

# Thinking of oneself as someone: the structure of minimal self-representation

## Abstract

When investigating the nature of self-representation, one standard question concerns the types of property that need to figure in their content. Here, authors have claimed that self-representations need to be about spatial, temporal, bodily, or mental properties. However, we can also ask a second question: *how* does a self-representation need to represent these properties? It is this latter question that I address. I argue that a distinction between egocentric and non-egocentric forms of representation – known primarily from the literature on spatial cognition – also applies to representations of other kinds of property. I use examples drawn from animal cognition and developmental psychology to show how creatures non-egocentrically represent their temporal, bodily, and cognitive properties. These representations are, I submit, minimal self-representations: they involve representing one's properties so that an explicit differentiation is made between the system's and other objects' properties (or between the system's actual and merely possible properties), they are directly relevant for action, and they exhibit a kind of immunity to error through misidentification. The upshot is a view on which different creatures may self-represent (in this minimal sense) more or fewer kinds of property. More substantive conceptions of self-representation (for instance, as exemplified by neurotypical adult human beings) then require integrated minimal self-representations of the right kinds of property.

## 1. Introduction

Representations may be about different kinds of property, and they may represent these properties in different ways. A red apple may be represented by an image of a red apple or the words 'red apple'. The matter is no different for cognitive (and mental) representations: I can token the phrase 'red apple' in inner speech or conjure up an image of a red apple. This paper concerns the special sorts of representation with which a system may represent its own properties. I argue that these representations are (minimal) *self*-representations when they represent properties in a specific way.

Self-representation doesn't depend on *what* properties are represented but only on *how* they are represented.

A minimal self-representation is the simplest kind of representation that is a genuine self-representation. A genuine self-representation, a representation with *de se* content, is a representation with which a system represents *itself qua itself*. This means it fulfils two conditions. First, the representation explicitly differentiates between the system and the world. A creature that represents a *tree to be located to the left* doesn't self-represent since the fact that the object is to the left *of itself* isn't made explicit. The representation must sport a representational token that stands for the system and one that stands for the tree, rather than, say, a single token specifying the distance to a tree. In the latter case, the representational token may be used by the system as denoting a distance to itself, but that isn't explicit. Such a representation merely *concerns* the self, but self-representations are *about* the self (Perry, 1993). However, explicitly differentiating between self and other doesn't ensure that the system knows what token refers to itself. An animal may look into a mirror, represent someone to have certain features, but fail to realise that it is that someone. This animal may represent itself, but fail to do so *qua itself*. This brings us to the second condition: a self-representer must be disposed to update and use the represented information in a specific kind of way that links it to sensation and behaviour (Evans, 1982). It must be disposed to use sensory information about the object in question to update its representation of itself. If someone were to draw a spot on the animal's forehead, it should represent itself as now having a spot on its forehead. And this representation must be taken to be directly relevant for behaviour: the animal must be disposed to use the information for behaviour in the relevant ways. For instance, if the animal prefers not to have a spot on its forehead, it should infer from its self-representation that certain behaviours are needed to bring to this state about.

I argue that minimal self-representations are coordinated non-egocentric representations. We know non-egocentric (or allocentric) representations from the literature on spatial cognition, where they appear in the guise of cognitive maps (see Rescorla, 2017). Like the maps on our phones, cognitive maps represent the locations of various objects relative to each other. A fruit bat, for instance, represents its nest and a fruit tree with two representational tokens (Tsoar et al., 2011). The relation between these tokens corresponds to the spatial relation between the two objects. Such representations contrast with egocentric representations such as those employed in path integration. The *Cataglyphis* desert ant, for instance, employs path integration to continuously update a representation of the *distance and bearing* to its nest (Collett, 2019; Gallistel, 1989; Reid et al., 2012; Srinivasan, 2015).

The differences between egocentric and non-egocentric spatial representations are striking and bear on the question of self-representation (see Grush, 2000). In the egocentric case, the token with which the nest is represented specifies a relation from the nest to the ant's current location. However, the fact that this relation is repres-

ented does not figure in the representation's explicit content. Such a representation concerns the ant (it's accuracy depends on the ant's location) but it isn't about the ant (no representational token stands for the ant). The non-egocentric case is different: the spatial relation is explicitly represented as a relation between representational tokens. By means of a representational token that stands for the ant, such a representation explicitly differentiates between the system and the rest of the world. This representational token is hence *about* the bat.

Another difference is striking, too: while the egocentric representation necessarily concerns the ant itself (always representing a relation *to the ant*), the representer may disappear in a non-egocentric representation. The bat may represent its cave to be such-and-such a distance away from a food source without locating itself of the cognitive map. Non-egocentric representations represent objects and their properties as in principle independent of the system's state. However, to use its non-egocentric representation for navigation, the bat needs to know how to use the map; it must employ a specific representational token as specifying *its own* location. When this happens, I call the representation *coordinated*. When the representation is coordinated, the system has a special sort of know-how regarding its non-egocentric representation. It's disposed to infer sensorimotor information from the non-egocentric representation (and, hence, adapt its behaviour given the non-egocentrically represented information) and update the non-egocentric representation given its sensorimotor representations. For instance, the bat can calculate the motor commands required to reach its cave based on the represented relation between its own location and the cave.

Spatial properties are only one among the many properties creatures represent with coordinated non-egocentric representations. I discuss evidence to this effect for representations of temporal, bodily, and cognitive properties. For instance, the forward models postulated by certain computational approaches to motor control imply an explicit differentiation between various possible bodily states. This, I argue, means they are non-egocentric representations. To use such a model to predict one's future bodily state, the model needs to be supplied with information that allows it to locate its actual state, which makes it a coordinated representation.

Coordinated non-egocentric representations explicitly differentiate between the system's and other objects' properties or between the system's actual and merely possible properties. Coordination entails that one token is used in a special way by the system, namely as specifying the system's own state. This means, first, that relations from this token to other tokens imply sensorimotor information and that, hence, the system is disposed to take the represented information to be directly relevant for its behaviour. Second, sensorimotor information implies relations between this special token and other tokens so that the system is disposed to explicitly self-attribute properties given relevant sensory input.

Since coordinated self-representation explicitly differentiates between self and other and are linked in the requisite way to behaviour and sensation, they are genuine

(if minimal) self-representations. Whenever a system represents *any* properties in this way, it instantiates a minimal self-representation. And as we will see in the examples I discuss, such self-representations are exemplified by infants and a range of non-human animals.

Moreover, the proposed account also explains why certain self-attributed properties are immune to certain errors of misidentification (Evans, 1982; Perry, 2010; Shoemaker, 1968; Wittgenstein, 2007). On my account, these are the properties inferred on the basis of sensorimotor information, and the immunity is due to the fact that such information is necessarily self-concerned. Self-attributed properties that aren't immune in this way are inferred by identifying oneself with some object – such as for instance my identification with the person I see reflected in the mirror in front of me.

I hope that this paper will prove interesting to both researchers in the empirical sciences and those more theoretically inclined. Cognitive scientists may find in this paper a proposal of the cognitive processes involved in self-representation that can be operationalised by extending the well-known paradigms from the study of spatial cognition without thereby giving up philosophical rigour. This should allow us to study if, when, and how human and non-human animals (and even artificial agents) exhibit self-representation. I believe that we will find that self-representation – at least of the simpler varieties – is more common than we might have thought. This, in turn, may prompt philosophers to broaden their investigations beyond the case of adult human beings using linguistic representations. Genuine (if minimal) self-representation doesn't require, as Musholt (2012) thinks, that we represent our mental states. It's also not essentially linked to spatial cognition (Grush, 2000; Ismael, 2008) or the representation of our bodies (Hohwy & Michael, 2017; Metzinger, 2003), and it doesn't require the representation of temporal properties (Campbell, 1999; Peacocke, 2014).

I call accounts that require self-representations to carry specific kinds of content *substantive*. While my account suggests that substantive accounts of self-representation don't explain self-representation's deepest puzzles, they are useful when it comes to explaining, for instance, the self-representations of neurotypical adult human beings. My account implies that whatever properties figure in a substantive self-representation, they need to figure in a coordinated non-egocentric representation. Moreover, the account suggests that many arguments in the philosophical literature may concern the kind of content needed to represent a self's essential properties. If this is right, then the disagreements between such accounts might concern the metaphysics of selfhood rather than questions of self-representation.

The paper is structured as follows. After this introduction follow four sections that look at how various creatures represent various kinds of properties by means of coordinated non-egocentric representations. The first of these sections discusses the representation of spatial properties and does double duty as an introduction to some of the main concepts employed throughout the paper. These sections all imply – but

do not explicitly argue for – a link between coordinated non-egocentric representation and self-representation. This link is made explicit in the sixth section, which argues that coordinated non-egocentric representations are genuine (if minimal) self-representations. Section seven links minimal self-representation to substantive self-representation. Section eight concludes.

## 2. Spatial cognition and minimal self-representation

The distinction between egocentric and non-egocentric representation arose in the literature on spatial cognition, and it's here that certain authors have first put forward links between non-egocentric representations and self-representation. This, in combination with the fact that the concepts of path integration and cognitive maps provide a uniquely intuitive starting point, makes spatial cognition the ideal first stop on the road to a general account of minimal self-representation. This section will distinguish between egocentric and non-egocentric representations in the spatial domain and give you a first taste of why coordinated non-egocentric representation is genuine self-representation.

The *Cataglyphis* desert ant has become a poster child for path integration (Collett, 2019; Gallistel, 1989; Reid et al., 2012; Srinivasan, 2015). These ants often leave their nests on lengthy foraging expeditions, and even though these can send them as far as 200 metres from home, the ants (almost) always manage to return. This feat is all the more impressive considering that the ants' monotonous desert environment is mostly devoid of landmarks. Moreover, if we pick up a homeward-bound ant and place it at a different location, it will walk in the direction in which the nest would have been, had we not related the ant (Wehner & Srinivasan, 1981).

In path integration, a creature encodes objects' locations through  $\langle \textit{distance}, \textit{bearing} \rangle$  tuples that specify relations from the creature's current location (and orientation). For instance, a foraging desert ant may represent its nest to be 200 steps (distance) just behind it (bearing). As it moves, the ant must update its representation to keep track of the object. These updates may be guided by proprioceptive information – for instance about the number of steps taken (Wittlinger et al., 2006) – which helps explain how an ant succeeds at navigating without external cues. Path integration may also account for the ant's behaviour when being picked up. In this scenario, proprioceptive information is lacking, therefore the ant does not update its representation and continues to navigate as if it had not been displaced.

Path integration involves the updating of *egocentric* spatial representations. In such representations, the origin of the represented space is necessarily the representer itself, and all objects' locations are given in relation to this source (Grush, 2000; Klatzky, 1998). A  $\langle \textit{distance}, \textit{bearing} \rangle$  tuple, for instance, necessarily represents the distance and bearing *from the representer's current location and orientation*. A

creature that represents various objects with such a representation arranges these objects around itself. One object may be *just behind*, another *some distance towards the front*, and yet another *far away to the left*. Note that the representational tokens all define relations that originate in the system itself and that no relations between (non-self) objects are represented.

As there exists, in the representation, a term or token that stands for the distance and bearing, these are *explicitly* represented. Something is explicitly represented ‘if and only if there actually exists in the functionally relevant place in the system a physically structured object [...] for which there is a semantics or interpretation, and a provision (a mechanism of some sort) for reading or parsing the formula’ (Dennett, 1982, p. 216). This is the case here: the representation contains a *⟨distance, bearing⟩* token, and (presumably) there exists a mechanism with which the ant can parse this token to employ it in navigation.

In contrast, the fact that the *⟨distance, bearing⟩* token tracks *a relation between the system’s location and orientation* and some object is only *tacitly* represented. The representation isn’t *about* the system itself, it merely *concerns* it (Perry, 1993). While the ant employs the representation to plot a path for *itself* to the nest, it does so without parsing the token to determine that a relation to itself is represented; it automatically employs *all* tokens as representing such relations. Dennett (1982), referencing Ryle (2009), argues that explicit representation needs to bottom out in non-representational *know-how*. The desert ant doesn’t represent its own location explicitly, but it knows how to use its egocentric representations to compute how it itself (and only it itself) can get to places. While the ant represents a relation, it doesn’t represent it *qua* relation (after all, we’re missing a *relatum*). Again, the fact that a relation is represented is tacit in the ant’s use of the representation.<sup>1</sup>

Another way of getting at the same point is to think about what makes the ant’s representation accurate or successful. Quite obviously, if the nest weren’t where the ant represented it to be, the representation would be inaccurate. But that’s not the only way such a representation can fail. If the very same representation were to be instantiated by a different ant (at a different location), the representation would likewise fail to successfully guide behaviour. This means that the ant’s state – the tacitly represented state implied by how the ant uses the representation – is an *unarticulated constituent* of the representation (Perry & Blackburn, 1986; see also Musholt, 2015). Even though the representation merely concerns the ant itself (and isn’t about it), the ant’s state enters into the representation’s success conditions.

As the representations employed in path integration do not contain a token standing for the system itself, they cannot underwrite an explicit differentiation between

---

<sup>1</sup> What I, after Dennett, call *tacit* representation, is sometimes called *implicit* representation in the philosophical literature (Dienes & Perner, 1999; Musholt, 2013). In the psychological literature, ‘implicit’ and ‘explicit’ often denote conscious and unconscious or automatic and nonautomatic processes. This is a different distinction.

self and other. As just mentioned, in such egocentric representations, *all* tokens are indiscriminately used as carrying information concerning relations to the system itself. These representations necessarily concern the self, but they are not about the self. This is why such representations cannot give rise to the self-representation (even of the minimal kind). They are, strictly speaking, selfless (Lewis, 1979; see also Recanati, 2024).

To find an instance of genuine self-representation, we need to look towards non-egocentric representations that enable certain sophisticated kinds of behaviour. In a recent experiment, Tsoar and colleagues (2011) captured fruit bats in their cave and released them in, or just outside, a crater about eighty kilometres away. The bats that found themselves within the crater, and unable to sense any of their home range's landmarks, engaged in random exploratory flights around the crater. Only once they had overcome the crater's cliffs did they head home. Bats released outside the crater did not engage in random exploratory flights and flew home immediately.

Tsoar and colleagues take the experiment to show that bats use *cognitive maps* (see Rescorla, 2017) to navigate their environment. We may think of these as representing spatial relations in a manner akin to our phones' navigation apps, with various tokens standing for the represented objects and the relations between tokens giving their spatial relations.<sup>2</sup> Importantly, such maps allow a system to represent spatial relations between non-self objects. This enables behavioural abilities unattainable with egocentric spatial representations alone: with a cognitive map, a system may, for instance, plot a novel route to a goal location that isn't currently given in perception.

And getting home along a novel route is precisely what fruit bats seem able to do. The crater was outside the bats' home range so that they couldn't have formed associative links connecting locations along a homewards trajectory. Moreover, because the bats were transported to the crater, they couldn't have used path integration to keep track of their home. The experimental design also strove to exclude the possibility that bats used magnetic, olfactory, or celestial navigation. The experiment hence suggests that bats represent the spatial relations between locations. By representing, for instance, their cave and landmarks such as hills and lights emanating from human settlements, they can compute a direct route home once they have established the distance and bearing to the landmark.

Cognitive maps are a type of non-egocentric spatial representation (Filimon, 2015; Grush, 2000; Klatzky, 1998). In a non-egocentric representation, the origin of the represented space isn't the representer itself. Instead, the origin might be some (non-self)

---

<sup>2</sup> Cognitive maps do not need to represent distances quite as straightforwardly as our everyday maps. As research on *predictive maps* argues, distances might instead represent the probabilities associated with getting from one place to another (Cothi et al., 2022; Stachenfeld et al., 2017; Stoewer et al., 2022). On such a view, token *A* being close to token *B* indicates that the system is likely to transition from location *A* to *B*. I believe such maps to be compatible with my proposal, but I cannot argue for this claim here.

object (what Grush calls *object-centered reference frames*). Here, the bat would represent the various objects as relations to the, say, the cave, so that the fruit tree is now *in front of the cave*. The represented space may also originate at a location without an object (what Grush calls *virtual points of view*). One example are maps with their top-down views of some location: all the various locations are represented as relations to some point somewhere above the terrain. What is crucial for our purposes isn't so much the location of the non-self origin of the space, but rather that non-egocentric representations explicitly relate (non-self) objects to one another. One representational token might be about the cave and another one about a distant hill, and the relation between them stands for their spatial relation. Note how these tokens do not, unlike in the egocentric case, tacitly refer to – that is, concern – the location of the representer itself. A non-egocentric representation can represent spatial relations as in principle independent of the system's location. The system's location does not need to (explicitly or tacitly) figure in the representation, and locomotion hence need not entail the need for accuracy-preserving updates.

To navigate with a non-egocentric representation, you need to know where *you* are. To plot a route home, a bat must first exit the crater and figure out how its own location relates to at least some of the represented objects. It must add to the cognitive map a token that stands for itself. When it does so, then then it instantiates a representation that explicitly differentiates between itself and the world, between its own location and the locations of other objects. This means the representation doesn't merely concern itself but is *about* itself.

However, self-representation requires more than a representation that sports a token specifying the system's location. With merely such a token, the representation may be about the system itself, but the system might be ignorant of that fact. Imagine some wily experimenter placing a mirror in front of a bat, and the bat, on seeing its own reflection, representing that bat (that is, itself) as located next to a tree. If the bat is ignorant of the fact that it's seeing itself in the mirror, it won't be able to use the represented location to, say, plot a route to the tree next to it.

What resurfaces here is a distinction familiar from the literature on *de se* thought, namely between thoughts that happen to be about oneself and thoughts where the agents knows that this is the case. It's only in the latter case that the subject thinks of itself *qua* itself. Consider this seminal example by Perry. He is in the supermarket when he notices a trail of sugar. He forms the belief that someone has a torn sack on their cart. As he follows the trail round and round the aisles, it eventually dawns on him: he is the one with the torn sack (Perry, 1979). Since we might say 'Perry believes that he has a torn sack' to describe Perry's belief both before and after his realisation, and since these two states are clearly distinct, it's become established to distinguish the second case by appending a \* to the relevant pronoun in indirect speech: Perry believes he\* has a torn sack (Castañeda, 1966).

I will introduce a similar bit of jargon in a distinction between self-tokens and

self\*-tokens, where, roughly speaking, the former involves a non-egocentric representation with a token that is merely about myself and the latter additionally involves knowledge that this is so. With a mere self-token, a non-egocentric representation has *de re* or *de dicto*<sup>3</sup> content about the self. This isn't sufficient for genuine self-representation, that is, a representation with *de se* content. Such a representation requires that the system know which token refers to itself, that is, that it instantiates a self\*-token. Since my interest is in self\*-tokens rather than mere self-tokens, and to keep the terminology as simple as possible, I will use 'self-token' to refer to self\*-tokens, except when the distinction becomes relevant again in later sections.

Knowing that some token is about oneself means being disposed to use that token in a specific sort of way when inferring information about current sensorimotor contingencies and when updating the non-egocentric representation based on sensorimotor information. In the case of a spatial representation, such sensorimotor information can be inferred from the vector whose initial point is the self\*-token and whose target point is the goal location. For instance, when the bat represents itself (self\*-token) to be such-and-such a distance away from the cave (object-token), the bat can infer the sensory input (for instance, visual impression of a dark looming hole) expected given certain motor commands (for instance, moving the wings in a certain fashion for a certain time). Similarly, if the bat has sensorimotor information regarding, say, a tree it is perceiving in front of itself, it can add a tree to its cognitive map by converting the sensorimotor information to a vector whose starting point is the self\*token and end point is the tree.

Sensorimotor representations are a kind of egocentric representation. A sensorimotor representation links motor outputs to sensory inputs, allowing a system to, for instance, calculate the motor outputs required to bring about some desired sensory input. The self doesn't explicitly figure in such representations, but the system's state *does* enter into their accuracy conditions. If we instantiated a given bat's sensorimotor representation in another bat (differently located and/or embodied), then the representation will fail to capture the links between motor output and sensory input.

Borrowing a term from Grush (2000; see also Ismael, 2008), I call a non-egocentric representation *coordinated* when a self\*-token is used to establish links to and from sensorimotor representations.<sup>4</sup> Roughly speaking, coordination is the lining up of the non-egocentric representation with the rest of the cognitive system, such that the representation entails, and is entailed by, sensorimotor information. As it is the relations between self\*-tokens and object-tokens that entail, and are entailed by, egocentrically represented information (such as sensorimotor information), coordination necessarily involves a self\*-token. Coordination entails that the system is now disposed to use the non-egocentrically represented information to guide its behaviour and to use in-

---

<sup>3</sup> The distinction between *de re* and *de dicto* content doesn't matter for our purposes.

<sup>4</sup> My use of the term 'coordination' should not be conflated with Grush's (2000). What I call coordination is only one variant of Grush's more general kind.

formation from the senses to update its non-egocentric representation.

A coordinated non-egocentric representation explicitly a system's properties, and it does so in a way that directly links them to behaviour and sensation. For instance, a bat that non-egocentrically represents itself to be close to a food source realises a representation that contains, as a minimum, tokens for itself and the food source. This bat explicitly differentiates between its own location and another object's location. Moreover, if the representation is coordinated, then the system is disposed to use relations between the self\*-token and object-tokens to infer sensorimotor information that guides its behaviour. Finally, the system is disposed to use information from the senses to update its non-egocentric representation. Hence, it appears that a system instantiating a non-egocentric representation fulfils both conditions on self-representation.

So far, I've only given a first sketch of the account. The rest of the paper fills in the gaps. Over the next few sections, I introduce instances of non-egocentric representation of temporal, bodily, and cognitive states. Thereafter, I turn to an in-depth argument why such representations are all genuine (if minimal) self-representations.

### 3. Non-egocentrically representing temporal states

A distinction akin to the one I draw between egocentric and non-egocentric representation has already been discussed in the domain of temporal cognition (Grush, 2008; Kort et al., 2005; McCormack, 2015). Moreover, the representation of temporal properties has long been seen as crucial for self-representation Goldie (2012), I must show how coordinated non-egocentric representations can represent such properties. I'll proceed in two steps: first, I look at egocentric representations of temporal properties and then argue that such representations are unable to underwrite certain more complex forms of behaviour. These instead require non-egocentric representations of temporal properties.

When food becomes scarce, slime mould amoebae merge into plasmodia that move around to forage for nutrients. Saigusa and colleagues (2008) exposed these plasmodia to multiple pulses of dry conditions, capitalising on the well-known fact that these plasmodia slow down when in a dry environment. They found that after just three exposures, slime moulds periodically slow down even if the dry condition stimulus is removed. This indicates that they have started to *anticipate* the dry conditions.

To behave anticipatorily, slime moulds need to keep track of time. It's likely that they do so with the help of regular internal oscillations (Saigusa et al., 2008) similar to those employed by a variety of other creatures (Gallistel, 1989). Slime moulds can use these oscillations to encode a *timespan* variable tracking the temporal distance to the next dry condition. Using such *interval timing* (Buhusi & Meck, 2005), a slime

mould may then slow down its forward movement when an impending dry condition is indicated by  $\langle \textit{timespan} \rangle$  approaching zero.

Interval timing and path integration exhibit striking similarities that indicate that they both employ egocentric representations. Where  $\langle \textit{timespan} \rangle$  specifies the temporal distance of some event *from the present*,  $\langle \textit{distance, bearing} \rangle$  specifies the spatial distance and bearing *from the current location and orientation*. Moreover, in both cases, the representations merely concern (and aren't about) relations *from the system's (current) actual state*. The system's actual state doesn't explicitly figure in the representations but enters as an unarticulated constituent that helps determine the representations' accuracy conditions. Moreover, because such representations *necessarily* concern relations to the system's actual state, egocentric temporal representations – just as egocentric spatial representations – need to be updated with each change in the system's state. While path integration is used to update spatial representation, egocentric temporal representations are kept accurate by means of internal oscillations that decrement  $\langle \textit{timespan} \rangle$ .

Since egocentric temporal representations necessarily represent temporal properties with a single token that merely concerns a relation to the present, there are limitations to what they can represent. For one, they 'are not carrying contents to the effect that this or that thing is happening at this or that time' (Grush, 2008, p. 156). They are unable to explicitly represent relations between moments in time and thus cannot represent that an event occurs at *some* moment in time. Consequently, egocentric representations of temporal properties cannot represent an event as happening before or after another (except tacitly, when one of the events is in the present). McCormack (2015) calls this a conception of *time as duration*, which is apt as the explicitly represented content is a duration from the present to some other moment in time.

Both Grush and McCormack contrast this form of temporal cognition with 'a more sophisticated kind' (Grush, 2008, p. 156) that does succeed at explicitly representing relations between events. A study by Arcediano and colleagues (2003) serves to illustrate this more sophisticated form of temporal cognition (see also Barnet et al., 1997; Taylor et al., 2014). They first exposed rats to two neutral stimuli (two sounds in this case) following one another with a five-second gap ( $S_2 \rightarrow S_1$ ). In a second phase, the same rats were exposed to an aversive stimulus *US* (a footshock) paired with  $S_1$  ( $US \rightarrow S_1$ ). In the test phase, the rats were exposed to  $S_2$ . The rats displayed decreased use of a lick bottle following this exposure, demonstrating a fear response.

Note how the rats seemed to anticipate the footshock on exposure to  $S_2$  even though they had not been conditioned on the relevant sequence ( $S_2 \rightarrow US$ ). The study's authors assert that the rats' capacity to anticipate without without being exposed to the relevant sequence of events shows that rats can integrate the two learnt temporal sequences to arrive at a sequence  $S_2 \rightarrow US \rightarrow S_1$ , leading the rats to expect *US* on being exposed to  $S_2$ . Arcediano and colleagues think that such integration requires a kind

of cognitive map already suggested by Tolman (1948), a temporal map.

Here's why we might think that egocentric representations cannot explain this behaviour and why non-egocentric representation are needed. Egocentric representations of temporal properties could be used to measure the time that passes between the events taking place in the first two phases of the experiment. In the first phase, a representation  $R1_{ego}$  could be instantiated to measure the duration between  $S2$  and  $S1$ . In the second phase, another representation  $R2_{ego}$  would measure the duration between  $US$  and  $S1$ . Both representations only tacitly represent the fact that they are relations *between* certain events. More specifically, the fact that these representations are about temporal relations is tacit in their activation at  $S2$  and  $US$ , respectively, and their subsequent employment in behaviour that is anticipatory of  $S1$ . That's why, even though  $R1_{ego}$  and  $R2_{ego}$  carry explicit content about a relation to  $S1$ , the rats cannot, say, subtract the value in  $R2_{ego}$  from  $R1_{ego}$  to arrive at the temporal distance between  $S2$  and  $US$ . Or, more precisely, while they *could* subtract one value from the other, they would not thereby have any indication of how to use the resulting representation; they would not, unlike in the case of  $R1_{ego}$  and  $R2_{ego}$  have formed a non-representational association between a certain event (here,  $S2$ ) and activation of the token.<sup>5</sup>

To integrate the two temporal sequences, rats must represent them non-egocentrically. When their representations  $R1_{non-ego}$  (about  $S2 \rightarrow S1$ ) and  $R2_{non-ego}$  (about  $US \rightarrow S1$ ) explicitly represent relations between events, they may integrate representations that share an element (Gür et al., 2018; Molet et al., 2012; Taylor et al., 2014). In the experiment, rats first learn that  $S2$  precedes  $S1$ . Then, in the second phase, they learn that  $US$  precedes  $S1$ . Because  $S2$  doesn't occur between  $US$  and  $S1$ , they seem to infer that  $S2$  occurs before  $US$ , arriving at  $S2 \rightarrow US \rightarrow S1$ .

In a non-egocentric representation, time functions as a framework (McCormack, 2015) that organises the temporal relations between events. These relations do not tacitly refer to the system's actual state, the present. The absence of such reference is evidenced by the fact that the rats' representation of  $S2 \rightarrow US \rightarrow S1$  may remain accurate as time passes even if it isn't updated. In short, temporal relations are represented in a way that is in principle independent of the system's actual state.

Since non-egocentric temporal representations do not necessarily concern the system's own state, coordination becomes an issue. Just as a creature may sometimes be ignorant about where it is, it may also be ignorant about *when* it is. The case of Arcedi and colleagues' rats illustrates this: Their representation of  $S2 \rightarrow US \rightarrow S1$  may be stored in memory without the rat knowing how those events temporally relate to the present. It may only be at, say, the occurrence of  $S2$  that the rat is able to infer

---

<sup>5</sup> We might attempt a different explanation: As the rats associate  $S1$  with the footshock ( $US$ ) in the second phase of the experiment,  $S1$  takes on a negative valence. Later on, when the rats are exposed to  $S2$  in the test phase, it's the expectation of  $S1$  rather than  $US$  that causes the fear response. Arcedi and colleagues (2003) are aware of this possibility and implemented an alternative version of the experiment to exclude it. I refer the interested reader to their paper for more information.

that *it itself* is now located at the same moment in time as  $S_2$ . It can then infer sensorimotor information from the non-egocentric representation and behave in a manner anticipatory of  $US$ , suppressing the use of the lick bottle.

Note how, just as in the case of spatial cognition, coordinated non-egocentric representation of one's temporal properties seems to be linked to self-representation. A creature with such a non-egocentric representation will explicitly differentiate between its own temporal properties and those of events in the past and future. Such a representation is about the system itself. Moreover, through coordination, the system comes to use a certain token in a special way, namely as a self\*-token that specifies the system's actual temporal state. The system is disposed to infer sensorimotor information from relations between the self\*-token and other tokens and to update the representation based on sensorimotor information. If this is right, a creature instantiating non-egocentric temporal representations is a genuine self-representer.

#### 4. Non-egocentrically representing bodily states

Many animals represent their own bodies. These body representations develop early in ontogenesis (Meltzoff et al., 2019; Slaughter & Brownell, 2011) and are said to prepare the ground for later representations of the body-external environment (Nguyen et al., 2020; Stoytchev, 2009). Some of these representations are non-egocentric and suggest a link to minimal forms of self-representation. Unlike in previous sections, and in the interest of brevity, I will directly turn to non-egocentric representations without contrasting them with their egocentric brethren.

Human infants begin to successfully reach for objects at around three to four months of age. At first, infants reach jerkily, inaccurately, and with movements that do not describe a straight line (Bruner & Koslowski, 1972; Hofsten, 1982; Thelen et al., 1996). Such movements expend more energy than would be required to reach the target state (Neil E. Berthier & Keen, 2006; Konczak et al., 1995; Rolf & Steill, 2012). Starting at around five months of age, infants anticipatorily adjust their hands to the shape of the object to be grasped at the end of the reach (Hofsten & Fazel-Zandy, 1984; Wentworth et al., 2000; Witherington, 2005), and by around seven months of age, infants can reach for objects efficiently and effectively (Hofsten, 1991; Thelen et al., 1996). They are now, for instance, able to drink from an open cup (Hofsten & Fazel-Zandy, 1984). At about two years, smoothness of motion approaches adult abilities (Neil E. Berthier & Keen, 2006).

Originally, researchers explained the early (jerky, inaccurate, zig-zag) movements with the supposed vision-guided nature of infant reaching (for overviews of that literature, see Neil E. Berthier (1996), Corbetta et al. (2018)). It was believed that infants would pause their reaching behaviour at various points along the trajectory in order

to check and correct their motions by gazing back and forth between the hand and the intended target. Only later would infants develop the sort of embodied sense of their own body that enables smooth bodily movements.

In light of a number of findings, this view has been largely abandoned. Von Hofsten and Lindhagen (1979) observed that infants do not shift their gaze back and forth between target and hand position but rather fixate on the object. Even more convincingly, Clifton and colleagues (1993) showed that infants can reach objects successfully even when they cannot see their own hands and arms. These findings indicate that infants need to *already* have an embodied sense of their hand position when they begin to engage in reaching behaviour (Corbetta et al., 2014).

Converging evidence from developmental psychology and robotics (Baranes & Oudeyer, 2013; Castellini et al., 2007; Desmurget & Grafton, 2000; Rolf & Steill, 2012; Schillaci et al., 2016; Takemura et al., 2018) indicates that infants' developmental trajectories and emergent abilities can be explained by internal body models: *forward models* and *inverse models*. Forward models (Grush, 2004; Körding & Wolpert, 2006; Wolpert & Ghahramani, 2000) allow systems to compute the bodily (and sensory) state to which a motor command will likely give rise. As forward models represent the bodily dynamics internally, they can be used to estimate the system's state even when sensory input is noisy. Moreover, because forward models are realised in the brain, their predictions are available long before proprioceptive signals have finished their trip back from the sensory surfaces. With a forward model, an infant could, say, predict that the motor commands it just issued will cause it to fail to reach a toy in front of it. Importantly, the infant could do so in the absence of – or before – visual sensory input that confirms the failure of the reaching movement. Inverse models are, as the name implies, the inverse of forward models. They allow computing the motor commands required to reach a target state and are hence implicated in goal-directed behaviour. An infant could employ her inverse model to compute the motor commands necessary to reach the toy without having to constantly visually align her hand with the target.

Forward and inverse models need information about the system's current bodily state to compute their outputs. For instance, depending on whether her arm is flat against the side or stretched out in front of her, an infant's inverse model should issue different predictions about the motor commands necessary to reach a toy. And the forward model, too, will predict different bodily states depending on the infant's current posture. Hence, the forward model relates possible bodily states in terms that may be inferred from motor commands, whereas the inverse model relates possible bodily states in terms that allow inferring the motor commands required to transition between the states.<sup>6</sup>

---

<sup>6</sup> Forward and inverse models only explicitly differentiate between possible bodily states if they are *articulated* (see Grush, 2004). This is likely to be the case. Grush references research showing correlations between certain groups of neurons and, for instance, elbow angle. Wolpert and Ghahramani note

Internal body models are non-egocentric representations. Not only do internal body models explicitly differentiate between various possible states, the relations that these bodily states are represented to bear on one another are in principle independent of the system's actual posture. When such a representation contains a token giving the system's actual state, it is hereby *about* the system's actual state.

To use such a model, a system needs to use one of the body state tokens in a special way, that is, it needs to use it as specifying its own actual state. Or, to put it in Coslett and colleagues' (2008) words, 'you are here' information is 'required in order to accurately reach toward an object' (p. 117). In other words, an internal body model needs to be coordinated to be used. This means, first, that the system must be disposed to infer sensorimotor information from relations from the self\*-token to tokens that represent merely possible bodily states. And it means, second, that the system must be disposed to infer possible bodily states from the relations between the self\*-token that correspond to sensorimotor information.

Coordinated non-egocentric representation of bodily states suggests a form of self-representation. An inverse model non-egocentrically represent bodily properties by representing the system as exemplifying a certain posture that is explicitly differentiated from other postures the system could exemplify if it were to issue certain motor commands. It is about the system's actual state (and doesn't merely concern it). Reminiscent of certain non-egocentric temporal representations, the *other* from which the system is explicitly differentiated here is the system's own merely possible state.<sup>7</sup>

## 5. Non-egocentrically representing cognitive states

Representations of our cognitive and mental properties have often been implicated in accounts of self-representation. We routinely represent ourselves to have certain beliefs, desires, and other mental states, and this is important at least for typical human-level self-representation. In addition, some authors even think that representing one's mental states is necessary for self-representation (see, for instance, Musholt, 2013, 2015). While this paper argues that the second claim is too strong – coordinated non-egocentric representation of any properties is self-representation – this doesn't dimin-

---

that the human muscular system comprises over 600 muscles and that the resulting complexity 'clearly prohibits a simple look-up table' (Wolpert & Ghahramani, 2000, p. 1212).

<sup>7</sup> The fact that mirror neurons may cause babies' (and adults') body representations to also activate when they watch another person reach or grasp complicates the picture (Lepage & Théoret, 2007; Marshall & Meltzoff, 2015). Represented bodily states are then only sometimes the infant's own, and this has been argued to underlie their ability to imitate others. The idea is, roughly speaking, that infants automatically mimic others because they (sometimes) fail to differentiate between perceived actions and their own target states. However, note that the infant still needs to use some token as a self\*-token and that this token cannot be what is confused with the perceived state. Observing the other would not entrain motor behaviour if the infant confused the other's state with her own since she would represent herself to already embody the target state.

ish the importance of showing that non-egocentric representations can underwrite self-ascriptions of mental and cognitive properties. As I'll show below, this is indeed the case and such representations are even instantiated by certain non-human animals.

Corvids, and in particular scrub-jays, have recently been the subjects of a slew of experiments that highlight their sophisticated forms of social cognition (Baciadonna et al., 2022; for instance, Bugnyar et al., 2016; Kort et al., 2005, 2005). Many of these studies make clever use of scrub-jays' predilection for caching food and pilfering conspecifics' caches. One such experiment by Clayton et al. (2007; see also Nathan J. Emery & Clayton, 2008) studied how cachers adapt their behaviour when a conspecific is present during cache recovery. At a first caching event, a cacher is observed by a conspecific (observer A) who looks on through a transparent partition. On the cacher's side, two trays – visible to both birds – constitute potential caching locations, though one of the trays (tray B) is inaccessible due to a transparent obstacle. This forces the cacher to hide the food in tray A. At the second caching event, the same cacher is observed by a different conspecific (observer B). This time, tray A is inaccessible, and the cacher hides food in tray B.

At recovery, the cacher behaves differently if observer A rather than B is present. For instance, when observer A is present, cachers tend to re-cache the food from tray A under tray B, often moving it back and forth several times first. They will not recover the food from tray B. Presumably, this behaviour aims at better protecting the food initially stored under tray A (that the observer saw being cached) while not giving away the location of the food under tray B (that the observer did not see being cached). When observer B is present, the behaviour is analogous but targets the other tray.<sup>8</sup>

The study's authors believe the experiment to reveal that corvids employ representations of conspecifics' mental states, though others think that such conclusions are premature (Derek C. Penn et al., 2008; Derek C. Penn & Povinelli, 2007). The experiment fails to show, they say, that corvids exhibit 'a sensitivity to what others have and have not seen' (Clayton et al., 2007, p. 519) or that they engage in 'knowledge attribution' (ibid.). The caching bird could, for instance, simply represent which observer had been present at which caching event. At recovery, the cacher could then re-cache food it hid while this particular observer had been present. Scrub-jay behaviour could then be explained by positing that corvids represent spatial and temporal properties, and not conspecifics' cognitive states.

However, the fact that only birds that have pilfered others' caches engage in cache protection suggests the need for an alternative explanation (N. J. Emery & Clayton, 2001). Corvids seem to infer others' behavioural patterns from their own, and this requires that they distinguish between their own informational states and others' informational states. Only then can the cacher differentiate between where it thinks the

---

<sup>8</sup> Clayton and colleagues (2007) went to considerable lengths to exclude the possibility that cache protection behaviour is cued by the observer's behaviour. More details can be found in their paper.

cache is located and where the conspecific thinks the cache is located. Moreover, the cacher must conceive of the informational states attributed to itself and the conspecific as being of the same kind. Only then can the scrub-jay be sensitive to the fact that these states determine behaviour in the same sort of way. In short, scrub-jays need to be able to represent that conspecifics may be in informational states of the same *kind* but differing in *value* from their own.

The spatial representations discussed in earlier sections were always tacitly about how the *system itself* estimated the world to be – as we just saw, this isn't sufficient to explain corvid caching behaviour. Butterfill and Apperly (2013) argue that scrub-jays may distinguish between their own and others' representations of spatial properties without representing mental states. Instead, they represent *registrations*. These are representations of others' relations to certain objects: 'an individual registers an object at a location if and only if she most recently encountered it at that location' (Butterfill & Apperly, 2013, p. 617). Like beliefs, registrations motivate behaviour and may be incorrect (when the object's no longer at the location at which it was registered). Representations of registrations can be used to predict others' behaviour across a wide (though limited<sup>9</sup>) range of situations. Seeing that some conspecific has registered food at a certain cache, a scrub-jay could thus infer that the conspecific will attempt to pilfer it.

Scrub-jays, on this proposal represent registrations by encoding several  $\langle \textit{individual}, \textit{location}, \textit{object} \rangle$  tuples. At cache recovery, a cacher first selects the registration (or registrations) whose *individual* corresponds to the conspecific currently present and whose *location* is the current location. The cacher then attempts to make it so that the *location* and *object* elements of the registration, which indicate the conspecific's registration of the *object* at the *location*, do not correspond to its own representation of cache location. For instance, if a scrub-jay represents the food to be located under tray A while representing the conspecific to have registered the food under that same tray, it will remove the food from the cache. It then represents the food to be at a location different from where the conspecific registered it.

We still need to shed light on how a scrub-jay compares a conspecific's registrations with its own representation of cache locations. After all, and as I've mentioned above, for the cacher to infer the observer's behaviour from its own, the informational states involved in the two cases need to be comparable. However, registrations aren't of the same kind as the scrub-jay's own representations of spatial properties and therefore cannot be directly compared to these. Hence, the scrub-jay needs to first translate its representation of spatial properties into  $\langle \textit{individual}, \textit{location}, \textit{object} \rangle$  tuples where the *individual* is it itself. It needs to instantiate a representational token that refers to itself and with which it attributes to itself certain registrations. These registrations can then be compared to those of conspecifics. How this is done doesn't matter for our pur-

---

<sup>9</sup> Registrations do not exemplify beliefs' full functional profile (Butterfill & Apperly, 2013). These details are not relevant to my argument.

poses; what matters is that translation is required and that it involves self-attributing registrations.

The resulting representation should be familiar from earlier sections: it sports tokens standing for objects (individuals) and their properties (registrations), and to employ the representation, the system must use some token as specifying its own state. This structure evidences the hallmarks of coordinated non-egocentric representation: object properties are represented as in principle independent of the system's own state, and use of the representation requires coordination.

As with non-egocentric representations of other kinds of properties, there exist tantalising connections to self-representation. A scrub-jay seems to explicitly differentiate between where *it itself* has encountered objects and where conspecifics have encountered them. The representational token that specifies its own registrations is *about* itself. Relations between its own registration and those of others imply certain sensorimotor information as can be seen in scrub-jays' flexible caching behaviour. Such non-egocentric representation therefore seem to imply a form of self-representation – scrub-jays seem to represent themselves (and others) as exemplifying cognitive states.

Note that even if it turns out to be true that scrub-jays self-represent some of their *cognitive* properties, they don't thereby represent *mental* properties.<sup>10</sup> Thus, authors such as Musholt (2013, 2015), who believe that representing one's mental states is required for self-representation would claim that corvids do not self-represent. I believe that this is false: any and all non-egocentric representation is self-representation. I now turn to giving substance to this claim.

## 6. Minimal self-representation

So far, I've argued that coordinated non-egocentric representations are used by a variety of creatures to represent a diverse set of properties. I've also hinted at connections between this kind of representation and *self*-representation. It's now time for me to explicitly argue why coordinated non-egocentric representation is minimal self-representation. If this is right, then any creature that non-egocentrically represents a property is a self-representer.

As mentioned before, an account of self-representation must, first, be an account regarding representations that are *about* the self – rather than an account regarding representations that merely *concern* the self. In the account that I propose, this distinction is captured by the differentiation between egocentric representations and non-egocentric representations.

---

<sup>10</sup> I understand mental properties to be those we know from folk psychology: beliefs, desires, and so forth. Cognitive properties include mental properties but encompass also other states that go between perception and action. Registrations are one example here.

Egocentric representations necessarily – and exclusively – represent the self tacitly. Hence, they necessarily concern the self and are never about the self. An ant that represents the distance and bearing to, say, its nest and some food source, instantiates two representational tokens whose  $\langle \textit{distance}, \textit{bearing} \rangle$  tuple necessarily concerns relations to the ant’s actual state (in particular, its location and orientation). Such a representation doesn’t contain a token that stands for the ant and, hence, isn’t about the ant. The matter is no different with egocentric representations of other properties. As we’ve seen, slime moulds encode  $\langle \textit{timespan} \rangle$  variables, which necessarily concern the temporal distance of some event to the slime mould’s actual state (here, the temporal present). The representation doesn’t contain a token that stands for the slime mould and, hence, isn’t about the slime mould.

Egocentric representations’ accuracy conditions necessarily involve the state of the representer itself. If another ant, at a different location, were to token the relevant representation, the representation would fail to capture the distance and bearing to the nest and food source. If another slime mould encoded a  $\langle \textit{timespan} \rangle$  variable with the same value, it would fail to capture the temporal relation to the represented event. Hence, although egocentric representations don’t contain a token that stands for the self, the self’s state is part of the accuracy conditions: it’s an *unarticulated constituent* of the representation (Perry & Blackburn, 1986). The format of egocentric representations is, as mentioned before, selfless but necessarily self-concerned Recanati (2024).

In contrast, object-tokens in a non-egocentric representations do not involve the self as an unarticulated constituent. We can see this in how object-tokens may represent objects and their properties as in principle independent of the system’s own state. According to a fruit bat’s cognitive map, it’s possible to fly from, say, the crater to the cave without having to update the tokens that stand for these locations. A scrub-jay may update its representation of cache locations (or, more precisely, its representation of its own registrations) without thereby needing to update its representation of others’ registrations.

Of course, non-egocentric representations can, and often do, represent the system’s own state. They do so by means of object-tokens that stand for system itself, which I call (mere) self-tokens. By means of such a token, the fruit bat may represent its location and the scrub-jay its registrations. In such representations, the self becomes an articulated constituent. Since this representation contains tokens for both the self and the other, it instantiates an explicit self-other differentiation. Note that in some of the cases I’ve discussed, such as internal body models, the self-token is the token that refers to the *actual* self, whereas the object-tokens refer to possible selves. This complication doesn’t change anything of substance, in either case we now have a representation that is *about* the (actual) self.

Since non-egocentric representations represent objects as in principle independent of the system’s own state, such representations’ usefulness is sharply limited, and this applies even for representations with a (mere) self-token. A fruit bat may have a

maximally detailed cognitive map that includes a token that stands its own location – but if it doesn't know what token stands for its\* location, it cannot use the representation for locomotion. Similarly, a rat representing a sequence of events may instantiate a token that stands for the present, but so long as it doesn't know what token is about its\* temporal state, it will not be able to use the representation to behave anticipatorily.<sup>11</sup>

To make its non-egocentric representations useful, a creature must know what representational token is about its\* state. Just like Perry needed to realise that he\* is the one with the torn sack of sugar to adjust the sack on his cart (1979) – or Mach that he\* is the dishevelled pedagogue to, say, do something about his look (1890) – the various creatures I've discussed must realise what representational token is about their\* state to employ these representations. According to my proposal, this realisation is the coordination of the non-egocentric representation with sensorimotor information. Coordination establishes systematic non-representational – or architectural (Ismael, 2008) – relations between these two types of representation so that the creature is disposed to infer sensorimotor information (in an egocentric format) from non-egocentric representation and vice versa.

Such coordination necessarily results in the instantiation of a special sort of token, the self\*-token. A self\*-token is, as mentioned before, distinct from a self-token. Self\*-tokens aren't individuated by the object for which they stand, but rather by how the system *employs* them. In short, they are employed by the rest of the cognitive system as being about the system itself. Think of spreadsheet software, which shows a view of data arranged in cells organised in rows and columns. When you move the select one of these cells, it gets highlighted. If you now enter some command, it affects only the cell that is selected. The highlighting indicates how the software is linked to the data, it indicates what piece of data it will operate on. The highlighting doesn't add any information to the spreadsheet, it merely establishes a link between the spreadsheet and the rest of the software. Similarly, self\*-tokens do not add information to non-egocentric representations, but rather indicate which (pre-existing) token the system is to use in a special way.

That special way of using a self\*-token consists in using it in inferences to and from sensorimotor information. I first look at how the self\*-token is involved in infer-

---

<sup>11</sup> Note that even uncoordinated non-egocentric representations can be used in certain situations. Imagine that fruit bats are able to tell one another the location of food sources. A bat could then tell a conspecific that some fruit tree is closer to the cave than another. To do so, the bat doesn't need to know its own location. However, while the contribution of the non-egocentric representation may be important, even necessary, it isn't be sufficient. The bat cannot infer the relevant sensorimotor contingencies from the uncoordinated non-egocentric representation alone but needs additional sensorimotor information about the communicative situation. For instance, if it knows that it can expect being treated favourably in the future if it helps the conspecific, it could then use the non-egocentric representation to figure out how to best help. It's sufficient for our purposes that coordination *is* required in many important cases.

ences from non-egocentric representation to sensorimotor information. Sensorimotor information is, as I've mentioned before, a kind of egocentric representation and, hence, necessarily self-concerning. In contrast, non-egocentric representation represents objects as in principle independent of the system's own state. Hence, for a non-egocentric representation to imply sensorimotor information, one of its tokens must be used as denoting the system's actual state. A fruit bat needs to represent not only the location of its cave but also its\* own location (by means of a self\*-token) to be able to infer the motor commands required to reach home. It's only when a rat uses one of the represented moments in time as specifying its\* own temporal state that the rat may use the representation to anticipate future events. And it's only when an infant uses one of the many bodily states represented by an internal model as specifying its\* actual bodily state that an inverse model can be used to infer the motor commands required to reach some goal state.

The links in the opposite direction can be used to update the non-egocentric representation with information in a sensorimotor format. On having gained sensorimotor information about the motor commands required to, say, lift her arm, an infant may want to use this information to enrich her internal model. A rat with sensorimotor information about some future event can only encode that information in its non-egocentric representation if it knows which representational token is about its\* temporal state (the present). A non-egocentric token's value can only be inferred on the basis of sensorimotor information *and* the self\*-token. Only then can the rat establish a correspondence between the sensorimotor information and a *relation* from its\* temporal state (that is, the self\*-token) to the event.

The above discussion makes it obvious that there's no hope in reducing self\*-tokens to mere self-tokens – mirroring the fact that certain indexicals seem to be essential. If we want to explain why some scrub-jay re-cached the food under some tray, we must make reference to a self\*-token. No matter what non-egocentrically represented information we give the bird (in the form of object-tokens, even including mere self-tokens), it cannot thereby infer sensorimotor information. Non-egocentric representations represent objects as in principle independent of the system's state, and sensorimotor information cannot be inferred without a self\*-token. It's only when the representation is coordinated, and a token is specified as self\*-token, that the cognitive system knows the origin of the non-egocentrically represented relation that implies sensorimotor information. Similarly, after Perry (1979) and Castañeda (1966), many authors (Babb, 2016; Bermúdez, 2017; for instance, Lewis, 1979; Stalnaker, 1981) argue that certain indexicals cannot be reduced to non-indexical terms.<sup>12</sup> It is impossible, they say, to explain why Perry adjusted the torn sack in his cart without saying that this is so because he realised that he\* is the one with the torn sack. When we replace 'he\*' with non-indexical descriptions – maybe 'the only bearded man in

---

<sup>12</sup> See Cappelen and Dever (2014) for a dissenting opinion.

the store' – this doesn't explain the change in behaviour if Perry doesn't know that he\* is that bearded man. No matter the description, we always also need Perry to realise that he\* is the person described.

Uncoordinated non-egocentric representation only contain *de re* or *de dicto* content, and that is never sufficient to infer *de se* content. No matter what (non-indexical) information we give the scrub-jay about its own registrations – or Perry about the person with the torn sack of sugar – this cannot enable identifying a self\*-token. We might think that the scrub-jay has some private information regarding its own registration, and it's using that to select some object token as a self\*-token. But speaking of knowledge of registrations is here simply a roundabout way of saying that the system has a self\*-token whose value indicates those registrations. We are then driven to ask on what basis the bird selected *that* self\*-token, at which point it should become clear that we're stuck in an infinite regress. We must conclude, as many others have also (Evans, 1982; Fuchs, 2017; MacKenzie, 2010; Peacocke, 2014; Shoemaker, 1968), that a system cannot select a certain token to be used as a self\*-token by identifying it with another of the representation's tokens.

Since properties are represented as in principle independent of the system's state, a system needs to go *beyond* what's contained in the non-egocentric representation to infer the value of the self\*-token. A system does so by establishing correspondences – making identity judgements – between sensorimotor information and *relations* in non-egocentric representations. For instance, an infant may possess sensorimotor information about how various motor commands would lead to body movements, and it can use this to search its internal body model for a self\*-token whose value implies this very sensorimotor information by its relation to various object-tokens. After all, sensorimotor information can be inferred from relations between a self\*-token and object-tokens, and so it stands to reason that by doing the reverse, a system can infer a self\*-token. Importantly, this only involves identity judgements between relations in a non-egocentric representation and egocentrically represented information – and not between two tokens of the non-egocentric representation, and therefore it's not subject to the above infinite regress.

The establishment of coordination is possible since sensorimotor information necessarily concerns the self. Hence, the origin of the relation in the non-egocentric representation that is inferred from sensorimotor information is an object token that necessarily refers to the self. This token is the articulation of the unarticulated component of the sensorimotor representation, and that unarticulated component is necessarily the self. Hence, there is here no possibility that this token attributes properties to an object other than the self, and because the possibility misattribution is excluded, the system can immediately designate that token as a self\*-token.

Because attributing properties to the wrong object is impossible when inferring a self-token from sensorimotor information, self\*-tokens inferred from sensorimotor information are immune to errors of misidentification (IEM) in the same way as certain

uses of the indexical 'I' (Evans, 1982; Perry, 2010; Shoemaker, 1968; Wittgenstein, 2007). When thinking 'I have a headache' or 'I see a tree', it's impossible to be mistaken regarding the person to whom the property is ascribed. If I see a tree, it makes no sense for me to wonder, 'Someone is seeing a tree, but is it "I"?' I might be mistaken about the content of the visual experience, but the fact that *I* am having the experience seems beyond doubt. If the above argument is on the right track, self\*-tokens are similar: while they may misrepresent a system's properties, they necessarily refer to the system itself. This is so because a coordinated non-egocentric representation 'allows for the possibility that it misrepresents the property that is being ascribed, while it cannot misrepresent the subject purportedly possessing that property' (Musholt, 2013, sec 2.3). A self\*-token may misrepresent the subject's properties, for instance when a bat infers its\* location based on mistaken sensorimotor information. However, it cannot misrefer because sensorimotor information is necessarily self-concerned.

My account of IEM shares some similarities with one that Recanati (Recanati, 2009, 2012, 2024) has recently advanced. Recanati, like me, attempts to give an account of why certain self-ascriptions do not require judgements of identity and hence goes beyond the *Simple View* that merely asserts *that* certain self-ascriptions do not require judgements of identity. He argues that certain self-ascriptions are based on experiences with a specific of mode. The mode of an experience is related to its functional role and is transparently accessible to the agent. So, the mode of a perceptual experience is different from the mode of an interoceptive experience, and that difference is due to differences in functional role and is experienced by the agent. He then argues that certain experiences' function is to provide information about the agent herself, and he calls these experiences of the *internal mode*. Because the role of such experiences is to always and only to provide information about the agent herself, '[t]he person in question is, as it were, pre-identified [as the system itself], being determined by the mode of the experiential state' (Recanati, 2024, p. 9). And because of this pre-determination, IEM is impossible. Recanati now still needs to explain why certain judgements based on experiences of the *external mode* are also IEM. This, he says, is because such experiences are also bound to be about the subject itself. When an agent sees the Eiffel Tower, then the Eiffel Tower is necessarily something that is present *in the subject's environment*. And hence, here too there cannot be any IEM.

I share with Recanati the conviction that IEM is due to a lack of the need for identification, and I also connect this lack to the fact that the basis on which IEM self-ascriptions are made is necessarily self-concerned. However, I don't believe IEM has anything to do with the mode of experiences. Indeed, it's rather awkward for Recanati that he begins by saying that experiences of the internal mode are IEM because their role is to provide information about the subject, only to then need a different kind of explanation to account for the fact that experiences of the external mode can also underlie self-ascriptions that are IEM. According to my view, one mechanism underlies IEM across all cases: self-ascriptions are IEM when they involve the articulation of

the unarticulated subject component of an egocentric representation. It has nothing to do with the fact that certain informational streams' role is to provide information about oneself. Proprioception *and* perception have an egocentric format, are necessarily self-concerning, and therefore are IEM.

My account isn't only more parsimonious than Recanati's, it also allows us to respond to some of the challenges which have been directed at his view. First, as Coliva and Palmira (2024) write, 'it's controversial to maintain that the content of one's proprioceptive experiences is subjectless' (p. 22). My view doesn't say anything about the nature of subjects' experiences, and it's hence possible, especially for subjects that do instantiate non-egocentric representations, that proprioceptive experiences are coloured, as it were, by their explicit ascription to themselves of certain properties. My account says that certain self-ascriptions are IEM because they originate from a transformation of egocentric to non-egocentric representations, and this claim only depends on the fact that the various sensory modalities provide information in an egocentric format.

A second objection concerns the case of explicit *de se* judgements, 'namely *de se* judgements whereby one makes the contrast between oneself and another person relevant' (Coliva & Palmira, 2024, pp. 22–23). Recanati argues that such judgements are IEM because of a process of *reflection* (2012), which is similar to how I show that coordination can issue in the articulation of sensorimotor informations' unarticulated subject component. Coliva and Palmira now argue that Recanati misses a hidden background assumption, which shows that the basis of explicit *de se* judgements isn't necessarily self-concerned. My proprioceptive experience of legs being crossed is only about myself 'if the background presupposition *that the person whose body is causing such and such proprioceptive experiences is myself* is not challenged' (Coliva & Palmira, 2024, p. 23, emphasis in the original). Now, Recanati's account seems to indeed be subject to this challenge since he at least implies that IEM is due to experiences of the internal mode necessarily supplying the self *as an object*. According to my view, this definitely isn't the case: IEM is due to the fact that sensorimotor information (of any kind) necessarily supplies the self as an unarticulated component. Even in the case of proprioceptive information, this is what drives IEM. If proprioception issues in a proprioceptive information to the extent that a certain motor output will lead to a certain sensory input, then the fact that this concerns *my* motor output and sensory input is beyond doubt *even if* the information comes from another body. It concerns me by dint of the fact that this how the representation is built into my cognitive system.

Only *some* self-ascriptions are IEM, and I still need to account for those that aren't. Think, for instance, of an animal that passes the mirror self-recognition test (or one of its variations). Such a creature looks in the mirror, sees an object reflected in it, and recognises that it\* is this object. It judges the creature in the mirror to be identical with itself, and consequently self-attributes the property of having, say, a spot of rouge on

the forehead. Such a self-attribution consists in instantiating a self\*-token specifying the property of having rouge on the forehead. Since this representation has a self\*-token, it is linked, in the manner discussed above, to sensorimotor information. The animal may, for instance, infer that if it moves its limbs in a certain way, it can bring the spot to disappear. However, unlike in previous examples, it's certainly possible for the creature to be mistaken regarding the object which has a spot on the forehead. The animal may, for instance, be standing next to a conspecific, and it's this conspecific that sports a spot of rouge on its forehead.

Whether a non-egocentrically represented property is IEM depends on the source of the information underlying the self-ascription (Evans, 1982). If, after an accident, I (visually) perceive a broken arm and form the belief that my arm is broken, this belief isn't IEM. It's possible that the arm I'm seeing isn't my arm and, hence, that the person with the broken arm isn't me. However, if I form the belief that my arm is broken on the basis of proprioceptive information, then it *is* IEM. Note that this distinction is sometimes made in the terms used by Wittgenstein (2007), who spoke of the use of 'I' as subject and the use of 'I' as object, where only the latter is IEM. But this risks hiding the fact that the distinction isn't so much about uses of 'I', but rather about the basis on which properties have been self-ascribed.

Earlier, I said that certain self-ascribed properties are IEM because they are based on egocentric information – but while that's true, it cannot be the full picture. After all, when I represent myself to have a broken arm on the basis of perceptual information, then that representation is also based on egocentric information. Rather, a self\*-token is IEM when its value stems from the egocentric representation's unarticulated (tacit) subject component, and it *isn't* IEM when it stems from the egocentric representation's articulated (explicit) object component. Hence, when I represent my arm to be broken based on sensorimotor information to the effect that arm movements of various sorts will lead to experiences of pain, then the self-ascription is IEM. In contrast, when I represent my arm to be broken because of sensorimotor information to the effect that if I move my head this and that way, I will see the arm from this and that angle (which indicates that it is my arm), then the self-ascription isn't IEM. In this latter case, I first instantiate an object-token that indicates that someone has a broken arm and then judge that the this object-token refers to the same object as the self\*-token. This judgement may be false as it's not guaranteed that the object-token refers to the self. This is why when such an identity judgement leads to the self-attribution of some property, this self-attribution isn't IEM even though it's linked up to sensorimotor information in the requisite way that makes it a genuine self-representation.

Note that it must be the case that at least *some* self\*-token values derive from the articulation of an egocentric representation's unarticulated subject component. Otherwise we cannot escape the infinite regress discussed above. A judgement of identity between a self\*-token and some object-token evidently requires a self\*-token. Now, if this self\*-token's value is itself derived from a judgement of identity with another

self\*-token, then that evidently requires an additional self\*-token. The infinite regress is only stopped when we arrive at a self\*-token whose value derives from the unarticulated subject component of a sensorimotor representation. Hence, coordinated non-egocentric representation requires that at least some self-attributed properties are IEM.

I have argued that non-egocentric representations explicitly differentiate between self and other and that coordination establishes links between such a representation and sensorimotor information. These links mean, first, that a creature with a coordinated non-egocentric representation is disposed to take the represented information as directly relevant for its behaviour. Second, the link entails that the creature is disposed to update the representation based on sensorimotor information supplied by the senses. When sensorimotor information is used to determine the value of a self\*-token, then this token is immune to errors of misidentification. If, in contrast, the self\*-token comes to have a certain value because of an identification with some object-token, then the self-ascribed properties aren't immune to misidentification. Given these characteristics, a coordinated non-egocentric representation is a genuine (if minimal) form of self-representation. Any creature representing any property with a coordinated non-egocentric representation is a self-representer. Minimal self-representation does not depend on the representation of specific kinds of properties, and as the examples throughout this paper have shown, minimal self-representation is rather more common than we might have thought. Scrub-jays, fruit bats, very young infants, and rats all self-represent some of their properties.

## 7. Substantive self-representation

What makes minimal self-representations relatively common is that they are, as the name implies, minimal; they don't demand that self-representation are about specific kinds of property. We might worry that minimal self-representation says a little too little about what is important about how, say, adult neurotypical human beings self-represent. Here, I want to make some cursory remarks about how minimal self-representation relates to such, more substantive, notions of self-representation.

Musholt believes that self-representation requires 'acquisition and application of the first person concept' (Musholt, 2013, sec. 4). This in turn, she believes, requires representing one's own and others' *mental states* (Musholt, 2012). This is a substantive view of self-representation: it argues that the representation of a specific kind of property – here, mental properties – is necessary for self-representation.

Many other kinds of property have been advanced as being necessary (or at least important in some privileged sense) for self-representation. Grush (2000), taking inspiration from Strawson (2011), writes that 'the subject/object distinction is the result of a cognizer's representation of space' (p. 62). Peacocke (2014) and Campbell (1999)

think that spatial properties are a good start, but that we additionally need temporal properties. The necessity of representing one's temporal properties is also one of the few things the literature on narrative selves can agree on (Goldie, 2012; Lamarque, 2004; Menary, 2008). Yet another group of writers focuses on representations of one's own body (Hohwy & Michael, 2017; Metzinger, 2003). Hohwy, for instance, aims to explain 'self-representation in terms of inferred hidden causes' (p. 374) where these 'deeply hidden causes stem from the organism itself' (p. 375).

If something like my account is correct, then the disagreements between these various proposals can't be about what constitutes a genuine self-representation. After all, any coordinated non-egocentric representation, no matter its content, is a genuine self-representation. To me, this suggests that one of two things is the case. First, it's possible that there's a genuine disagreement about the kinds of property that need to be represented for the content of the self-representation to be sufficiently rich to capture a self's essential properties. Second, it's possible that there's only a surface disagreement and that these accounts try to capture different kinds of – possibly complementary – self-representation.

Peacocke (2014) seems to think that a self essentially exemplifies both temporal and spatial properties. A creature with just a non-egocentric spatial representation only represents '*this place on the map is here*' (p. 30, emphasis in the original) and *de se* content requires representing one's trajectory in time. But if that *here* location is connected to sensorimotor information in the sense discussed for the self\*-token throughout this paper, then both a self\*-token in just a spatial representation and one that specifies both spatial and temporal properties seems to genuinely self-represent. What's left, then, seems to be the claim that selves are essentially temporally extended and that a genuine self-representation must capture this property.

Similarly, Musholt seems to think that a self essentially exemplifies mental properties, and that failure to represent these *qua* mental properties is a failure to represent a self *qua* self. She writes that 'for a subject to realize that other subjects are distinct beings with their own mental states, and as such are similar to herself [...] she [...] needs to understand that others possess first person information about their mental states which might differ from her own' (Musholt, 2012, p. 78). It is therefore only at around one-and-a-half years of age that a 'child begins to explicitly differentiate between self and other' (p. 79). Note how this argument seems to presuppose that the child must understand what kind of being she is, where this essentially involves having mental states. Without representing oneself as a bearer of mental states, the child fails to represent herself as she is and thus fails to self-represent.

If this is really what these accounts disagree about, then their disagreement is about selves' essential properties. In other words, the disagreements seem to draw on differing answers regarding the metaphysics of selfhood rather than questions particular to self-representation.

Alternatively, we may want to reinterpret the differences in accounts of substant-

ive self-representation not as disagreements but simply describing different kinds of self-representation. As we've seen throughout this paper, creatures may minimally self-represent different kinds of property. Of course, a creature can also minimally self-represent *a range* of properties of different kinds. Thus, while I used fruit bats to illustrate *spatial* self-representation, it seems likely that they also temporally self-represent – this would allow them to infer that certain motor commands will make it so that they arrive at a target location *at some moment in the future*.

The differences between the various accounts I introduced could then be seen as differences in the richness of the self-representations involved. Such differences are important, and depending on the use to which one wants to put one's concept of self-representation, we might prefer one over the other.

I want to make two quick notes about building up more substantive self-representations from minimal ones. First, whatever properties may be represented by a substantive representation, they need to be represented in coordinated non-egocentric representations. Otherwise, they fail to fulfil even the minimum criteria for genuine self-representation. Thus, the fruit bat would need to realise both a spatial and temporal non-egocentric representation.

However, and this is the second point, this isn't sufficient; the relevant non-egocentric representations need to also be integrated. The representations need to be coordinated with sensorimotor information such that sensorimotor information is inferred as a function of the content of the temporal *and* spatial representation. The fruit bat represents that *it itself is at some location now* because it uses information about temporal *and* spatial properties together to infer sensorimotor information. Without such integration, the bat might realise two self-representations, but there would be no (tacit) representation of the fact that these representations' self\*-tokens refer to the same subject.

Future empirical research should investigate which creatures non-egocentrically represent which properties and to what extent these representations are integrated.

Human beings are then distinguished from bats by having very highly integrated and high-dimensional self-representations. A self\*-token in such a representation self-ascribes a multitude of properties: in addition to spatial, temporal, bodily, and mental properties, we also represent social properties, our character traits, narrative links between events, and much much more. The difference between our self-representations and those exemplified by bats and scrub-jays is one of degree.

## 8. Conclusion

I've suggested that some organisms use coordinated non-egocentric representations to navigate space, keep track of temporal relations, control their body movements, and monitor what other people know about the world. In these representations, self

and other are explicitly differentiated, and a self\*-token enables coordination with sensorimotor information, which links the representation to behaviour and sensation. Moreover, the account can explain why some self-attributed properties are immune to error through misidentification whereas others are not.

Whenever a system represents a property with a coordinated non-egocentric representation, it realises a minimal self-representation. As we have seen, young infants and a diverse roster of non-human animals are capable of realising representations of this type, making self-representation more common than many would have believed. By integrating minimal self-representations of a variety of properties, more substantial forms of self-representation – such as those displayed by neurotypical adult human beings – may emerge.

## Acknowledgements

I want to thank Mark Sprevak for feedback on early versions of the ideas developed in this paper, Hadeel Naeem for reading drafts as the work progressed, and members of the Philosophy of Cognitive Science reading group at LOGOS (Universitat de Barcelona) for detailed discussions.

## Author statements

**Financial Support / Funding statement:** The author was supported by the grant P500PH\_202829 / 1 of the Swiss National Research Foundation.

**Competing Interest statement:** The author hereby declares not to have any conflict of interests (financial or non-financial) regarding the present research.

## References

- Arcediano, Francisco, Escobar, Martha, & Miller, Ralph R. (2003). Temporal integration and temporal backward associations in human and nonhuman subjects. *Animal Learning & Behavior*, 31, 242–256. <http://dx.doi.org/10.3758/bf03195986>
- Babb, Matthew. (2016). The essential indexicality of intentional action. *The Philosophical Quarterly*, 66, 439–457. <http://dx.doi.org/10.1093/pq/pqw023>
- Baciadonna, Luigi, Jerwood, Georgia M., Farrar, Benjamin G., Clayton, Nicola S., & Emery, Nathan J. (2022). Investigation of mirror-self recognition in ravens (*corvus corax*). *Journal of Comparative Psychology*, 136, 194–198. <http://dx.doi.org/10.1037/com0000319>
- Baranes, Adrien, & Oudeyer, Pierre-Yves. (2013). Active learning of inverse models with intrinsically motivated goal exploration in robots. *Robotics and Autonomous Systems*, 61, 49–73. <http://dx.doi.org/10.1016/j.robot.2012.05.008>

- Barnet, Robert C., Cole, Robert P., & Miller, Ralph R. (1997). Temporal integration in second-order conditioning and sensory preconditioning. *Animal Learning & Behavior*, 25, 221–233. <http://dx.doi.org/10.3758/bf03199061>
- Bermúdez, José Luis. (2017). Yes, essential indexicals really are essential\*. *Analysis*, 77, 690–694. <http://dx.doi.org/10.1093/analys/anx137>
- Berthier, Neil E. (1996). Learning to reach: A mathematical model. *Developmental Psychology*, 32, 811–823. <http://dx.doi.org/10.1037/0012-1649.32.5.811>
- Berthier, Neil E., & Keen, Rachel. (2006). Development of reaching in infancy. *Experimental Brain Research*, 169, 507–518. <http://dx.doi.org/10.1007/s00221-005-0169-9>
- Bruner, J. S., & Koslowski, Barbara. (1972). Visually preadapted constituents of manipulatory action. *Perception*, 1, 3–14. <http://dx.doi.org/10.1068/p010003>
- Bugnyar, Thomas, Reber, Stephan A., & Buckner, Cameron. (2016). Ravens attribute visual access to unseen competitors. *Nature Communications*, 7. <http://dx.doi.org/10.1038/ncomms10506>
- Buhusi, Catalin V., & Meck, Warren H. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nature Reviews Neuroscience*, 6, 755–765. <http://dx.doi.org/10.1038/nrn1764>
- Butterfill, Stephen A., & Apperly, Ian A. (2013). How to construct a minimal theory of mind. *Mind & Language*, 28, 606–637. <http://dx.doi.org/10.1111/mila.12036>
- Campbell, John. (1999). *Past, space, and self*. NetLibrary, Inc.
- Cappelen, Herman, & Dever, Josh. (2014). *The inessential indexical: On the philosophical insignificance of perspective and the first person* (1 edition). Oxford University Press.
- Castañeda, Hector-Neri. (1966). 'He': A study in the logic of self-consciousness. *Ratio*, 8(December), 130–157.
- Castellini, Claudio, Orabona, Francesco, Metta, Giorgio, & Sandini, Giulio. (2007). Internal models of reaching and grasping. *Advanced Robotics*, 21, 1545–1564. <http://dx.doi.org/10.1163/156855307782148587>
- Clayton, Nicola S., Dally, Joanna M., & Emery, Nathan J. (2007). Social cognition by food-caching corvids. The western scrub-jay as a natural psychologist. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362, 507–522. <http://dx.doi.org/10.1098/rstb.2006.1992>
- Clifton, Rachel K., Muir, Darwin W., Ashmead, Daniel H., & Clarkson, Marsha G. (1993). Is visually guided reaching in early infancy a myth? *Child Development*, 64, 1099–1110. <http://dx.doi.org/10.1111/j.1467-8624.1993.tb04189.x>
- Coliva, Annalisa, & Palmira, Michele. (2024). Immunity to error through misidentification: Some trends. *Philosophical Psychology*, 1–36. <http://dx.doi.org/10.1080/09515089.2024.2371423>
- Collett, Thomas S. (2019). Path integration: How details of the honeybee waggles dance and the foraging strategies of desert ants might help in understanding its mechanisms. *Journal of Experimental Biology*, 222. <http://dx.doi.org/10.1242/jeb.205187>

- Corbetta, Daniela, Thurman, Sabrina L., Wiener, Rebecca F., Guan, Yu, & Williams, Joshua L. (2014). Mapping the feel of the arm with the sight of the object: On the embodied origins of infant reaching. *Frontiers in Psychology*, 5. <http://dx.doi.org/10.3389/fpsyg.2014.00576>
- Corbetta, Daniela, Wiener, Rebecca F., Thurman, Sabrina L., & McMahon, Emalie. (2018). The embodied origins of infant reaching: Implications for the emergence of eye-hand coordination. *Kinesiology Review*, 7, 10–17. <http://dx.doi.org/10.1123/kr.2017-0052>
- Coslett, H. Branch, Buxbaum, Laurel J., & Schwoebel, John. (2008). Accurate reaching after active but not passive movements of the head: Evidence for forward modeling. *Behavioural Neurology*, 19, 117–125.
- Cothi, William de, Nyberg, Nils, Griesbauer, Eva-Maria, Ghanamé, Carole, Zisch, Fiona, Lefort, Julie M., Fletcher, Lydia, Newton, Coco, Renaudineau, Sophie, Bendor, Daniel, Grieves, Roddy, Duvelle, Éléonore, Barry, Caswell, & Spiers, Hugo J. (2022). Predictive maps in rats and humans for spatial navigation. *Current Biology*, 32, 3676–3689.e5. <http://dx.doi.org/10.1016/j.cub.2022.06.090>
- Dennett, Daniel C. (1982). Styles of mental representation. *Proceedings of the Aristotelian Society*, 83, 213–226. <https://www.jstor.org/stable/4545000>
- Desmurget, Michel, & Grafton, Scott. (2000). Forward modeling allows feedback control for fast reaching movements. *Trends in Cognitive Sciences*, 4, 423–431. [http://dx.doi.org/10.1016/S1364-6613\(00\)01537-0](http://dx.doi.org/10.1016/S1364-6613(00)01537-0)
- Dienes, Zoltan, & Perner, Josef. (1999). A theory of implicit and explicit knowledge. *Behavioral and Brain Sciences*, 22, 735–808. <http://dx.doi.org/10.1017/S0140525X99002186>
- Emery, N. J., & Clayton, N. S. (2001). Effects of experience and social context on prospective caching strategies by scrub jays. *Nature*, 414, 443–446. <http://dx.doi.org/10.1038/35106560>
- Emery, Nathan J., & Clayton, Nicola S. (2008). How to build a scrub-jay that reads minds. In Shoji Itakura & Kazuo Fujita (Eds.), *Origins of the social mind: Evolutionary and developmental views*. Springer Japan.
- Evans, Gareth. (1982). *The varieties of reference* (John Henry McDowell, Ed.). Clarendon Press; Oxford University Press.
- Filimon, Flavia. (2015). Are all spatial reference frames egocentric? Reinterpreting evidence for allocentric, object-centered, or world-centered reference frames. *Frontiers in Human Neuroscience*, 9. <http://dx.doi.org/10.3389/fnhum.2015.00648>
- Fuchs, Thomas. (2017). Self across time: The diachronic unity of bodily existence. *Phenomenology and the Cognitive Sciences*, 16(2), 291–315.
- Gallistel, C. R. (1989). Animal cognition: The representation of space, time and number. *Annual Review of Psychology*, 40(1), 155–189. <https://doi.org/10.1146/annurev.ps.40.020189.001103>
- Goldie, Peter. (2012). *The mess inside: Narrative, emotion, and the mind*. Oxford Uni-

- versity Press.
- Grush, Rick. (2000). Self, world and space: The meaning and mechanisms of ego- and allocentric spatial representation. *Brain and Mind*, 1(1), 59–92.
- Grush, Rick. (2004). The emulation theory of representation: Motor control, imagery, and perception. *Behavioral and Brain Sciences*, 27, 377–396. <http://dx.doi.org/10.1017/S0140525X04000093>
- Grush, Rick. (2008). Temporal representation and dynamics. *New Ideas in Psychology*, 26(2), 146–157. <http://www.sciencedirect.com/science/article/pii/S0732118X07000475>
- Gür, Ezgi, Duyan, Yalçın Akın, & Balcı, Fuat. (2018). Spontaneous integration of temporal information: Implications for representational/computational capacity of animals. *Animal Cognition*, 21, 3–19. <http://dx.doi.org/10.1007/s10071-017-1137-z>
- Hofsten, Claes von. (1982). Eye–hand coordination in the newborn. *Developmental Psychology*, 18, 450–461. <http://dx.doi.org/10.1037/0012-1649.18.3.450>
- Hofsten, Claes von. (1991). Structuring of early reaching movements: A longitudinal study. *Journal of Motor Behavior*, 23, 280–292. <http://dx.doi.org/10.1080/00222895.1991.9942039>
- Hofsten, Claes von, & Fazel-Zandy, Shirin. (1984). Development of visually guided hand orientation in reaching. *Journal of Experimental Child Psychology*, 38, 208–219. [http://dx.doi.org/10.1016/0022-0965\(84\)90122-x](http://dx.doi.org/10.1016/0022-0965(84)90122-x)
- Hofsten, Claes von, & Lindhagen, Karin. (1979). Observations on the development of reaching for moving objects. *Journal of Experimental Child Psychology*, 28, 158–173. [http://dx.doi.org/10.1016/0022-0965\(79\)90109-7](http://dx.doi.org/10.1016/0022-0965(79)90109-7)
- Hohwy, Jakob, & Michael, John. (2017). Why should any body have a self? In Frédérique de Vignemont & Adrian J. T. Alsmith (Eds.), *The subject's matter: Self-consciousness and the body* (pp. 363–391). The MIT Press.
- Ismael, Jenann. (2008). *Situated self*. Oxford University Press, Incorporated.
- Klatzky, Roberta L. (1998). Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections. In *Lecture Notes in Computer Science* (pp. 1–17). Springer Berlin Heidelberg. [http://dx.doi.org/10.1007/3-540-69342-4\\_1](http://dx.doi.org/10.1007/3-540-69342-4_1)
- Konczak, Jürgen, Borutta, Maïke, Topka, Helge, & Dichgans, Johannes. (1995). The development of goal-directed reaching in infants: Hand trajectory formation and joint torque control. *Experimental Brain Research*, 106. <http://dx.doi.org/10.1007/bf00241365>
- Körding, Konrad P., & Wolpert, Daniel M. (2006). Bayesian decision theory in sensorimotor control. *Trends in Cognitive Sciences*, 10(7), 319–326. <http://linkinghub.elsevier.com/retrieve/pii/S1364661306001276>
- Kort, Selvino R. de, Dickinson, Anthony, & Clayton, Nicola S. (2005). Retrospective cognition by food-caching western scrub-jays. *Learning and Motivation*, 36, 159–176. <http://dx.doi.org/10.1016/j.lmot.2005.02.008>
- Lamarque, Peter. (2004). On not expecting too much from narrative. *Mind and Lan-*

- guage*, 19(4), 393–408. <http://doi.wiley.com/10.1111/j.0268-1064.2004.00265.x>
- Lepage, Jean-François, & Théoret, Hugo. (2007). The mirror neuron system: Grasping others? Actions from birth? *Developmental Science*, 10, 513–523. <http://dx.doi.org/10.1111/j.1467-7687.2007.00631.x>
- Lewis, David. (1979). Attitudes de dicto and de se. *The Philosophical Review*, 88, 513. <http://dx.doi.org/10.2307/2184843>
- Mach, Ernst. (1890). The analysis of the sensations. antimetaphysical. *The Monist*, 1(1), 48–68. <http://www.jstor.org/stable/27896829>
- MacKenzie, Matthew. (2010). Enacting the self: Buddhist and enactivist approaches to the emergence of the self. *Phenomenology and the Cognitive Sciences*, 9(1), 75–99.
- Marshall, Peter J., & Meltzoff, Andrew N. (2015). Body maps in the infant brain. *Trends in Cognitive Sciences*, 19, 499–505. <http://dx.doi.org/10.1016/j.tics.2015.06.012>
- McCormack, Teresa. (2015). The development of temporal cognition. *Handbook of Child Psychology and Developmental Science*, 1–47. <http://dx.doi.org/10.1002/9781118963418.childpsy215>
- Meltzoff, Andrew N., Saby, Joni N., & Marshall, Peter J. (2019). Neural representations of the body in 60-day-old human infants. *Developmental Science*, 22. <http://dx.doi.org/10.1111/desc.12698>
- Menary, Richard. (2008). Embodied narratives. *Journal of Consciousness Studies*, 15(6), 63–84.
- Metzinger, Thomas. (2003). *Being no one: The self-model theory of subjectivity*. MIT Press.
- Molet, Mikael, Miguez, Gonzalo, Cham, Henry X., & Miller, Ralph R. (2012). When does integration of independently acquired temporal relationships take place? *Journal of Experimental Psychology: Animal Behavior Processes*, 38, 369–380. <http://dx.doi.org/10.1037/a0029379>
- Musholt, Kristina. (2012). Self-consciousness and intersubjectivity. *Grazer Philosophische Studien*, 84(1), 63–89.
- Musholt, Kristina. (2013). Self-consciousness and nonconceptual content. *Philosophical Studies*, 163(3), 649–672. <http://link.springer.com/10.1007/s11098-011-9837-8>
- Musholt, Kristina. (2015). *Thinking about oneself: From nonconceptual content to the concept of a self*. MIT Press.
- Nguyen, Phuong D. H., Eppe, Manfred, & Wermter, Stefan. (2020). *Robotic self-representation improves manipulation skills and transfer learning*. <http://arxiv.org/abs/2011.06985v1>
- Peacocke, Christopher. (2014). *The mirror of the world: Subjects, consciousness, and self-consciousness*. Oxford University Press.
- Penn, Derek C., Holyoak, Keith J., & Povinelli, Daniel J. (2008). Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences*, 31, 109–130. <http://dx.doi.org/10.1017/S0140525X08003543>

- Penn, Derek C., & Povinelli, Daniel J. (2007). On the lack of evidence that non-human animals possess anything remotely resembling a 'theory of mind'. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362, 731–744. <http://dx.doi.org/10.1098/rstb.2006.2023>
- Perry, John. (1979). The problem of the essential indexical. *Noûs*, 13(1), 3–21.
- Perry, John. (1993). *The problem of the essential indexical: And other essays*. Oxford University Press.
- Perry, John. (2010). Selves and self-concepts. In Joseph Keim Campbell, Michael O'Rourke, & Harry Silverstein (Eds.), *Time and identity* (pp. 229–247). MIT Press.
- Perry, John, & Blackburn, Simon. (1986). Thought without representation. *Aristotelian Society Supplementary Volume*, 60, 137–166. <http://dx.doi.org/10.1093/aristotelian/supp/60.1.137>
- Recanati, François. (2009). De re and de se. *Dialectica*, 63, 249–269. <http://dx.doi.org/10.1111/j.1746-8361.2009.01194.x>
- Recanati, François. (2012). Immunity to error through misidentification: What it is and where it comes from. In *Immunity to error through misidentification: New essays* (pp. 180–201). Cambridge University Press.
- Recanati, François. (2024). IEM explained. *Philosophical Psychology*, 1–20. <http://dx.doi.org/10.1080/09515089.2024.2359493>
- Reid, Chris R., Latty, Tanya, Dussutour, Audrey, & Beekman, Madeleine. (2012). Slime mold uses an externalized spatial 'memory' to navigate in complex environments. *Proceedings of the National Academy of Sciences*, 109(43), 17490–17494. <https://www.pnas.org/content/109/43/17490>
- Rescorla, Michael. (2017). Maps in the head? In *The routledge handbook of philosophy of animal minds*. Routledge.
- Rolf, Matthias, & Steill, Jochen J. (2012). Goal babbling: A new concept for early sensorimotor exploration. *Proceedings of Workshop on Developmental Robotics*.
- Ryle, Gilbert. (2009). *The concept of mind*. Routledge.
- Saigusa, Tetsu, Tero, Atsushi, Nakagaki, Toshiyuki, & Kuramoto, Yoshiki. (2008). Amoebae anticipate periodic events. *Physical Review Letters*, 100(1), 018101. <https://link.aps.org/doi/10.1103/PhysRevLett.100.018101>
- Schillaci, Guido, Hafner, Verena V., & Lara, Bruno. (2016). Exploration behaviors, body representations, and simulation processes for the development of cognition in artificial agents. *Frontiers in Robotics and AI*, 3. <http://dx.doi.org/10.3389/frobt.2016.00039>
- Shoemaker, Sydney. (1968). Self-reference and self-awareness. *Journal of Philosophy*, 65(October), 555–567.
- Slaughter, Virginia, & Brownell, Celia A. (Eds.). (2011). *Early development of body representations*.
- Srinivasan, Mandyam V. (2015). Where paths meet and cross: Navigation by path integration in the desert ant and the honeybee. *Journal of Comparative Physiology*

- A, 201, 533–546. <http://dx.doi.org/10.1007/s00359-015-1000-0>
- Stachenfeld, Kimberly L., Botvinick, Matthew M., & Gershman, Samuel J. (2017). The hippocampus as a predictive map. *Nature Neuroscience*, 20, 1643–1653. <http://dx.doi.org/10.1038/nn.4650>
- Stalnaker, Robert C. (1981). Indexical belief. *Synthese*, 49, 192–151.
- Stoewer, Paul, Schlieker, Christian, Schilling, Achim, Metzner, Claus, Maier, Andreas, & Krauss, Patrick. (2022). Neural network based successor representations to form cognitive maps of space and language. *Scientific Reports*, 12. <http://dx.doi.org/10.1038/s41598-022-14916-1>
- Stoytchev, A. (2009). Some basic principles of developmental robotics. *IEEE Transactions on Autonomous Mental Development*, 1, 122–130. <http://dx.doi.org/10.1109/tamd.2009.2029989>
- Strawson, Peter Frederick. (2011). *Individuals: An essay in descriptive metaphysics* (Reprinted, transferred to digital printing). Routledge.
- Takemura, Naohiro, Inui, Toshio, & Fukui, Takao. (2018). A neural network model for development of reaching and pointing based on the interaction of forward and inverse transformations. *Developmental Science*, 21, e12565. <http://dx.doi.org/10.1111/desc.12565>
- Taylor, Kathleen M., Joseph, Victory, Zhao, Alice S., & Balsam, Peter D. (2014). Temporal maps in appetitive pavlovian conditioning. *Behavioural Processes*, 101, 15–22. <http://dx.doi.org/10.1016/j.beproc.2013.08.015>
- Thelen, Esther, Corbetta, Daniela, & Spencer, John P. (1996). Development of reaching during the first year: Role of movement speed. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 1059–1076. <http://dx.doi.org/10.1037/0096-1523.22.5.1059>
- Tolman, Edward C. (1948). Cognitive maps in rats and men. *Psychological Review*, 55, 189–208. <http://dx.doi.org/10.1037/h0061626>
- Tsoar, Asaf, Nathan, Ran, Bartan, Yoav, Vyssotski, Alexei, Dell’Omo, Giacomo, & Ulanovsky, Nachum. (2011). Large-scale navigational map in a mammal. *Proceedings of the National Academy of Sciences*, 108. <http://dx.doi.org/10.1073/pnas.1107365108>
- Wehner, Rüdiger, & Srinivasan, Mandyam V. (1981). Searching behaviour of desert ants, genus *cataglyphis* (formicidae, hymenoptera). *Journal of Comparative Physiology A*, 142, 315–338. <http://dx.doi.org/10.1007/bf00605445>
- Wentworth, Naomi, Benson, Janette B., & Haith, Marshall M. (2000). The development of infants’ reaches for stationary and moving targets. *Child Development*, 71, 576–601. <http://dx.doi.org/10.1111/1467-8624.00169>
- Witherington, David C. (2005). The development of prospective grasping control between 5 and 7 months: A longitudinal study. *Infancy*, 7, 143–161. [http://dx.doi.org/10.1207/s15327078in0702\\_2](http://dx.doi.org/10.1207/s15327078in0702_2)
- Wittgenstein, Ludwig. (2007). *The blue and brown books: Preliminary studies for the "philosophical investigations"*. Blackwell.

- Wittlinger, Matthias, Wehner, Rüdiger, & Wolf, Harald. (2006). The ant odometer: Stepping on stilts and stumps. *Science*, 312, 1965–1967. <http://dx.doi.org/10.1126/science.1126912>
- Wolpert, Daniel M., & Ghahramani, Zoubin. (2000). Computational principles of movement neuroscience. *Nature Neuroscience*, 3, 1212–1217. <http://dx.doi.org/10.1038/81497>

DRAFT