of neural mechanism is to represent, and to establish this, one needs not to warrant that there must be a gap between representational explanations and neural explanations. In this case, integration of evidence at multiple levels of organization of mechanisms seems to be much more important than the purported autonomy of talk at individual levels.

References

Anderson, M.L., and Gregg Rosenberg. 2008. “Content and Action: The Guidance Theory of Representation.” *Journal of Mind and Behavior* 29 (1-2): 55–86.

Bechtel, William. 1998. “Representations and Cognitive Explanations: Assessing Dynamicist’s Challenge in Cognitive Science.” *Cognitive Science* 22 (3): 295–318.

———. 2008. *Mental Mechanisms*. New York: Routledge (Taylor & Francis Group).

———. 2011. “Mechanism and Biological Explanation.” *Philosophy of Science* 78 (4) (October): 533–557. doi:10.1086/661513.

Bechtel, William, and Adele Abrahamsen. 2005. "Explanation: a Mechanist Alternative." *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 36 (2): 421–441. doi:10.1016/j.shpsc.2005.03.010.

Benhamou, S. 1996. "No Evidence for Cognitive Mapping in Rats." *Animal Behaviour* 52 (1): 201–212.

Bernstein, Nikolai. 1967. *The Co-ordination and Regulation of Movements....* Oxford ;;London ;;P. ;;Edinburgh: Pergamon Press.

Bernstein, Peter S, Marten K Scheffers, and Michael G. H. Coles. 1995. "'Where Did I Go Wrong?' A Psychophysiological Analysis of Error Detection." *Journal of Experimental Psychology: Human Perception and Performance* 21 (6) (December): 1312–1322. doi:10.1037/0096-1523.21.6.1312.

Bickhard, Mark H. 1993. "Representational Content in Humans and Machines." *Journal of Experimental & Theoretical Artificial Intelligence* 5 (4) (October 1): 285–333. doi:10.1080/09528139308953775.

———. 1998. "Levels of Representationality." *Journal of Experimental & Theoretical Artificial Intelligence* 10 (2) (April): 179–215. doi:10.1080/095281398146824.

Boccignone, Giuseppe, and Roberto Cordeschi. 2012. "Predictive Brains: Forethought and the Levels of Explanation." *Frontiers in Psychology* 3 (November): 2–3. doi:10.3389/fpsyg.2012.00511.

Bongaardt, R., and O G Meijer. 2000. "Bernstein's Theory of Movement Behavior: Historical Development and Contemporary Relevance." *Journal of Motor Behavior* 32 (1) (March 31): 57–71. doi:10.1080/00222890009601360.

Cao, Rosa. 2011. "A Teleosemantic Approach to Information in the Brain." *Biology & Philosophy* 27 (1) (October 15): 49–71. doi:10.1007/s10539-011-9292-0.

Carnap, R. 1947. *Meaning and Necessity*. 1st ed. Chicago: Univ. of Chicago Press.

Chemero, Anthony. 2009. *Radical Embodied Cognitive Science*. Cambridge, Mass.: MIT Press.

Cheung, Allen, David Ball, Michael Milford, Gordon Wyeth, and Janet Wiles. 2012. "Maintaining a Cognitive Map in Darkness: The Need to Fuse Boundary Knowledge with Path Integration." Ed. Olaf Sporns. *PLoS Computational Biology* 8 (8) (January): e1002651. doi:10.1371/journal.pcbi.1002651.

Clark, Andy. 2001. "Visual Experience and Motor Action: Are the Bonds Too Tight?" *Philosophical Review* 110 (4) (October 1): 495–519. doi:10.1215/00318108-110-4-495.

———. 2013. "Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science." *The Behavioral and Brain Sciences* 36 (3) (June 10): 181–204. doi:10.1017/S0140525X12000477.

Coles, Michael G.H., Marten K. Scheffers, and Lisa Fournier. 1995. "Where Did You Go Wrong? Errors, Partial Errors, and the Nature of Human Information Processing." *Acta Psychologica* 90 (1-3) (November): 129–144. doi:10.1016/0001-6918(95)00020-U.

Collier, John. 2008. "Simulating Autonomous Anticipation: The Importance of Dubois' Conjecture." *Bio Systems* 91 (2) (February): 346–54. doi:10.1016/j.biosystems.2007.05.011.

Conklin, John, and Chris Eliasmith. 2005. "A Controlled Attractor Network Model of Path Integration in the Rat." *Journal of Computational Neuroscience* 18 (2): 183–203. doi:10.1007/s10827-005-6558-z.

Craver, Carl F. 2007. *Explaining the Brain. Mechanisms and the Mosaic Unity of Neuroscience*. Oxford: Oxford University Press.

Cruse, Holk, and Rüdiger Wehner. 2011. "No Need for a Cognitive Map: Decentralized Memory for Insect Navigation." Ed. Karl Friston. *PLoS Computational Biology* 7 (3) (March): e1002009. doi:10.1371/journal.pcbi.1002009.

Cummins, Robert. 1996. *Representations, Targets, and Attitudes*. Cambridge, Mass.: MIT Press.

Cummins, Robert, and Martin Roth. 2012. "Meaning and Content in Cognitive Science." In *Prospects for Meaning*, ed. Richard Schantz. Berlin & New York: de Gruyter.

Dennett, Daniel C. 1969. *Content and Consciousness*. London: Routledge and Kegan Paul.

Derdikman, Dori, and Edvard I Moser. 2010. "A Manifold of Spatial Maps in the Brain." *Trends in Cognitive Sciences* 14 (12) (December): 561–9. doi:10.1016/j.tics.2010.09.004.

Dewey, John. 1896. "The Reflex Arc Concept in Psychology." *Psychological Review* 3 (4): 357.

Dretske, Fred. 1986. "Misrepresentation." In *Belief: Form, Content, and Function*, ed. Radu Bogdan, 17–37. Oxford: Clarendon Press.

Dubois, Daniel. 2003. "Mathematical Foundations of Discrete and Functional Systems with Strong and Weak Anticipations." In *Anticipatory Behavior in Adaptive Learning Systems Foundations, Theories, and Systems*, ed. Martin Butz, Olivier Sigaud, and Pierre Gérard, 2684:107–125. Berlin Heidelberg: Springer. doi:10.1007/978-3-540-45002-3_7.

Egiazaryan, Galina G, and Konstantin V Sudakov. 2007. "Theory of Functional Systems in the Scientific School of P.K. Anokhin." *Journal of the History of the Neurosciences* 16 (1-2) (January): 194–205. doi:10.1080/09647040600602805.

Fiorillo, Christopher D, Philippe N Tobler, and Wolfram Schultz. 2003. "Discrete Coding of Reward Probability and Uncertainty by Dopamine Neurons." *Science (New York, N.Y.)* 299 (5614) (March 21): 1898–902. doi:10.1126/science.1077349.

Fodor, Jerry A. 1983. *The Modularity of Mind*. Cambridge, Mass.: MIT press.

Frege, G. 1960. *Translations from the Philosophical Writings of Gottlob Frege*. Ed. Peter T. Geach and Max Black. Oxford: Blackwell.

Friston, Karl, and Stefan Kiebel. 2011. “Predictive Coding: A Free-Energy Formulation.” In *Predictions in the Brain: Using Our Past to Generate a Future,* ed. Moshe Bar, 231–246. Oxford: Oxford University Press.

Friston, Karl, Jérémie Mattout, and James Kilner. 2011. “Action Understanding and Active Inference.” *Biological Cybernetics* 104 (1-2) (February): 137–60. doi:10.1007/s00422-011-0424-z.

Garzon, F. C. 2008. “Towards a General Theory of Antirepresentationalism.” *The British Journal for the Philosophy of Science* 59 (3) (September 1): 259–292. doi:10.1093/bjps/axl007.

Van Gelder, T. 1995. “What Might Cognition Be, If Not Computation?” *The Journal of Philosophy* 92 (7): 345–381.

Glennan, Stuart. 2002. “Rethinking Mechanistic Explanation.” *Philosophy of Science* 69 (S3) (September): S342–S353. doi:10.1086/341857.

Holst, Erich, and Horst Mittelstaedt. 1950. “Das Reafferenzprinzip.” *Die Naturwissenschaften* 37 (20): 464–476. doi:10.1007/BF00622503.

Hutto, Daniel. 1999. “Cognition Without Representation?” In *Understanding Representation in the Cognitive Sciences*, ed. A. Riegler, M. Peschl, and A.von Stein, 57–74. New York: Kluwer Academic/Plenum Publishers.

Hutto, Daniel, and Erik Myin. 2013. *Radicalizing Enactivism: Basic Minds Without Content*. Cambridge Mass.: MIT Press.

Keijzer, Fred. 2001. *Representation and Behavior*. Cambridge, Mass.: MIT Press.

Levy, Arnon, and William Bechtel. 2013. “Abstraction and the Organization of Mechanisms.” *Philosophy of Science* 80 (2) (April): 241–261. doi:10.1086/670300.

Lisman, John, and A David Redish. 2011. “Prediction, Sequences, and the Hippocampus.” In *Predictions in the Brain: Using Our Past to Generate a Future,* ed. Moshe Bar, 258–270. Oxford: Oxford University Press.

Machamer, Peter, Lindley Darden, and Carl F Craver. 2000. “Thinking About Mechanisms.” *Philosophy of Science* 67 (1): 1–25.

MacKay, Donald MacCrimmon. 1969. *Information, Mechanism and Meaning*. Cambridge: M.I.T. Press.

Meulders, Michel. 2010. *Helmholtz: From Enlightenment to Neuroscience*. Tran. Laurence Garey. Cambridge, Mass.: MIT Press.

Mill, John. 1875. “A System of Logic, Ratiocinative and Inductive Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation”. London :: Longmans, Green, Reader, and Dyer,.

Millikan, Ruth Garrett. 1984. *Language, Thought, and Other Biological Categories : New Foundations for Realism*. Cambridge, Mass.: The MIT Press.

———. 1989. “Biosemantics.” *Journal of Philosophy* 86 (6): 281–297.

———. 2002. “Biofunctions: Two Paradigms.” In *Functions: New Essays in the Philosophy of Psychology and Biology*, ed. Andrew Ariew, Robert Cummins, and Mark Perlman. New York: Oxford University Press, USA.

Miłkowski, Marcin. 2013. *Explaining the Computational Mind*. Cambridge, Mass.: MIT Press.

Naqshbandi, Mariam, and William A Roberts. 2006. “Anticipation of Future Events in Squirrel Monkeys (Saimiri Sciureus) and Rats (Rattus Norvegicus): Tests of the Bischof-Kohler Hypothesis.” *Journal of Comparative Psychology (Washington, D.C. : 1983)* 120 (4) (November): 345–57. doi:10.1037/0735-7036.120.4.34.

Neisser, Ulric. 1976. *Cognition and Reality: Principles and Implications of Cognitive Psychology*. San Francisco: W.H. Freeman.

Nielsen, Keld Stehr. 2010. “Representation and Dynamics.” *Philosophical Psychology* 23 (6) (December): 759–773. doi:10.1080/09515089.2010.529045.

O’Keefe, John., and Lynn Nadel. 1978. *The Hippocampus as a Cognitive Map*. Oxford; New York: Clarendon Press ; Oxford University Press.

Pezzulo, Giovanni. 2008. “Coordinating with the Future: The Anticipatory Nature of Representation.” *Minds and Machines* 18 (2) (February 7): 179–225. doi:10.1007/s11023-008-9095-5.

———. 2011. “Grounding Procedural and Declarative Knowledge in Sensorimotor Anticipation.” *Mind & Language* 26 (1) (February 14): 78–114. doi:10.1111/j.1468-0017.2010.01411.x.

Piccinini, Gualtiero. 2007. “Computing Mechanisms.” *Philosophy of Science* 74 (4) (October): 501–526. doi:10.1086/522851.

Poli, Roberto. 2010. “The Many Aspects of Anticipation.” *Foresight* 12 (3): 7–17. doi:10.1108/14636681011049839.

Putnam, Hilary. 1975. “The Meaning Of’meaning'.” In *Mind, Language, and Reality. Philosophical Papers, Volume 2*, 3–52.

Pylyshyn, Zenon W. 1984. *Computation and Cognition: Toward a Foundation for Cognitive Science*. Cambridge, Mass.: MIT Press.

Ramsey, William M. 2007. *Representation Reconsidered*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511597954.

Récanati, François. 2012. *Mental Files*. Oxford: Oxford University Press.

Redish, A.D. 1999. *Beyond the Cognitive Map: From Place Cells to Episodic Memory*. Cambridge, Mass.: The MIT Press.

Rescorla, Michael. 2009. "Cognitive Maps and the Language of Thought." *The British Journal for the Philosophy of Science* 60 (2) (April 7): 377–407. doi:10.1093/bjps/axp012.

Rosen, Robert. 1991. *Life Itself: a Comprehensive Inquiry into the Nature, Origin, and Fabrication of Life*. New York: Columbia University Press.

———. 2012. *Anticipatory Systems: Philosophical, Mathematical, and Methodological Foundations*. 2nd ed. New York: Springer.

Ross, D., and D. Spurrett. 2007. "Notions of Cause: Russell's Thesis Revisited." *The British Journal for the Philosophy of Science* 58 (1): 45–76. doi:10.1093/bjps/axl027.

Russell, Bertrand. 1912. "On the Notion of Cause." *Proceedings of the Aristotelian Society* 13 (January 1): 1–26. doi:10.2307/4543833.

Schultz, W, and A Dickinson. 2000. "Neuronal Coding of Prediction Errors." *Annual Review of Neuroscience* 23 (January): 473–500. doi:10.1146/annurev.neuro.23.1.473.

Sporns, O., G.M. Edelman, and O.G. Meijer. 1998. "Bernstein's Dynamic View of the Brain: The Current Problems of Modern Neurophysiology (1945)." *Motor Control* 4 (2).

Stepp, Nigel, Anthony Chemero, and Michael T. Turvey. 2011. "Philosophy for the Rest of Cognitive Science." *Topics in Cognitive Science* 3 (2) (April 6): 425–437. doi:10.1111/j.1756-8765.2011.01143.x.

Stepp, Nigel, and M T Turvey. 2010. "On Strong Anticipation." *Cognitive Systems Research* 11 (2) (June 1): 148–164. doi:10.1016/j.cogsys.2009.03.003.

Sudakov, K. 1998. "To the Centenary of P.K. Anokhin, a Great Russian Physiologist." *Integrative Physiological and Behavioral Science : the Official Journal of the Pavlovian Society* 33 (2): 171–5.

Tolman, E C. 1948. "Cognitive Maps in Rats and Men." *Psychological Review* 55 (4) (July): 189–208.

Uexküll, Jakob Von. 2001a. "An Introduction to Umwelt." *Semiotica* 2001 (134) (July): 107–110. doi:10.1515/semi.2001.017.

———. 2001b. "The New Concept of Umwelt: A Link Between Science and the Humanities." *Semiotica* 2001 (134) (July): 111–123. doi:10.1515/semi.2001.018.

Valerio, Stephane, and Jeffrey S Taube. 2012. “Path Integration: How the Head Direction Signal Maintains and Corrects Spatial Orientation.” *Nature Neuroscience* 15 (10) (October): 1445–53. doi:10.1038/nn.3215.

Wessel, Jan R, Claudia Danielmeier, J Bruce Morton, and Markus Ullsperger. 2012. “Surprise and Error: Common Neuronal Architecture for the Processing of Errors and Novelty.” *The Journal of Neuroscience : the Official Journal of the Society for Neuroscience* 32 (22) (May 30): 7528–37. doi:10.1523/JNEUROSCI.6352-11.2012.

Whishaw, Ian Q. 1991. “Latent Learning in a Swimming Pool Place Task by Rats: Evidence for the Use of Associative and Not Cognitive Mapping Processes.” *The Quarterly Journal of Experimental Psychology. B, Comparative and Physiological Psychology* 43 (1) (February): 83–103.

Wilson, A. George, and Jonathon D Crystal. 2012. “Prospective Memory in the Rat.” *Animal Cognition* 15 (3) (May): 349–58. doi:10.1007/s10071-011-0459-5.