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

Julian Hauser

#### Abstract

One question we can ask when investigating the nature of selfrepresentation concerns the types of property that must figure in its content. Here, authors have claimed that self-representations must be about spatial, temporal, bodily, or mental properties. However, we can also ask a second question: *how* do we need to represent a property to *self* -represent it? I address this latter question. I argue that a distinction between egocentric and non-egocentric forms of representation – known from 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 minimal self-representations: they represent one's properties so that an explicit differentiation is made between the system other objects (or between the system's actual and merely possible properties), they are directly linked to behaviour and sensation, and they are immune to error through misidentification. The upshot is a view on which different creatures may self-represent more or fewer kinds of property. More substantive forms of self-representation (for instance, as exemplified by neurotypical adult human beings) then require integrated minimal self-representations of the right kinds of property.

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 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 representation with which a system represents 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 has *de se* content; with it, a system represents *itself qua itself*. To do so, the representation must fulfil two

conditions. First, it must explicitly differentiate between the system's properties and those of other objects. 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 the creature* isn't made explicit. We say that such a representation merely *concerns* the self, whereas a self-representation – containing a representational token that refers to the self – is *about* it (Perry, 1993).

Even a system instantiating a representation about itself may fail to know that this is the case – and hence fail to genuinely self-represent. An animal may see itself in a mirror but fail to recognise itself. This animal represents itself but fails to do so *qua itself*. Here, we arrive at the second condition: a self-representer must be disposed to update and use the represented information in a specific way that directly links it to sensation and behaviour (Evans, 1982). The creature must be disposed to use sensory information about itself to update its representation of itself. If someone were to draw a spot on the animal's forehead, the animal should represent itself as now having a spot on its forehead. The animal must be disposed to directly use the represented information to guide behaviour. For instance, the animal should infer that it needs to engage in certain motor actions to have an unstained forehead.

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 (O'Keefe & Nadel, 1978; Rescorla, 2017). 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, continuously updates a representation of the *distance and bearing* to its nest (Gallistel, 1989; Reid et al., 2012).

The distinction between egocentric and non-egocentric spatial representations has intriguing links to self-representation (see Grush, 2000). The token the ant uses to represent the nest in the egocentric case specifies a relation between the nest and the ant's position. Since this relation isn't explicitly represented, the representation concerns the ant (accuracy depends on the ant's location) but isn't about the ant (no representational token stands for the ant). The non-egocentric case is different: relations between representational tokens explicitly represent spatial relations. With a representational token that stands for the ant, such a representation explicitly differentiates between properties attributed to the system and to other objects. This representation is *about* the bat.

Another striking difference: while an egocentric representation cannot but concern the ant (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 on the map; non-egocentric representations represent objects as in principle independent of the system. Hence, to use its representation for navigation, the bat must employ a specific representational token as specifying *its* location. When this happens, I call the representation *coordinated*. The system now knows how to use its non-egocentric representation. For instance, it knows how to calculate the motor commands required to reach its cave based on the represented relation between its location and the cave.

Spatial properties constitute merely one kind of property that organisms represent with coordinated non-egocentric representations. I discuss evidence to this effect for representations of temporal, bodily, and cognitive properties. For instance, forward models (Grush, 2004; Wolpert & Ghahramani, 2000) imply an explicit differentiation between possible bodily states, making them non-egocentric representations. To predict a future bodily state, the model needs to be supplied the system's actual state, making 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 involves the system using a certain token in a special manner, namely as specifying its own state. Coordination entails, first, that relations from this token to other tokens imply sensorimotor information and that, hence, the system is disposed to use the represented information to guide behaviour. Second, coordination entails that sensorimotor information implies relations between this special token and other tokens so that the system is disposed to update the representation given relevant sensory input.

Since coordinated self-representations explicitly attribute properties to the self and link in the requisite way to behaviour and sensation, they are genuine (if minimal) self-representations. Whenever a system represents *any* property this way, it instantiates a minimal self-representation. 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 errors through misidentification (Evans, 1982; Shoemaker, 1968; Wittgenstein, 2007). A self-representation that results from a system inferring its properties from sensorimotor information is immune in this sense as sensorimotor information is inherently self-concerned. Self-attributed properties that aren't immune to errors through misidentification are inferred by identifying oneself with some object – such as my identification with the person I see reflected in the mirror in front of me.

The present work offers cognitive scientists an operationalizable proposal for studying self-representation by extending paradigms from spatial cognition research. The resulting studies can then reveal how various agents – from human beings to animals to artificial systems – exhibit self-representation, which may be more common than previously thought. This, in turn, may prompt philosophers to broaden their investigations beyond the case of adult human beings employing linguistic representations. Notably, genuine self-representation doesn't require

representations of mental states (Musholt, 2012), nor is it essentially tied to spatial cognition (Grush, 2000; Ismael, 2008), bodily representation (Hohwy & Michael, 2017; Metzinger, 2003), or 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 explain, for instance, the self-representations of neurotypical adult human beings. According to my account, properties must be represented non-egocentrically to figure in such a substantive self-representation. Moreover, my account suggests that many arguments in the philosophical literature may concern the kinds of content needed to represent a self's essential properties, rendering them disagreements about the metaphysics of selfhood rather than self-representation.

The paper begins with four sections examining coordinated non-egocentric representations across different creatures, starting with a section on spatial representations that introduces key concepts. These sections imply a link between coordinated non-egocentric representation and self-representation. Section 5 argues that coordinated non-egocentric representations are genuine self-representations, while Section 6 looks at how creatures infer their properties. Section seven links minimal selfrepresentation to substantive self-representation. Section eight concludes.

#### 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 some (Grush, 2000) have first suggested links between non-egocentric representations and self-representation. This section will distinguish between egocentric and non-egocentric representations in the spatial domain and gives 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 (Gallistel, 1989; Reid et al., 2012). These ants can return home from long foraging expeditions despite their featureless desert habitat. Moreover, if we capture a homewardbound ant and transfer it to a different location, it will proceed in the direction where the nest would have been, had we not relocated it (Wehner & Srinivasan, 1981).

In path integration, a creature encodes objects' locations through  $\langle distance, bearing \rangle$  tuples specifying relations from the current location. For instance, a desert ant may represent its nest to be 200 steps behind it. As it moves, the ant must update this representation, which it does using proprioceptive information about steps taken (Wittlinger et al., 2006). This explains successful navigation without external cues and why displaced ants continue in the original direction when proprioceptive inputs are absent.

Path integration employs *egocentric* spatial representations where the origin of the represented relations is the representer itself (Grush, 2000). A  $\langle distance, bearing \rangle$  tuple represents spatial relations *from the representer's current location*. A creature employing such a representation arranges objects around itself: one may be *just behind*, another *some distance towards the front*. These representations cannot represent relations between (non-self) objects.

As there exists a 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). The ant's representation fulfils this condition: it contains a  $\langle distance, bearing \rangle$  token, and (presumably) there exists a mechanism with which the ant can parse this token.

In contrast, the fact that the  $\langle distance, bearing \rangle$  token tracks a relation between the system's location (and orientation) and some object is only tacitly represented. The representation merely concerns the system rather than being about it (Perry, 1993). The ant uses the representation to navigate without explicitly representing that representational tokens relate to itself – it simply employs *all* tokens as representing such relations, using them to compute how it (and only it) can get to places.<sup>1</sup>

We arrive at the same conclusion if we think about what makes the ant's representation successful or accurate. Quite obviously, if the nest weren't where the ant represents it to be, the representation would be inaccurate. But an egocentric representation can also fail in another way. If another ant (at a different location) instantiated the same representation, the representation would fail to successfully guide behaviour. The ant's state is hence an *unarticulated constituent* of the representation (Perry & Blackburn, 1986).

As the representations employed in path integration do not contain a token standing for the system, they cannot underwrite an explicit differentiation between self and other. These representations concern but are not about the self, making them strictly speaking selfless (Lewis, 1979; see also Recanati, 2024) and, hence, not genuine selfrepresentations.

To find an instance of genuine self-representation, we need to look towards creatures whose more sophisticated behavioural profile relies on non-egocentric representation. In a recent experiment, Tsoar et al. (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

<sup>&</sup>lt;sup>1</sup> What I, following Dennett (1982), call *tacit* representation is sometimes called *implicit* representation in the philosophical and psychological literature. Additionally, in the psychological literature, 'implicit' and 'explicit' often denote conscious and unconscious or automatic and nonautomatic processes. This is a different distinction.

after clearing the crater's cliffs did they head home. Bats released outside the crater flew home immediately.

Tsoar et al. (2011) take the experiment to show that bats use *cognitive maps* (see O'Keefe & Nadel, 1978; Rescorla, 2017) to navigate their environment. These represent spatial relations, with tokens for objects and their spatial relations.<sup>2</sup> Importantly, such maps can represent relations between non-self objects, enabling behavioural abilities unattainable otherwise.

With cognitive maps, bats can plot novel routes to goal locations even when these aren't directly perceivable. After determining distances and bearings to landmarks like hills and settlement lights, bats compute direct routes home. The experiment carefully excluded alternative explanations: the crater was outside their home range (ruling out associative links), they were transported (excluding path integration), and the design controlled for magnetic, olfactory, and celestial navigation. This suggests bats represent the spatial relations between locations.

Cognitive maps are *non-egocentric* spatial representations (Grush, 2000). In such representations, the origin of the represented space isn't the representer itself. Instead, the origin might be some (non-self) object, so that the bat represents various locations in relation to, for instance, the cave. Alternatively, the represented space may originate at a location without an object as it does in maps that show the territory from above.

Crucially, non-egocentric representations explicitly relate objects to one another without having to involve to the representer's location. Unlike in the egocentric case, representational tokens don't tacitly represent – that is, concern – the location of the representer itself, and, consequently, they do not need to be updated during locomotion.

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 some of the represented objects. It must add to its map a token that stands for itself, explicitly differentiating between its own and other objects' locations. This representation is *about* the self.

However, genuine self-representation requires more than just a token that happens to be about oneself – the system must also know that this token refers to itself. Consider a bat seeing itself in a mirror but failing to recognise its reflection. Though the resulting representation includes a token about the bat, the bat cannot use this information to guide its behaviour. Following the literature on *de se* thought (Castañeda, 1966; Perry, 1979), I distinguish between mere self-tokens (that happen to be about oneself) and self\*-tokens (where one knows they are about oneself). Only self\*-tokens enable genuine self-representation with *de se* content. To keep the terminology simple,

<sup>&</sup>lt;sup>2</sup> Cognitive maps do not need to represent distances straightforwardly. As research on *predictive maps* argues, such maps might instead represent the probabilities associated with getting from one place to another (Stachenfeld et al., 2017). I believe such maps to be compatible with my proposal, but I cannot argue for this claim here.

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 way that links it to inherently egocentric sensorimotor information. Specifically, the system will be disposed to infer sensorimotor information based on nonegocentrically represented information and vice versa. In spatial representation, such sensorimotor information can be inferred from the vector whose origin is the self\*token and which ends at the token giving 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 some time). Similarly, if the bat has sensorimotor information regarding, say, a tree given in perception, it can add the tree to its cognitive map by converting the sensorimotor information to a vector which originates at the self\*token and terminates at a token representing the tree.

Borrowing a term from Grush (2000; see also Ismael, 2008), I call a non-egocentric representation *coordinated* when a self\*-token establishes links to and from sensor-imotor representations.<sup>3</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, egocentric-ally represented information, coordination necessarily involves a self\*-token.

A coordinated non-egocentric representation fulfils both conditions for selfrepresentation: it explicitly represents system properties and directly links them to behaviour and sensation. So far, I've only given a sketch of the argument; 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. Following this, I argue why such representations are genuine self-representations.

# Non-egocentrically representing temporal states

Several authors discuss the distinction between egocentric and non-egocentric representation in temporal cognition (Grush, 2008; Kort et al., 2005; McCormack, 2015). Many also argue that agents need to – or at least often do – *self*-represent temporal properties (Campbell, 1999; Goldie, 2012; Lamarque, 2004; Menary, 2008; Peacocke, 2014), making it important to show that non-egocentric representations can account for these. I'll proceed in two steps: first, I examine how egocentric representations of temporal properties fail to support certain complex behaviours, then I turn to non-

<sup>&</sup>lt;sup>3</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.

egocentric representations and their links to self-representation.

When food becomes scarce, slime mould amoebae merge into plasmodia that forage for nutrients. Saigusa et al. (2008) exposed these plasmodia to pulses of dry conditions, knowing they slow down in dry environments. After three exposures, slime moulds periodically slow down even without the dry condition stimulus, indicating they have begun to *anticipate* them.

To behave anticipatorily, slime moulds track time using regular internal oscillations (Saigusa et al., 2008). These oscillations encode a  $\langle timespan \rangle$  variable tracking temporal distance to the next dry condition. Using such *interval timing* (Buhusi & Meck, 2005), a slime mould can slow its movement when  $\langle timespan \rangle$  approaches zero.

Path integration and interval timing share key features suggesting both use egocentric representations. Where  $\langle timespan \rangle$  specifies the temporal distance from the present,  $\langle distance, bearing \rangle$  specifies the spatial distance and bearing from the current location (and orientation). In both cases, the representations concern (but aren't about) relations from the system's actual state. Hence, egocentric temporal and spatial representations require updating with each state change. Path integration updates spatial representation, while internal oscillations decrement  $\langle timespan \rangle$  to track time.

Egocentric temporal representations are restricted to representing properties with a single token concerning a relationship to the present. They 'are not carrying contents to the effect that this or that thing is happening at this or that time' (Grush, 2008, p. 156) and cannot explicitly represent relations between moments in time. Consequently, they 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*, as the explicitly represented content is a duration from the present to some other moment.

Grush and McCormack contrast this form of temporal cognition with 'a more sophisticated kind' (Grush, 2008, p. 156) that explicitly represents relations between events. A study by Arcediano et al. (2003) illustrates this. They exposed rats to two neutral sounds with a five-second gap ( $S_2 \rightarrow S_1$ ). In a second phase, the rats were exposed to an aversive stimulus *US* (a footshock) paired with  $S_1$  ( $US \rightarrow S_1$ ). In the test phase, rats exposed to  $S_2$  displayed decreased use of a lick bottle, demonstrating a fear response.

The rats anticipated the footshock (*US*) on exposure to *S*<sup>2</sup> despite never being conditioned on the sequence  $S_2 \rightarrow US$ . The authors argue this shows rats can integrate the two learnt temporal sequences to arrive at  $S_2 \rightarrow US \rightarrow S_1$ , leading the rats to expect *US* when exposed to *S*<sub>2</sub>. Arcediano et al. (2003) suggest this integration requires a temporal map, a kind of cognitive map already proposed by Tolman (1948).

A creature limited to egocentric representations could measure distances between events but couldn't use this to anticipate *US*. Consider how it might use  $R_{I_{ego}}$  to measure the duration between *S*<sub>2</sub> and *S*<sub>1</sub>, and  $R_{2_{ego}}$  for the duration between *US* and *S*<sub>1</sub>. These representations represent relations only tacitly through their activation at *S*<sup>2</sup> and *US*, and their use in anticipating *S*<sup>1</sup>. While one might suggest the creature could subtract  $R_{2_{ego}}$  from  $R_{1_{ego}}$  to find the correct duration, it would lack the cognitive wherewithal to do so. The egocentric nature of these representations means the creature cannot grasp that both relate to the same event. Moreover, even if it performed the calculation, it couldn't utilise the result – the output would lack the crucial non-representational link between event (*S*<sub>2</sub>) and token activation present in  $R_{1_{ego}}$  and  $R_{2_{ego}}$ .<sup>4</sup>

To integrate temporal sequences, rats must represent them non-egocentrically. Their representations  $R_{I_{non-ego}}$  (about  $S_2 \rightarrow S_1$ ) and  $R_{2_{non-ego}}$  (about  $US \rightarrow S_1$ ) explicitly represent relations between events, allowing integration of representations sharing common elements (Molet et al., 2012). In the experiment, rats learn that  $S_2$  precedes  $S_1$  and that US precedes  $S_1$ . Since  $S_2$  doesn't occur between US and  $S_1$ , they infer that  $S_2$  occurs before US, arriving at  $S_2 \rightarrow US \rightarrow S_1$ .

In non-egocentric representation, time functions as a framework (McCormack, 2015) organising temporal relations between events. These relations don't reference the system's actual state (the present), as evidenced by the rats' representation of  $S_2 \rightarrow US \rightarrow S_I$  remaining accurate without updates as time passes. Temporal relations are represented independently of the system's actual state.

Since non-egocentric temporal representations don't necessarily concern the system's state, coordination becomes an issue. Just as creatures may be ignorant of where they are, they may be ignorant of when they are. The rats illustrate this: Their stored representation of  $S_2 \rightarrow US \rightarrow S_1$  doesn't indicate how these events relate to the present. Only when  $S_2$  occurs can a rat infer that it is temporally co-located with  $S_2$ . It can then derive sensorimotor information from the non-egocentric representation and behave anticipatorily.

Coordinated non-egocentric representation of temporal properties links to self-representation. A creature with such a representation explicitly differentiates between its temporal properties and those of past and future events. Such representation is *about* the system instantiating it. Through coordination, the system uses 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. This suggests that creatures with non-egocentric temporal representations are genuine self-representers.

<sup>&</sup>lt;sup>4</sup> A simpler explanation might be that  $S_1$  takes on a negative valence after being paired with the footshock (*US*), and later exposure to  $S_2$  causes fear through expectation of  $S_1$  rather than *US*. Arcediano et al. (2003) addressed this possibility with an alternative version of their experiment.

#### Non-egocentrically representing bodily states

Many animals represent their own bodies. These representations develop early in ontogenesis (Meltzoff et al., 2019) and prepare the ground for representing the body-external environment (Stoytchev, 2009). Some of these representations are non-egocentric and suggest a link to minimal forms of self-representation. Unlike previous sections, I'll directly address non-egocentric representations without contrasting them with their egocentric counterparts.

Infants begin successfully reaching for objects at three to four months. Initially, their reaches are jerky and inaccurate, not following straight lines (Hofsten, 1982; Thelen et al., 1996). By five months, infants anticipatorily adjust their hands to object shapes (Hofsten & Fazel-Zandy, 1984; Witherington, 2005), and by seven months, they can reach efficiently (Thelen et al., 1996) and drink from open cups (Hofsten & Fazel-Zandy, 1984). At two years, their motion smoothness approaches adult levels (Berthier & Keen, 2006).

Researchers initially attributed early jerky movements to vision-guided reaching (see Corbetta et al., 2018), theorising that infants paused to check and correct their motions by gazing back and forth between hand and target. They thought smooth movements emerged only later, after developing a sense of the body. This view has been largely abandoned. Hofsten & Lindhagen (1979) found that infants fixate on the target object rather than alternating their gaze. More decisively, Clifton et al. (1993) demonstrated successful reaching even when infants cannot see their hands and arms, indicating they must already possess an embodied sense of hand location (Corbetta et al., 2014).

Evidence from developmental psychology and robotics (Baranes & Oudeyer, 2013; Desmurget & Grafton, 2000; Schillaci et al., 2016) explains these abilities through internal body models: *forward models* and *inverse models*. Forward models (Grush, 2004; Wolpert & Ghahramani, 2000) compute likely bodily states resulting from motor commands, enabling state estimation even with noisy sensory input. Hence, their predictions manifest in the brain before proprioceptive signals return from the sensory surfaces. Thus, an infant could predict failing to reach a toy before visual confirmation. Inverse models compute the motor commands needed for target states, enabling goaldirected behaviour. An infant could thus calculate reaching commands without visual hand-target alignment.

Forward and inverse models need information about the system's current bodily state. 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. The forward model, too, will predict different bodily states depending on the infant's current posture. The forward model relates possible bodily states in terms inferable from motor commands, whereas the inverse model relates them in terms allowing inference of required motor commands.<sup>5</sup>

Internal body models are non-egocentric representations. Not only do internal body models explicitly differentiate between various possible states, but the relations 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's thereby *about* the system's actual state.

To employ such a model, a system must use one body state token to specify the system's own actual state. As Coslett et al. (2008) put it, 'you are here' information is required 'to accurately reach toward an object' (p. 117). In other words, an internal body model needs to be coordinated for the system to use it. Coordination enables the system to infer sensorimotor information from relations between the self\*-token and tokens representing possible bodily states and to infer possible bodily states from sensorimotor information.

Coordinated non-egocentric representation of bodily states suggests a form of selfrepresentation. An inverse model represents the system as exemplifying a certain posture, explicitly differentiated from other postures it could exemplify. It is about the system's actual state (not merely concerning it). Like in certain non-egocentric temporal representations, the *other* from which the system is explicitly differentiated is its own merely possible state.<sup>6</sup>

# Non-egocentrically representing cognitive states

We routinely represent ourselves to have beliefs, desires, and other mental states, and some authors consider this ability necessary for self-representation (see Musholt, 2013). It's hence important to explain how non-egocentric representations can underwrite self-ascriptions of mental and cognitive properties. This section shows there are no principled difficulties in extending my account to such properties and demonstrates that such representations exist in certain non-human animals.

Recent experiments highlight sophisticated forms of social cognition in corvids, particularly scrub-jays (for instance, Bugnyar et al., 2016; Kort et al., 2005). Clayton et al. (2007) studied scrub-jays' caching and recovery behaviour in the presence of conspecific observers, exploiting the birds' tendency to cache food and pilfer others' caches. During a first caching event, an observer (A) watches through a transparent

<sup>&</sup>lt;sup>5</sup> Forward and inverse models only explicitly differentiate between possible bodily states if *articulated* (see Grush, 2004). Such articulation is likely the case. See Wolpert & Ghahramani (2000) and Grush (2004).

<sup>&</sup>lt;sup>6</sup> Mirror neurons activate body representations in babies (and adults) when observing others' actions (Lepage & Théoret, 2007). While this has been taken to suggest that imitation occurs because infants fail to differentiate between perceived and own target states, the infant must still maintain some distinct self\*-token. Otherwise, perceiving another's target state would not motivate imitation, as they would already see themselves as embodying it.

partition as the cacher hides food in tray A (tray B being visible but inaccessible). At a second event, a different observer (B) watches the cacher hide food in tray B (tray A now being inaccessible).

At recovery, cacher behaviour depends on which observer is present. With observer A present, cachers frequently move food from tray A to tray B, often relocating it multiple times, while leaving tray B's contents untouched. Thereby, the cacher attempts to protect the food observer A saw being cached while not revealing the food the observer doesn't know about. With observer B present, the behaviour is analogous but targets the other tray.<sup>7</sup>

The study's authors believe the experiment reveals corvids' representations of conspecifics' mental states, though others (Penn et al., 2008) consider such conclusions premature. These detractors argue the experiment fails to demonstrate that corvids exhibit 'a sensitivity to what others have and have not seen' (Clayton et al., 2007, p. 519) or engage in 'knowledge attribution' (ibid.). The caching bird might simply represent which observer was present at which caching event and re-cache food accordingly. This would explain scrub-jay behaviour through representations of spatial and temporal properties rather than cognitive states.

However, only birds who have pilfered others' caches engage in cache protection, which suggests an alternative explanation (Emery & Clayton, 2001). Corvids appear to infer others' behavioural patterns from their own, requiring them to distinguish between their own informational states and others'. Only then can the cacher differentiate between where it thinks the cache is located and where the conspecific thinks it is located. The cacher must recognise these informational states as being of the same kind, allowing it to understand that these states guide behaviour similarly. Thus, scrub-jays must represent that conspecifics can be in informational states of the same *kind* but differing in *value* from their own.

Butterfill & Apperly (2013) propose that scrub-jays distinguish between their own and others' spatial representations without representing mental states. Instead, they represent *registrations*, where '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 guide behaviour and can be incorrect. Through registrations, a scrub-jay can predict others' behaviour across many (though limited<sup>8</sup>) situations. Seeing a conspecific register food at a cache location, a cacher can anticipate attempted pilfering.

Scrub-jays represent registrations as *(individual, location, object)* tuples. At cache recovery, they select registrations matching the present conspecific and location. The cacher then acts to ensure the *location* and *object* elements of these registrations do

 $<sup>^{7}</sup>$  See Clayton et al. (2007) for details on how the experiment excluded the possibility that cache protection behaviour is cued by the observer's behaviour.

<sup>&</sup>lt;sup>8</sup> Registrations do not exemplify beliefs' full functional profile (Butterfill & Apperly, 2013). These details are not relevant to my argument.

not match its own representation of cache location. For example, if a scrub-jay represents food to be under tray A while a conspecific has registered it there, it will move the food to a new location.

We still need to understand how scrub-jays compare conspecifics' registrations with their own non-egocentric representations of cache location. For the cacher to infer others' behaviour from how their own informational states link to behaviour, the creature's representation of its own and others' informational states must be comparable. Since registrations differ from the scrub-jay's spatial representations, they must convert these into  $\langle individual, location, object \rangle$  tuples – where *individual* refers to the scrub-jay itself. It must create a representational token that refers to itself and with which it attributes registrations to itself, enabling comparison with conspecifics' registrations. How this is done doesn't matter for our purposes; what matters is that translation is required and that it involves self-attributing registrations.

This representation matches the pattern seen earlier: tokens represent objects (individuals) and their properties (registrations), while the system must use one token to specify its own state. This shows the key features of coordinated non-egocentric representation: object properties are represented as in principle independent of the system's state and using the representation requires coordination.

As with other non-egocentric representations, this has intriguing links to self-representation. A scrub-jay explicitly differentiates between where it and where others have encountered objects. The representational token that specifies its own registrations is *about* itself. Relations between its own registration and those of others imply sensorimotor information as evidenced by scrub-jays' caching behaviour. Such non-egocentric representation hence seems to imply a form of self-representation – scrub-jays seem to represent themselves (and others) as exemplifying cognitive states.

Note that we might still doubt whether representing registrations entails self-representation of mental states, seeing it instead as representing merely cognitive states. Authors like Musholt (2013), who require the representation of mental states for self-representation, would thus deny that corvids self-represent. I disagree: any coordinated non-egocentric representation constitutes self-representation. I now turn to defending this claim.

#### Minimal self-representation

I've argued that coordinated non-egocentric representations are used by various creatures to represent various kinds of property and have hinted at connections to *self*-representation. I now show why coordinated non-egocentric representation is minimal self-representation by demonstrating how it fulfils two conditions: such representation explicitly attributes properties to the representer, and the representer

knows that the representation is about itself through direct links to behaviour and sensation.

To be *about* the system, and fulfil the first condition on self-representation, a nonegocentric representation must explicitly represent the system's own state. It does so with an object-token that stands for the system, which I call (mere) self-token. With such a token, the fruit bat may represent its location and the scrub-jay its registrations, making the self an *articulated* constituent and explicitly differentiating between self and other. Note that in cases like internal body models, the self-token may relate to object-tokens referring to *the system's own* merely possible states, but this doesn't change the fundamental point: we have a representation *about* the self.

However, a (mere) self-token alone is insufficient – the system must know that it\* is the represented object. A fruit bat's cognitive map may include its location, but without knowing that this token represents its\* location, it cannot use it for locomotion and doesn't self-represent. Similarly, a rat representing event sequences needs to know which token represents its\* temporal state to behave anticipatorily – and to self-represent.

Self-representation requires that the non-egocentric representation is directly linked to sensation and behaviour. The bat must know where it\* is and the rat when it\* is. This parallels examples from the literature on self-locating beliefs: Perry must realise he\* has the torn sugar sack to adjust it (1979), and Mach must realise he\* is the dishevelled pedagogue to reach for the comb (1890).<sup>9</sup>

According to my proposal, a non-egocentric representation becomes a self-representation when *coordinated*. Coordination establishes systematic non-representational – or architectural (Ismael, 2008) – relations between sensorimotor and non-egocentric representation, enabling the creature to infer sensorimotor information (in an egocentric format) from non-egocentric representation and vice versa.

Such coordination necessarily creates a self\*-token – a special representational token, distinct from mere self-tokens, through which a system represents itself in a way that directly links to behaviour and sensation. A self\*-token is individuated not by what it represents but by how the cognitive system *employs* it. Consider spreadsheet software: selecting a cell highlights it, indicating which piece of data your command will affect. Like this highlighting, a self\*-token doesn't add information but indicates which token is linked to the system in the special way that manifests the system's know-

<sup>&</sup>lt;sup>9</sup> This isn't to say that a creature may never use *uncoordinated* non-egocentric representations. Imagine fruit bats can communicate food locations – a bat could then tell others that one fruit tree is closer to the cave than another without knowing its\* location. However, while the contribution of the uncoordinated non-egocentric representation may be important, even necessary, for behaviour, it isn't sufficient. The bat needs additional egocentric information concerning the communicative situation to infer relevant sensorimotor contingencies. For instance, if it expects favourable treatment for helping conspecifics, it could use the uncoordinated non-egocentric representation to determine how to help. For our purposes, it suffices that coordination *is* required in many important cases.

how that this token stands for itself.

A system knows that the self\*-token is about its\* state through the token's role in inferences between sensorimotor and non-egocentric information. Sensorimotor information, being egocentric, is self-concerning, while non-egocentric representation treats objects as independent of the system's state. For a non-egocentric representation to imply sensorimotor information, one of its tokens must be *used as* denoting the system's state. A fruit bat inferring the motor commands to get home needs a self\*-token indicating its\* position alongside the object token standing for the cave. Similarly, a rat predicting future events must use one represented moment as its\* temporal state, and an infant can only infer motor commands when she uses one bodily state as specifying her\* actual state.

Since creatures require self\*-tokens to use non-egocentric representations, these tokens are necessary to explain the resulting behaviour. This parallels the observation by Castañeda (1966) and Perry (1979) that certain indexicals are essential to action explanations.<sup>10</sup> We cannot explain why Perry adjusted the torn sack without referring to his realisation that he\* had the torn sack. Replacing 'he\*' with descriptions like 'the only bearded man in the store' fails to explain the behaviour if we don't attribute to Perry the knowledge that he\* is that man. Similarly, explaining a scrub-jay's recaching requires reference to a self\*-token. No amount of non-egocentric information (even including mere self-tokens) allows the bird to infer sensorimotor information without a self\*-token. Only through coordination and specification of a self\*-token can a system know the origin of the non-egocentrically represented relation that implies sensorimotor information.

Links in the opposite direction enable systems to update non-egocentric representations with egocentric information. A rat can only encode sensorimotor information about future events in its non-egocentric representation if it knows which token represents its\* temporal state. A non-egocentric token's value can only be inferred using both sensorimotor information *and* the self\*-token. Only then can the rat establish a correspondence between sensorimotor information and a relation from its\* temporal state to the event. Similarly, when an infant gains sensorimotor information about the motor commands required to, say, lift her arm, she needs a self\*-token to update her internal model since the sensorimotor information corresponds to a relation between her\* actual state and a goal state.

I have argued that non-egocentric representations explicitly differentiate between properties attributed to the self and others, while coordination links such representation to sensorimotor information. These links ensure that creatures with coordinated non-egocentric representations take the information as directly relevant to behaviour and update their representations based on sensory input.

<sup>&</sup>lt;sup>10</sup> Cappelen & Dever (2014) disagree, arguing that 'there is no such thing as essential indexicality, irreducibly *de se* attitudes, or self-locating attitudes' (p. 3). Millikan (1990) argues that the relevant mental tokens *are* essential, but not indexical. I am sympathetic to this latter view, but cannot discuss it here.

Hence, any creature representing any property with a coordinated non-egocentric representation is a self-representer. As I have shown through the examples of scrubjays, fruit bats, young infants, and rats, this means that minimal self-representation may be more common than expected.

#### How to find out who you are

A crucial question remains: how does a system determine which token to use as a self\*-token to coordinate non-egocentric representation with sensorimotor information? In this section, I show how a creature can gain information about itself in two ways: by making explicit the unarticulated subject component of its sensorimotor representations or by identifying itself with a non-egocentrically represented object. Self-attributions of the first kind are immune to certain errors through misidentification, which many philosophers consider essential for self-representation (Evans, 1982; Musholt, 2013; Shoemaker, 1968).

The most important way a system can learn about its<sup>\*</sup> properties is by establishing correspondences between sensorimotor information and relations in the non-egocentric representation. An infant with sensorimotor information linking motor commands to bodily movements can search her internal body model for a self\*-token whose relations to object-tokens imply this sensorimotor information. Just as a system can infer sensorimotor information from relations between a self\*-token and object-tokens, it can do the reverse and establish coordination by inferring a self\*-token from sensorimotor information and object-tokens.

Such coordination is possible because sensorimotor information concerns the self. When matching sensorimotor information to a relation in the non-egocentric representation, the object-token at the relation's origin articulates the sensorimotor representation's unarticulated subject component. Since this component refers to the self, its articulation as an object-token also refers to the self, allowing the system to designate that token as a self\*-token without possibility of error.

Such self\*-tokens are immune to errors through misidentification (IEM) like certain uses of the indexical 'I' (Evans, 1982; Shoemaker, 1968; Wittgenstein, 2007). These uses – what Wittgenstein (2007) called uses of 'I' as subject – make it impossible to ascribe the property to the wrong person. When thinking 'I see a tree', it makes no sense to wonder, 'Someone is seeing a tree, but is it *I*?' While I might be mistaken about the tree (maybe it's a cardboard cut-out), it seems impossible to be confused about the referent of 'I'. Similarly, a self\*-token inferred from the subject component of a sensorimotor 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). Even if a bat infers its self\*-token from inaccurate sensorimotor information, leading to misrepresentation of its properties, the self\*-token still refers to the system itself.

My explanation of why certain self-ascriptions are IEM aligns with the view developed by Recanati (2012a, 2024). Recanati, like me, traces the IEM status of certain self-ascriptions to the content of their grounds. Take, for example, when I self-ascribe the property of having crossed legs based on proprioceptive experience. According to Recanati, such experience has selfless content – when I sense my legs being crossed, I simply experience *crossed legs*. Since these grounds do not represent the self, no identification with an object is needed for self-ascription, thus ruling out misidentification. Rather, the relevant object – the self – is given by the *mode* of the experience. This mode, determined by the functional role of each experience type, is phenomenally accessible to the agent, allowing for the distinction between different kinds of experience independent of their content. Proprioception's functional role, for instance, is to convey information about one's own body, and experiences with this role have a distinctive feel. Given that proprioception exclusively provides information about the agent's own body, '[t]he person in question is, as it were, pre-identified, being determined by the mode of the experiential state' (Recanati, 2024, p. 9). Thus, I cannot misattribute the property when moving from a proprioceptive experience of crossed legs to self-attributing crossed legs.

This approach faces a difficulty: how to explain why *external mode* experiences – such as perception, whose role isn't to provide information about the self – can also ground IEM self-ascriptions? Recanati argues this is because these experiences are also 'bound to be about the subject of experience' (2024, p. 10). When seeing the Eiffel Tower, it is necessarily present *in the subject's environment*. Thus, my belief that *I am standing in front of the Eiffel Tower* is IEM. However, note that the mode no longer does any work: experiences of the internal and external mode can ground IEM self-ascriptions.

In my view, self-ascriptions are IEM when grounded in egocentric representations and, since experiences of the internal *and* external mode can be egocentrically represented, both may provide the relevant grounds. The crucial factor isn't that egocentric representations are selfless – when I look in a mirror, I see *myself*, yet this can ground the IEM self-ascription *I stand in front of a mirror*. Rather, IEM stems from the fact that these representations are *self-concerning* (viz. Ismael, 2012); whenever a self-ascription articulates an egocentric representation's unarticulated subject components, it is IEM.<sup>11</sup>

<sup>&</sup>lt;sup>11</sup> García-Carpintero (2013) argues that Recanati's account of *mental files* (Recanati, 2012b; see also Peacocke, 2014) is in tension with his account of IEM. The mental files account holds that agents maintain files containing object information, individuated by *epistemically rewarding* (*ER*) *relations*. For instance, seeing a tree generates a file storing properties like *is tall* and *is green*. The self-file, in this view, is a standard file where 'the relevant ER relation is the identity relation' (Recanati, 2012b, p. 68). This overlooks that self-representations aren't mere information stores but link to behaviour and sensation in a special way (that is relevant for IEM). My approach aligns more with Goodman and Gray's

In contrast, a self-ascription isn't IEM when it's based on identification with an object-token. If I visually perceive a broken arm after an accident and believe it's my<sup>\*</sup> arm, this belief isn't IEM, as the arm I see may not be mine. Here, the judgement that I have a broken arm is ultimately based on the articulated (explicit) *object* component of a sensorimotor representation. My seeing of the broken arm might instantiate sensorimotor information to the effect that if I moved my head this or that way, I would see it from this or that angle. From this, I might instantiate an object-token representing someone with a broken arm, then judge that I am identical with this person, creating a self<sup>\*</sup>-token attributing the broken arm to myself. Since the identity judgement is fallible, the self-ascription isn't IEM.

Note that a self-ascription's status as IEM depends on the ascription's grounds and not content (Evans, 1982). If I self-ascribed the property of having a broken arm based on proprioceptive information, it would ultimately derive from the subject component of an egocentric representation and be IEM.

At least *some* self\*-token values must derive from the articulation of an egocentric representation's unarticulated subject component. A judgement of identity between a self\*-token and some object-token obviously 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 requires yet another self\*-token. This well-known infinite regress (Evans, 1982; Peacocke, 2014; Shoemaker, 1968) can only be stopped by a self\*-token deriving from a sensorimotor representation's unarticulated subject component.

To conclude, self-ascriptions are IEM when derived from a sensorimotor representation's subject component but not when they stem from identification with an object-token. My account's explanatory power further evidences that coordinated non-egocentric representations are genuine self-representations.

# Substantive self-representation

Minimal self-representations are relatively common because they don't require selfrepresentation to be about specific kinds of property. This might raise concerns that minimal self-representation says too little about how adult neurotypical human beings self-represent. Hence, I examine how minimal self-representation relates to more substantive notions of self-representation.

Many kinds of property have been proposed as necessary (or at least privileged) for self-representation. Musholt believes that self-representation requires 'acquisition and application of the first person concept' (Musholt, 2013, sec. 4), which depends

<sup>(2022)</sup> mental filing account, where object-representations are individuated by their input and output relations. This alternative framework offers additional advantages, particularly in handling relational properties. Where the classical mental files view struggles with the awkward question of which file should store relational information, the alternative view sees relations as represented by the relations between object-tokens.

on representing one's own and others' *mental states* (Musholt, 2012). Grush (2000), following 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) argue that temporal properties are also necessary. This emphasis on temporal properties is also one of the few points of agreement in the literature on narrative selves (Goldie, 2012; Lamarque, 2004; Menary, 2008). Another group focuses on bodily representations (Hohwy & Michael, 2017; Metzinger, 2003) with Hohwy explaining 'self-representation in terms of inferred hidden causes' (p. 374) that 'stem from the organism itself' (p. 375).

My account suggests that these disagreements aren't about what constitutes genuine self-representation since any coordinated non-egocentric representation qualifies as such. One possibility is that these disagreements are about the properties a creature must represent to represent a self *qua* self. The dispute is then about the metaphysics of selfhood rather than self-representation as such.

Peacocke (2014) argues that self-representation requires representing temporal and spatial properties. A creature with only non-egocentric spatial representation represents '*this place on the map is here*' (p. 30, emphasis in the original), but *de se* content requires representing temporal properties. However, if that *here* location links to sensorimotor information in the relevant ways, then it qualifies as a self\*-token and thus a self-representation. 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 argues 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). Only at around eighteen months does a 'child begins to explicitly differentiate between self and other' (p. 79). Note how this argument appears to presuppose that the child must understand what kind of being she is in order to self-represent.

Alternatively, these differences might only appear to conflict and simply describe distinct – but complementary – kinds of self-representation. As shown throughout this paper, creatures can minimally self-represent various properties. A creature can, of course, also self-represent *a range* of kinds of property. While fruit bats illustrate *spatial* self-representation, they likely also temporally self-represent, enabling them to infer that certain motor commands will bring them to a target location *at some moment in the future*.

If this is correct, the differences between accounts mostly boil down to variations in substantive self-representations' richness. Such differences are important, and depending on our use of the concept of self-representation, we might prefer one over the other. According to this view of the disagreement, it remains important to remember that there are many kinds of self-representation and that research – empirical and philosophical – should investigate how these interrelate.

I want to make two quick notes about building substantive self-representations. First, to qualify as self-representations, they must represent all properties with coordinated non-egocentric representations. For a fruit bat to self-represent temporal and spatial properties, it must realise both spatial and temporal non-egocentric representations.

Second, the relevant representations must be integrated. They 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 is at some location now* by using 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 the relevant self\*-tokens refer to the same subject.

Human beings are then distinguished by having strongly integrated, highdimensional self-representations. Our self\*-tokens ascribe numerous properties: in addition to spatial, temporal, bodily, and mental properties, we also self-represent social properties, character traits, narrative links between events, and more. While debates about our self-representations' integration continue – see for example the disagreement between Ismael (2008) and Clark (2007) – my account shows that the difference between human and animal self-representation is one of degree.

#### Conclusion

I've suggested that some organisms use coordinated non-egocentric representations to navigate space, keep track of temporal relations, control their bodily movements, and monitor what other people know about the world. These representations differentiate between self-attributed and other-attributed properties and contain a self\*-token that coordinates with sensorimotor information to directly link representation with behaviour and sensation. This account also explains why certain self-attributed properties are immune to error through misidentification while others are not.

Systems realising coordinated non-egocentric representations exemplify minimal self-representation. As we have seen, young infants and a diverse roster of non-human animals can realise representations of this type, making self-representation more common than many might have believed. Through the integration of these minimal self-representations across various properties, more substantive forms emerge.

### References

Arcediano, Francisco, Escobar, Martha, & Miller, Ralph R. (2003). Temporal integration and temporal backward associations in human and nonhuman subjects. *An*- imal Learning & Behavior, 31, 242–256. http://dx.doi.org/10.3758/bf03195986

- 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
- 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
- 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 & Amp; 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.
- Clark, Andy. (2007). Soft selves and ecological control. In *Distributed cognition and the will: Individual volition and social context* (pp. 101–122). MIT Press.
- 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
- 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/k r.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.
- Dennett, Daniel C. (1982). Styles of mental representation. Proceedings of the Aris-

totelian 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
- 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/1 0.1038/35106560
- Evans, Gareth. (1982). *The varieties of reference* (John Henry McDowell, Ed.). Clarendon Press; Oxford University Press.
- 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
- García-Carpintero, Manuel. (2013). The self file and immunity to error through misidentification. *Disputatio*, *5*, 191–206. http://dx.doi.org/10.2478/disp-2013-0017
- Goldie, Peter. (2012). *The mess inside: Narrative, emotion, and the mind*. Oxford University Press.
- Goodman, Rachel, & Gray, Aidan. (2022). Mental filing. *Noûs*, *56*, 204–226. http: //dx.doi.org/10.1111/nous.12354
- 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/S0732118X070 00475
- 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, & 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
- 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
- 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.
- Ismael, Jenann. (2012). Immunity to error as an artefact of transition between representational media. In Simon Prosser & François Recanati (Eds.), *Immunity to error through misidentification: New essays* (pp. 62–80). Cambridge University Press.

- 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 Language*, 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
- McCormack, Teresa. (2015). The development of temporal cognition. *Handbook of Child Psychology and Developmental Science*, 1–47. http://dx.doi.org/10.1002/97811 18963418.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.
- Millikan, Ruth Garrett. (1990). The myth of the essential indexical. *Noûs*, *24*, 723. http://dx.doi.org/10.2307/2215811
- 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*, *16*3(3), 649–672. http://link.springer.com/10.1007/s11098-011-9837-8
- O'Keefe, John, & Nadel, Lynn. (1978). *The hippocampus as a cognitive map*. Clarendon Press ; Oxford University Press.
- 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

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, & Blackburn, Simon. (1986). Thought without representation. *Aristotelian Society Supplementary Volume*, 60, 137–166. http://dx.doi.org/10.1093/aristotelia nsupp/60.1.137
- Recanati, François. (2012a). Immunity to error through misidentification: What it is and where it comes from. In Simon Prosser & François Recanati (Eds.), *Immunity to error through misidentification: New essays* (pp. 180–201). Cambridge University Press.

Recanati, François. (2012b). Mental files. Oxford 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://ww w.pnas.org/content/109/43/17490
- Rescorla, Michael. (2017). Maps in the head? In *The routledge handbook of philosophy of animal minds*. 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.
- 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
- Stoytchev, A. (2009). Some basic principles of developmental robotics. *IEEE Transactions on Autonomous Mental Development*, *1*, 122–130. http://dx.doi.org/10.1109/t amd.2009.2029989
- Strawson, Peter Frederick. (2011). *Individuals: An essay in descriptive metaphysics* (Reprinted, transferred to digital printing). Routledge.
- 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/10.1037/009 6-1523.22.5.1059
- 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, & Ulan-

ovsky, 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
- 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
- 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/s cience.1126912
- Wolpert, Daniel M., & Ghahramani, Zoubin. (2000). Computational principles of movement neuroscience. *Nature Neuroscience*, *3*, 1212–1217. http://dx.doi.org/1 0.1038/81497