From a Sensorimotor to a Sensorimotor++
Account of Embodied Conceptual Cognition

Joel Parthemore

Centre for Cognitive Semiotics, University of Lund, Sweden
joel.parthemore@semiotik.lu.se

Abstract. Since the publication of O'Regan and Noë's original article in Behavioral and Brain Sciences in 2001, which first set out the sensorimotor account by which sensory experience and motor engagement are inextricably intertwined, there have been not just one but many sensorimotor accounts. However, in many ways that original article remains the canonical account. In this paper, I discuss a particular theory of concepts from philosophy of mind – the unified conceptual space theory, based on Peter Gärdenfors' conceptual spaces theory – and, in that light, set out what I take to be the key points of the 2001 account, along with its strengths and weaknesses. I discuss the ways in which the 2001 account aligns with, and departs from, the unified conceptual space theory; and I offer an extension to it that I call sensorimotor++, which adds to the 2001 account a key role for emotional affect and the somatosensory system, with which one might ground salience, and a key role for (so-called 'mental') representation, properly understood. I argue that sensorimotor++ makes for a better theory of concepts – one that is not just embedded and embodied but enactive – and, perhaps, a better sensorimotor theory more broadly.

Keywords: sensorimotor theory, concepts, representations, enaction, enactive, circular causality, conceptual spaces theory, unified conceptual space theory.

1 Introduction

Theories of concepts represent a sub-domain of philosophy of mind with considerable overlap into psychology and cognitive science, represented most prominently by such contemporary writers as Jesse Prinz [43], Jerry Fodor [10], and Peter Gärdenfors [13], with input from e.g. Ruth Millikan [28] and Alva Noë [31], and with a healthy criticism of the whole endeavour from e.g. Edouard Machery [24]. Concepts may be understood as the building blocks of systematically, productively, compositionally, and endogenously controlled structured thought, while conceptual abilities are those skills by which certain agents we identify as conceptual agents are able to cognize in a systematically, productively, etc. – and, above all, flexibly – structured fashion. I trace the notions of systematicsity (the same concepts can be used in more or less the same fashion across unboundedly
many contexts) and productivity (a finite number of concepts can be combined to form an unbounded number of complex concepts) to Gareth Evans’ Generality Constraint, set out in [8, 100-104]. Compositionality (the ability of concepts to be joined together or taken apart) follows directly from the first two properties. I owe the phrase ‘endogenously controlled’ to Prinz (see e.g. [43, p. 197], who offers it as an alternative to the potentially misleading ‘(Kantian) spontaneity’: concepts are not just passively given to the conceptual agent but somehow actively under her control. Note that nothing in this list entails that conceptual agents necessarily possess (human-style) language.

One may approach concepts and conceptual cognition in one of two ways, which I take to be equivalent. One talks of concepts as reified entities: the aforementioned “building blocks” of structured thought. The other talks of the abilities by which certain agents are able to engage with the world in cognitively creative ways. They are, to me, as two sides of a coin.

Meanwhile, the sensorimotor\(^1\) account [34] offers a theory about the nature of cognition more broadly, where ‘cognition’ may be understood in rough-and-ready terms as the encounter of cognitive mind with physical world (in a way that rejects any kind of Cartesian substance dualism – more on that below); or, in philosophical terms, the means by which certain agents effectively create an online/offline distinction in their interactions with their environment: there is the world, and then there are thoughts about the world.

### 1.1 Toggling Between Perspectives

Within the tea cup that is the field of concept studies, many a storm has brew ed over whether concepts are (or are best understood) as abstract (objects) or concrete (abilities), representational or non-representational, public or private, atomic or structured (see Figure 1): so e.g. on Jerry Fodor’s informational atomism account [10], concepts are atomic, public, representational, and abstract; on Jesse Prinz’s prototypes account [43] or Peter Gärdenfors’ conceptual spaces account [13], they are structured, both public and private, representational, and abstract; on Noë’s sensorimotor-theory-based account [31], I believe they are best understood as structured, public, non-representational, and concrete.

The problem is not, I think, that the concept of concepts is polysemous, as many might claim. The problem is more that we are standing too close to what we are trying to examine. On my own, unified conceptual space account [39,36,40], which is based on Gärdenfors’ work, concepts are either abstract (objects) or concrete (abilities), either representational or non-representational, either structured or atomic, depending on which of two perspectives – both necessary to any proper theory of concepts – one is taking at any particular time. If, most of the time, conceptual agents must, logically, get on with possessing and employing concepts without stopping to think about their concepts as concepts

\(^1\) The term ‘sensorimotor’ is meant to capture that which is necessarily and simultaneously sensorial and motor-based: senses and motor system are not two separate things to be brought together but two sides of one coin.
(the one perspective); then sometimes – for certain conceptual agents at least – the concepts themselves become the focus of attention (the other perspective). I believe that self-reflective conceptual agents toggle between these two perspectives constantly and, for the most part, non-reflectively: i.e., they do not stop to reflect on their own reflection (for practical reasons, as much as any fear of infinite regress!). Both perspectives are required, for anything like a complete view; but they cannot be resolved into a single perspective, on pain of contribution. Doing so would require setting our conceptual nature aside and stepping outside the observational system of which we are, inextricably, a part.

At the same time, on the unified conceptual space account, concepts are both public and private: again, depending on what perspective one is taking. One can talk about concepts for any given conceptual agent; for conceptual agents who are socially organized, one can talk about shared concepts for the group, which both relate closely to and yet differ from the individual agents’ concepts, as described very nicely by Prinz [43, p. 159]; for conceptual agents who possess human-like language, one can talk yet again about the words of a language by which the shared concepts are (perhaps imperfectly) lexicalized.

1.2 Escaping the Cognitivist Trap

For all of the frequent talk by Noë, by Vittorio Gallese and George Lakoff [12], and by many others of concepts as concrete abilities or direct engagements with
the world, concepts still often come off sounding highly abstract and far removed from the world. The target of complaint is variously described as cognitivism, symbolic AI, GOF AI ('good old-fashioned AI') – a term coined by John Haugeland in [18], or Cartesian mind/body dualism. The concern is that cognition comes off as disembodied and that the 'offline' mode is understood as actual detachment from – rather than inattention to – the environment. It creates an unintended and unnecessary *explanatory gap* between mind and body, cognitive agent and physical world: how to bridge the divide?

The peculiar, regress-inviting way by which we (necessarily, I think) conceptualize about our concepts should not – these researchers would say – mislead us into over-intellectualizing their nature: concepts ultimately are not about intellectual 'exercise' but hands-on activity. In keeping with them and *contra* e.g. Fodor's *informational atomism*, I want to reject the notion of concepts as physical symbol systems [30]: one stage in an input-output-based, *SMPA (sense-motivate-plan-act)* architecture easily describable in terms of a software program running on a digital computer. As Noë writes [31, p. 2]:

... We ought to reject the idea – widespread in both philosophy and science – that perception is a process in the brain whereby the perceptual system constructs an internal representation of the world. No doubt perception depends on what takes place in the brain, and very likely there are internal representations in the brain (e.g., content-bearing internal states). What perception is, however, is not a process in the brain, but a kind of skillful activity on the part of the animal as a whole.

Remember what I wrote earlier about concepts being one thing when we stop and reflect on them but logically another when we get on with possessing and employing them non-reflectively. The trick is that, the moment we reflect on what it is to possess and employ them non-reflectively, we bring them within the domain of (seemingly internal) representations! That does not mean, however, that we are powerless to say anything about what possessing and employing them non-reflectively could mean; and here something like sensorimotor theory, grounded in the empiricist tradition, seems to me to offer the best way forward, offering a way to trick ourselves out of the cognitivist trap.

### 1.3 The Unified Conceptual Space Theory

Space prevents me from recapitulating the *unified conceptual space theory* here (see the above references), other than to offer a very brief summary. In brief, however, conceptual spaces theory, on which the unified conceptual space theory is built, is a prototype- or similarity-space-based theory of concepts couched in the language of geometry, whereby concepts arise and evolve by the progressive partitioning and re-partitioning of conceptual spaces described as Voronoi tessellations (see Figure 2). The theory of prototypes – best examples of a category – derives from the work of Eleanor Rosch [45,44]; while a *similarity space* locates concepts in an abstract space defined by a set of integral dimensions with
a predefined metric, such that the closer two concepts (points or sub-regions) are located within the space, the more similar they are judged to be.

The unified conceptual space theory attempts to fill in some of the missing detail in conceptual spaces theory, whilst pushing it in a more algorithmically amenable, more empirically explorable way, inspired by the work in prototype theories in general and conceptual spaces theory in particular. It provides a specific algorithm – a kind of recipe – by which one can move from protoconcepts (foundational elements that fail to meet one or more of the properties offered above) to concepts to concepts of concepts. That algorithm is currently best described as semi-formal, though clearly detailed enough to allow direct translation of the theory into a mind-mapping software program (see Figure 2). The intention is to develop the algorithm into a properly formal (non-monotonic) logic, albeit one that allows for the apparent inconsistencies in most people’s conceptual frameworks. The thinking is that inconsistencies at the global level can be tolerated, provided they are sufficiently spatially removed from each other in the unified space (e.g., thinking X at Time T and not X at Time T+100); what cannot be tolerated is localized inconsistencies (e.g., thinking X and not X at Time T).

The recipe looks like this: for any given conceptual agent, concepts (of whatever type) are located within a common space of spaces that brings together all the many conceptual spaces described by conceptual spaces theory. An analogous space is meant to exist at the group or societal level, as inspired by [22] and, even more so, [49].

Concepts have both proximal and distal connections to one another. The proximal connections are along the three dimensions that define the unified space: what I call the axis of generalization (the familiar concept hierarchy: a dog is a mammal is an animal...), the axis of abstraction (from 'lower order' / concrete / 'physical' to 'higher order' / abstract / 'mental'), and the axis of alternatives (obtained by varying the values of any one or more integral dimension according to a pre-given metric: e.g., colour has the integral dimen-

---

2 That such a translation is possible should not be surprising: the theory was designed, from the beginning, with such an application in mind. For more on mind-mapping programs, which are intended to allow users to brainstorm ideas and to 'externalize' their understanding of one or another conceptual domain, see e.g. [33] and [47, pp. 77-82].

3 For some readers, the worry will persist that any logic-based treatment of concepts or conceptual cognition will both over-intellectualize matters and fail to capture such a dynamic view of cognition as the enactive approach (see below) is committed to. The implied claim – that an appropriately designed algorithm or logic can capture such a dynamic view of cognition – is beyond the remit of this paper. Here, I will simply note Rick Grush and Patricia Churchland’s response to Roger Penrose, where they point out that it remains very much an open question whether there is any phenomenon in the physical universe that cannot, at least in principle, be algorithmically described (p. [15, p. 190]). I readily allow that, in practice, an algorithm for concepts is likely – indeed, almost certain – to leave much out; but the unified conceptual space theory makes no claim to be the final word on concepts.
Fig. 2. A sample of Voronoi tessellations, created by the mind-mapping program described in Chapter 8 of [36] as a direct translation of the unified conceptual space theory into software.

sions of hue, saturation, and brightness). All three dimensions are divergent in both directions, meaning that the geometry of the space is hyperbolic rather than Euclidean. Meanwhile, distal connections are of three types: some concepts describe a component (mereological) relation to other concepts; some describe parameters or properties to other concepts; and all may be contextually related to (commonly associated with, but not required by) other concepts.

In Section Two, I set out what I take to be the pluses and minuses of sensorimotor theory, with respect to theories of concepts in general and the unified conceptual space theory in particular. Section Three proceeds to describe in greater detail how the canonical sensorimotor account and the unified conceptual space theory view cognition differently. Section Four presents sensorimotor++: an extended version of the sensorimotor account formulated as part of the unified conceptual space theory. Section Five summarizes the discussion and offers some conclusions.

2 Sensorimotor Theory: Pluses and Minuses

Before I discuss how sensorimotor theory, as described in [34], may be adapted to work with the unified conceptual space theory, I first must address how it does – and does not – fit that theory.

2.1 Pluses

The key points of sensorimotor theory, as described in [34] and implied by the quote above, I take to be these, all of which are strongly endorsed by the unified conceptual space account:

– All mental content, conceptual or otherwise, must be grounded in specific sensorimotor engagements, making sensorimotor engagements at least partly
constitutive of that content. This places sensorimotor theory squarely in the empiricist tradition embraced by proxytotypes theory, conceptual spaces theory, and the unified conceptual space theory and in contrast to the rationalist tradition embraced by informational atomism.  

Mental content is not fixed but dynamic and, indeed, contingent: if I do this, then I will experience that: a position that has much in common with so-called ecological psychology [14]. This accords as well with the enactive philosophy implicit in conceptual spaces theory (as I have described e.g. in [39]) and explicit in the unified conceptual space theory. Associated with Francisco Varela, Humberto Maturana, and Evan Thompson among others (see e.g. [27,26,54,51]), the spirit of enactivism is perhaps best summarized in a poem of Anotonio Machado, translated by Varela [53, p. 63]):

Wanderer, the road is your footsteps, nothing else; wanderer, there is no path, you lay down a path in walking. In walking you lay down a path...

- In keeping with this, mental content is part-and-parcel with interaction: input is logically inseparable from output, except in certain - very limited and circumscribed - cases that should not be mistaken for the general case. This, too, accords with enactivism, by which neither pre-conceptual agent nor pre-conceptual environment are recoverable from their interaction: I have proposed using the term enactive to... evoke the idea that what is known is brought forth, in contrast to the more classical views of either cognitivism or connectionism [27, p. 255]. Indeed, Noë for a time called his version of sensorimotor theory, developed in the years following the 2001 paper with O’Regan, ‘enactive’ (see e.g. [31]), though in more recent years he has preferred to talk of actionism (see e.g. [32]).

O’Regan and Noë introduce the key phrase sensorimotor contingencies to

---

4 Contrast this with e.g. the position described in [12], where all mental content, including the most abstract, just is specific sensorimotor engagements, albeit with parts of those engagements routinely suppressed (when one thinks or talks of democracy, one need not ‘do’ democracy): that is to say, there is no residue to be explained once sensorimotor engagements have been accounted for. The account is oddly reminiscent of George Berkeley’s discussion [1] of triangles and his argument that no one has an abstract concept of triangle that is anything more or other than a specific triangle instance (for Berkeley, a mental picture of a triangle). Like Berkeley, Gallese and Lakoff reject the notion of abstract classes in favour of ‘concrete’ instances.

5 Sensorimotor theory constitutes a refinement and renewal of classical empiricism (I have in mind in particular the British empiricism of John Locke, David Hume, and George Berkeley) – as Prinz has tried to do in his own way with his proxytotypes theory. Contra rationalism, for all of the importance of reasoning to (human) cognition, cognition is ultimately grounded not in reason but experience. For all that reasoning may conclude that something surely must logically be impossible, experience can always come along and say: no, in fact; here is the proof that it cannot be!
describe what governs these interactions. Here is how they describe holding a bottle [34, p. 945]:

In fact... you may well have very little sensory stimulation coming from the bottle at the present moment. Yet, you actually have the feeling of 'having a bottle in your hand' at this moment. This is because your brain is 'tuned' to certain potentialities: if you were to slide your hand very slightly, a change would come about in the incoming sensory signals which is typical of the change associated with the smooth, sliding surface of glass.

- Most of cognition has little to do with representation, in any but the most loosely metaphorical of senses. Whatever representation there is and whatever representation is about, 'there is no "re"-presentation of the world inside the brain' [34, p. 940]. This is because, most of the time, cognition is not something we think about: it's something we do, an 'exploratory activity' [34, p. 940] where one experiences the world by exploring it. What one cannot actively explore, one cannot experience (see e.g. [19]). Remember what I said earlier that, most of the time, conceptual agents just get on with possessing and employing their concepts non-reflectively; it is only when they stop to examine their concepts that their concepts become representational.

- Sensorimotor theory recognizes the important difference between so-called mental imagery and pictures - a point that Göran Sonesson has long argued (for a recent treatment, see [48]). According to the classical imagist account of concepts, as described in [43, pp. 25-31], concepts just are conscious pictures in the mind. He writes [43, p. 139]: 'to bring concept empiricism up to date, one must abandon the view that concepts are conscious pictures'. Instead, for Prinz, concepts are partly unconscious multimodal 'images'. According to the unified conceptual space theory, conflating mental imagery with pictures (conscious or otherwise) confuses the very nature of what they have in common: namely, what makes them both (under certain circumstances) representations (see Section 3.3 and [38]). Unlike pictures, concepts are not primarily things that we 'see', even metaphorically.

In sum, sensorimotor theory consciously tries to avoid over-intellectualizing cognition: a surely laudable effort. If we think too much about how we think, we risk losing our very target in an elaborate set of mind games (not to mention, we will never be seen to do anything, because we will always be thinking about what to do). Sensorimotor theory takes on board the best of Rodney Brooks and the lesson that, if an agent can offload some amount of its cognition onto the world - knowing that the world will be one way and not another (the source of the - in my mind - misleading and ill-advised proverb that 'the world is its own best model' [2, p. 5]) - surely it will. Putting this another way, sensorimotor theory reduces the number of steps needed for explaining a number of key cognitive abilities. If you don’t need to build more in to make it work, don’t: a fine

6 Cf. [34, p. 939]: 'the outside world serves as its own, external, representation'.
application of Occam’s razor. The lesson is one that theories of concepts—for all their inherent tendency to over-intellectualize cognition—would do well to mind.

2.2 Minuses

As described in the original 2001 paper, sensorimotor theory assumes—indeed, depends on—a particular, realist metaphysical position without making that commitment clear: a charge I would similarly lay on Fodor’s informational atomism theory, which stands or falls on its brand of realism. If one does not share that position, then one might not feel that the so-called hard problem of consciousness[5]—which the 2001 paper boldly purports to dissolve—has really been addressed. Indeed, both conceptual spaces theory and the unified conceptual space theory are inclined toward—if not necessarily dependent on—an antirealist metaphysics (a position that, indeed, much of the enactive community, with its emphasis on the inseparability of agent and environment, tends toward as well). At one point, the 2001 paper does come close to acknowledging that its position is more metaphysical than empirical [34, p. 948]: 'we are providing a general framework for the study of vision, and it is not possible to subject a general framework to direct verification. Our new framework provides scientists with new problems and it makes some old problems appear as non-problems...'.

Both the 2001 paper and, even more so, Noë’s subsequent book [31], make strong assumptions about what people experience that one need not necessarily share: e.g., 'many people say that they have the impression that when they see, the entire visual field is present to consciousness in all its nearly infinite detail’[34, p. 961]. Perhaps ‘many’ people do; but many, I would claim, do not. Similarly, Noë writes [31, p. 56]: 'if you reflect on what it is like for you to look at the wall, you will notice that it seems to you as if the whole wall is there, at once...’ and [31, p. 57] 'of course, it does seem to us as if we have perceptual access to a world that is richly detailed, complete, and gap-free. And indeed we do! We take ourselves to be confronted with and embedded in a high-resolution environment.' This strikes me as very much a philosopher’s and not a lay person’s way of putting things: the very error Noë is so anxious to avoid. As a philosopher, I am frequently and painfully made aware that non-philosophers (or other philosophers for that matter) do not necessarily view the world as I do. In any case, a theory of concepts—if it lays any claim to being a general theory of concepts—must be able to capture that portion of people’s experience that is conceptually structured7 in all its richness or poverty.

Although paying lip service to other sensory modalities—with the sensorimotor account is also meant to explain—the 2001 paper focuses much if not most

7 I leave aside, for now, the question of whether experience is entirely conceptually structured, as the so-called conceptualists would hold, or a mixture of the conceptual and non-conceptual, as the non-conceptualists contend—other than to say that, on this point, I side with the non-conceptualists. What matters here is that both sides of the debate agree that at least part of experience is conceptually structured.
of its attention on visual cognition and experience. To the extent that one is not attempting to explain all of cognition and experience, this is appropriate: one explains a part in the hope that it might be useful to the explanation of other parts. To the extent though that one is privileging vision above the other sensory modalities – as, I believe, many philosophers are implicitly inclined to do – one may well be, to pardon the metaphor, distorting the picture. Certainly in the context of conceptual cognition, too narrow of a focus on vision (or indeed any one modality) is not helpful, as the earlier discussion about classical imagism should have made clear.

O’Regan and Noë make a primary-vs.-secondary quality distinction – echoing John Locke – of which I am quite skeptical, given the spectre it raises of mind/body dualism. Primary qualities, of course, are meant to be ‘in’ an entity, independent of observer or context, whereas secondary qualities depend on observer and context. O’Regan and Noë apply this distinction not only to the objects of visual attention but to vision itself. There are, for example – at least in the 2001 paper – two kinds of sensorimotor contingencies [34, p. 943]:

Sensorimotor contingencies of the first sort – those that are determined by the character of the visual apparatus itself – are independent of any categorization or interpretation of objects and can be considered a fundamental, underlying aspect of visual sensation. Sensorimotor contingencies of the second sort – those pertaining to visual attributes – are the basis of visual perception.

I am inclined to push the earlier point about interaction further and suggest that both primary and secondary qualities, and both kinds of sensorimotor contingencies, arise from the interaction of agent and environment in such a way that observer and observed cannot cleanly be disentangled; i.e., the interaction is fundamental or foundational. This is, I think, what separates sensorimotor theory from enactivism – or, if you will, from other enactive accounts. This is how enactivism incorporates, but ultimately goes beyond, notions of embeddedness or situatedness (the cognitive agent is always located in a specific environment, which shapes its cognition) and embodiment (the cognitive agent always takes a particular physical form, which likewise substantively shapes its cognition). With the ‘secondary’ qualities and the one kind of sensorimotor contingencies, the role of the observer is explicit and unavoidable; with the ‘primary’ qualities and the other kind of sensorimotor contingencies, the role of the observer is implicit: the observer sits in the background, out of sight. Sometimes, attention will be more on the cognitive agent to the exclusion of physical environment, making one set most prominent; other times, attention will be more on the physical environment to the exclusion of the agent, and the other set will be emphasized. Either way, the observer is always present and must be accounted for: that is, both ‘primary’ and ‘secondary’ qualities, and both kinds of sensorimotor contingencies, depend on the interaction.

---

8 Note that this is really a metaphysical claim rather than one argued from the evidence: that is, it is taken as a starting assumption, whose merits or lack thereof lie in what conclusions it may lead to: if one assumes this, then this is what follows.
on the observer — without whom, there are neither qualities nor contingencies. As Humberto Maturana writes, 'everything that is said, is said by an observer to another observer that could be himself' [25, p. 30]. Inman Harvey writes, 'the underlying assumption of many is that a real world exists independently of any observer; and that symbols are entities that can 'stand for' objects in this real world in some abstract and absolute sense. In practice, the role of the observer in the act of representing something is ignored' [17, p. 5].

O'Regan and Noë are at pains to suggest that sensorimotor theory — unlike, presumably, its competitors — avoids any recourse to magic: e.g., [34, p. 946] '...by taking the stance that the experience of vision is actually constituted by a mode of exploring the environment, we escape having to postulate magical mechanisms to instill experience into the brain'. Avoiding magic is fair enough; but what exactly is magic? The supernatural, like the natural, is often discussed and rarely defined — as if everyone just knew already what it means. I am inclined to borrow a page from the science fiction author Arthur C. Clarke and suggest that magic is that which we do not currently understand — with the caveat that some things may not be possible for us to understand, even in principle, because they lie beyond our conceptual abilities (such as e.g. what it means to imagine a tesseract in all its four-dimensional glory). If that is what magic is, then, while I am inclined to agree with O'Regan and Noë that a sensorimotor-based approach leaves less explanatory residue, that is not to say there will be none remaining. If, by a 'fully naturalized account', one means an account that is both complete and consistent, then I would borrow a page from Douglas Hofstadter [20] in suggesting that, for any sufficiently expressive system, completeness and consistency rarely if ever sit comfortably together: one should expect explanatory residue (and not only when describing tesseracts!). Indeed, in [36], I argue — as part of the unified conceptual space theory — that, contra Roger Penrose [41, esp. pp. 72-77], there are good reasons to believe that conceptual understanding in general is knowably bounded even while there is no reason to think that the mind-independent world is similarly bounded.

Finally, for all of the importance of the 2001 paper and ensuing discussions, the sensorimotor account was never quite as new as it presented and presents itself. Its biggest contribution, I think, lies not in what it brings new to the table but in its way of describing things: things that many of us at least partly understood but lacked the adequate language for.

3 From Sensorimotor Toward Sensorimotor++

So far I have attempted to describe how the unified conceptual space theory aligns with, and where it departs from, the original sensorimotor account. Before I describe what the sensorimotor++ account, first described in [40] and again in [36], adds to to that account — i.e., what each of the pluses stands for — I need

---

9 The name 'sensorimotor++' was suggested to me by Peter Gärdenfors following a discussion with him about the draft of the paper that would become [40], later incorporated into Chapter 7 of [36].
to address in more detail how its outlook on cognition (and, by extension, that of the unified conceptual space theory of which it forms a part) differs.

3.1 Causality

Sensorimotor theory favours a linearly structured account; sensorimotor++ opts for a circular causality.

On the canonical account, sensorimotor engagements give rise to sensorimotor profiles and sensorimotor profiles to so-called higher cognition: a largely if not strictly uni-directional, bottom-up process from the sensorimotorly concrete to the conceptually abstract, from mechanically driven associations to flexible conceptual structures. This may be a consequence (see Section 3.2) of sensorimotor theory’s often very strong externalism and the consequent primacy it gives (on my reading) to the environment driving the agent rather than the agent driving the environment.

Clearly – as I am using the terms – not all cognitive agents are conceptual agents: that is, regardless of whether cognition ‘goes all the way down’ to the simplest organisms, as enactivism tends to favour (see e.g. [50]), conceptual cognition does not.\(^\text{10}\) In the case of cognitive agents who are not conceptual agents – which I take to be a majority – the bottom-up linear-causal account may well be the most appropriate. Such agents will be like the purely stimulus-response-driven automata Descartes envisioned all non-human animals to be.

Nevertheless, in the case of conceptual agents – those that meet all the desiderata offered in Section 1 – it seems to me that one can equally turn the perspective around, to look at how sensorimotor engagements consist of or are built upon conceptually structured perceptions: how mind constrains the (experienced) world. Experience gives rise to concepts, which, in turn, structure experience; it is logically impossible, as a conceptual agent, to set one’s concepts aside and step outside of the loop, to see the world ‘as it really is’: to do so would be like the dragon swallowing its own tail. One can at most gesture at what the world ‘outside the loop’ must be like. As I wrote in [40, p. 297]:

Concept acquisition and application go hand in hand. Acquiring concepts is a process of applying concepts, which may themselves change in the process of acquiring new concepts. … Our conceptual spaces, individually and collectively, are both the product of our interaction with our environment and the basis for it. The model of causality is not linear but circular.

\(^{10}\) Indeed, I am more inclined to the sort of account of ‘how far down cognition goes’ offered by Andy Clark and Rick Grush [6,16], where cognitive agents are those who, at least to a limited extent, are able to step back from the here-and-now and create a functional distinction to the ‘was’ or ‘might be’. I am indebted to one of the anonymous reviewers for suggesting the clarification and the references.
Note that, a circular causal account (see Figure 3) is only coherent if the two 'sides’ of the circle are considered independently of one another; i.e., as two separate instances of linear causality. Considered jointly, no consistent interpretation can be given.

![Circular Causality](image)

Fig. 3. A model of circular causality: $a$ causes $b$ causes $a$.

### 3.2 Externalism

*Sensorimotor theory is often strongly externalist; in keeping with its enactive perspective, sensorimotor++ seeks to avoid what it sees as the equally undesirable extremes – the Scylla and Charybdis if you will – of externalism and internalism.*

Remember that, for the enactive perspective even far more than for sensorimotor theory, everything keeps coming back to interaction: both experienced agent and experienced environment are a product of that interaction; agent does not separate cleanly from environment nor mind from world. On an enactive account, cognition in general and concepts in particular are not in the (brain of) the agent (internalism) nor in the agent’s environment (various forms of externalism); insofar as they can be located anywhere, they are metaphorically 'in' the interaction of agent with environment. It is in the nature of concepts always to be setting boundaries and creating categories; it is in the nature of the enactive perspective always to view those boundaries as flexible and, indeed, dynamic: conceptually drawn lines that mask underlying continuities. As the author has heard one prominent researcher in the field to say, enactivism does not think much of fixed boundaries.

In keeping with the enactive tradition, when talking about cognition or cognitive phenomena, one must, I think, take great care using terms like 'inside' and 'outside'. As I argue in [37], these are terms that apply to physical volumes, and their usage in any other contexts should be understood as loosely metaphorical at best: neither mind nor cognition nor sensory perception is *prima facie* a physical quantity even as they are all physically realized.
The result of sensorimotor theory’s externalism is, I believe, a tendency to overlook or downplay the role of the agent’s bodily states as interoceptively and proprioceptively experienced; and, in particular, to overlook or downplay the agent’s emotions. Indeed, this is the diagnosis offered by Anthony Morse and Tom Ziemke in an unpublished paper from 2010, where they call their proposal the somatic sensory hypothesis \[29\]. As they argue there, without an account of emotions and consequent motivations, sensorimotor theory cannot ground salience: it particular, it cannot explain why some affordances are salient and others not, nor why and how those saliences change over time. Although many philosophers, including Aaron Sloman\[12\] see rationality and emotion as opposed to each other, the unified conceptual space theory is inclined to see emotional affect as part of the necessary foundation to cognition in general – and, therefore, to conceptual cognition in particular.

### 3.3 Representationalism

Sensorimotor theory is often read as moderately or even strongly anti-representationalist; sensorimotor++ favours a qualified (or ‘modest’) representationalism.

Noë \[31\], in particular, I take to be anti-representationalist, in a way not so far removed from Brooks \[3\] or John Perry \[42\]. An informal poll would suggest that many if not most enactive philosophers are strongly anti-representationalist as well: consider e.g. the radical enactivism advocated by Daniel Hutto \[21\], for whom there is no such thing as mental content, let alone representation.

I share with O’Regan and Noë their distaste for so-called internal representations and the pride of place so often and uncritically, in certain circles – notably the cognitivist and GOF AI circles noted earlier – given to them. Indeed, in \[38\] I question the very coherence of talking about internal or mental representations as something ontologically distinct from any other kind of representation. Instead, I take the cue from Harvey \[17\] in restricting representation to a four-place function whereby agent $Q$ uses $R$ to stand in for $S$ for agent $T$ (who could, in certain circumstances, be agent $Q$ herself). In \[38\] I take representation to be not a thing – let alone a reification – but rather an intentionally reflective perspective that certain agents, in certain circumstances, take toward one aspect or another of their experience, be it a picture or painting, or a thought in the mind.\[13\] As such – and in keeping with any perspective one takes – they are

---

\[11\] It should be said that O’Regan acknowledges (personal communication) that the lack of any somatosensory account is one of the significant limitations of the 2001 paper – one that his notion of ‘bodiliness’ (see e.g. \[35\]) goes some way toward addressing; see also \[4\] for one suggestion of a formalized way to better take bodily states into account.

\[12\] Email correspondence.

\[13\] In this way my definition is, indeed, more restrictive than Harvey’s, since Harvey does not require any degree of intentionality and so is willing to attribute ‘minimal representations’ to his artificial-life creations.
neither 'internal' nor 'external'. Like 'inside' and 'outside', these terms apply most appropriately to physical volumes and only in a loose metaphorical sense to anything else.

Further, as noted in Section 2.1 and implied at several points through this paper, I share with O'Regan and Noë – and, indeed, Brooks and Perry and others – their belief that much if not most cognition can be explained without resort to representations anywhere except in the eye of the observer. Nevertheless, as I said in the introduction, when it comes to concepts and conceptual cognition it seems that representations and representational language are difficult if not impossible to avoid. If, as is doubtless true, most of the time we possess and employ our concepts without reflecting on them as concepts; nevertheless, when we do, representations are what we find: concepts do not simply re-present the world (as a photograph is often naively thought to do) but represent it as being one way and not another, with the possibility that the 'picture' they describe could be (according to intersubjective consensus at least) wrong.

3.4 Empiricism vs. Rationalism

Calling to mind Gilbert Ryle’s classic distinction [46], sensorimotor theory focuses on knowing how to the exclusion of knowing that. Sensorimotor++ suggests that, when it comes to concepts, concepts are neither precisely knowledge how nor knowledge that but something of neither and both. Most properly: concepts sit between knowledge how and knowledge that.

The longstanding debate in philosophy between rationalism and empiricism can be understood as a disagreement over which is ultimately foundational to (what we recognize as) cognition: reason grounded in knowledge that (the sort of stuff one can reflect upon or represent to oneself and others) or perceptual discovery grounded in knowledge how (the sort of stuff one can generally reflect upon only poorly and represent or describe inadequately: consider the difference between knowing how to ride a bike and knowing precisely what that knowledge consists of).

In keeping with sensorimotor theory, sensorimotor++ agrees that, when it comes to cognition in general, empiricism trumps rationalism, and the rationalist tendency to over-intellectualize should be strongly resisted. At the same time, sensorimotor++ – and the unified conceptual space theory of which it forms a part[40,36] – sees conceptual cognition14 as a special case: one in which both perspectives are needed and neither can be held as primary or ultimately correct. Concepts are both something that conceptual agents do and things they possess. They are both – as I suggested in the introduction – building blocks and abilities, and settling on one to the exclusion of the other is to miss not just half the story but – in a very real way – the whole thing.

14 Remember that, on the unified conceptual space theory, conceptual cognition is taken to be only one small part of cognition.
This focus on processes over abstract objects, fluid and dynamic descriptions over reifications, is clearly evident in many places through the 2001 paper, but particularly in O’Regan and Noé’s description of qualia [34, p. 960]:

Qualia are meant to be properties of experiential states or events. But experiences, we have argued, are not states. They are ways of acting. They are things we do. There is no introspectively available property determining the character of one’s experiential states, for there are no such states. . . . Experience is something we do and its qualitative features are aspects of this activity.

The problem with qualia, according to sensorimotor++ and the unified conceptual space account, is that all of experience is meant to be structured from individuable quales. Instead, as described in [40,36], all experience – for the conceptual agent – is better understood as a mixture of conceptual and non-conceptual content (see Footnote 7); and while the conceptual content may be – indeed, probably is – appropriately structured, the non-conceptual content need not be.

4 The Sensorimotor++ Account

Beyond these differences of emphasis and perspective, the sensorimotor++ account, as one component of the unified conceptual space theory, adds two key ingredients to the 2001 version of the sensorimotor account: namely, an account of salience centered on the agent (albeit highly preliminary) and – for conceptual agents – a representationally disposed cognitive mechanism to actualize it. Sensorimotor++ is sensorimotor engagements plus somatic and other bodily information (per Damasio [7] and Morse and Ziemke [29]) plus (appropriately qualified) representational language, as situated within a conceptual-spaces-based framework.

4.1 From Meaninglessness to Salience

Like the first airplane designers attempting flight, sensorimotor cognition faces a fundamental difficulty getting itself off the ground. Unless meaning is there from the very beginning, one is compelled to ask where does meaning come from? Unless one believes that (even abstract) cognition consists in nothing more than specific sensorimotor engagements – as per e.g. Gallese and Lakoff [12] (see Section 4.2) – one faces a difficulty in how one moves beyond specific sensorimotor engagements: i.e., how one generalizes to the sensorimotor profiles one needs to explain affordances, nevermind abstract conceptual thought.

As said in Section 3.2, an improved sensorimotor account needs to give a key role to emotional affect and the somatosensory system more broadly. Without that, sensorimotor++ argues, it cannot offer an adequate (or perhaps any) account of salience centred on the agent.\footnote{One could argue, of course, that it is not the role of a sensorimotor account to explain salience, that that should be left to another theory.} If the cognitive agent is not somehow
predisposed to find certain aspects of the environment salient and others not, there will be nothing to guide or structure its sensorimotor engagements. If the conceptual agent is not somehow predisposed to seek out certain patterns in its environment and disregard others, there will be nothing to guide its developing conceptual frameworks down one or another path in the way that Gärdenfors describes [13, p. 221]:

The prime problem is that the information received by the receptors is too rich and unstructured. What is needed is some way of transforming and organizing the input into a mode that can be handled on the conceptual or [according to Gärdenfors, more abstract] symbolic level. This basically involves finding a more *economic* form of representation: going from the subconceptual to the conceptual level usually involves a *reduction of the number of dimensions* that are represented.

Such salience for the conceptual agent must, seemingly, be grounded in something far more basic: most likely, something that applies to all cognitive agents. Here, *sensorimotor* does not have much to offer beyond a promissory note and an appeal to the attempt by many in the enactive community to ground ‘minimal’ salience in the survival of the organism (see e.g. [52]); what is salient is what enables the organism to survive. As the agent becomes more complex and develops a somatosensory system, that system then comes to play a key role.

As I suggested in Section 3.2, talk of ‘inside’ and ‘outside’ with respect to cognitive phenomena are unhelpful at best, misleading at worst; so it is with the familiar distinction between the five ‘external’ senses of taste, touch, smell, sight, and hearing, and the ‘internal’ senses of interoception (awareness of bodily states) and proprioception (awareness of bodily positions and movements).\(^\text{16}\)

Certainly, on the perspective where the observer is pushed into the background (if not ignored altogether), it makes sense to talk of how the different sensory modalities are integrated in the brain; but on the competing perspective, where the observer is in the foreground and one’s focus is on mind more than world, then it makes much more sense to talk about how one starts from undifferentiated experience, which then gets divided out into the various modalities and – for the self-reflective conceptual agent – subsequently conceptualized into realms of ‘internal’ self and ‘external’ non-self/other/world.\(^\text{17}\)

### 4.2 From Protoconcepts to Concepts

As noted in the previous section (see also Footnote 4), on Gallese and Lakoff’s account [12], all concepts – even the most abstract ones – are nothing more than specific sensorimotor engagements, albeit with parts of those engagements

---

\(^\text{16}\) This distinction is clearer in some disciplines than others; as one of the anonymous reviewers of this paper noted, in biology, proprioception and touch are often seen to be closely related.

\(^\text{17}\) I owe this point to psychologist Marek McGann, whom I have heard to make it in several conference presentations.
suppressed; no additional cognitive mechanism is required. The example they make the most reference to is the concept of grasp. All that the concept amounts to, on any given usage, is a specific occasion of (physical) grasping, even if one does not physically carry it out: that is, the concept is no more nor anything other than a variation on the familiar physical action.

As I have suggested, such an approach seems inadequate for cognition in general; but, when it comes to conceptual cognition, its shortcomings are particularly striking. Here, one might well accuse Gallese and Lakoff of having chosen a deliberately 'concrete' concept; yet, like the gaping differences between so-called mental imagery and pictures such as those hanging on the wall, the concept of grasping – which allows grasping ideas and intentions as much as door handles or hammers – is, in many ways, not like an 'actual' physical grasping at all. In particular: to the extent that one’s concept of grasping is a representation of grasping – and I have argued that, for the conceptually reflective agent, all concepts take on this representational aspect – the representation may have as much, or as little, to do with the represented as e.g. a representation of a dog, such as a painting of a dog, has to do with an actual dog. What goes for grasping goes for all the more abstract concepts like procrastination or ennui or the concept of concept itself. In keeping with the sensorimotor account, sensorimotor++ and the unified conceptual space theory agree that, certainly, sensorimotor engagement is necessary to the foundation of even the most abstract of concepts; at the same time, it is not sufficient. Pace Gallese and Lakoff, some additional cognitive mechanism is required.

As the quote from Gärdenfors suggests, one of the key roles of concepts is to simplify: precisely to distance oneself from the world in order to better understand it. Concepts abstract away from the moment, from the particulars of context, from – pace Gallese and Lakoff and Berkeley – any particular application of them. They allow the agent to step back from the immediacy of their sensorimotor engagement; in which case, some additional mechanism or mechanisms is required, to actualize the path from meaninglessness to salience.

I propose that the unified conceptual space theory provides this ‘representationally disposed’ mechanism, described through its algorithm or ‘recipe’ for constructing and de-constructing concepts. Like Gallese and Lakoff’s account, it rejects a prior ontological class/instance distinction; but, whereas Gallese and Lakoff (and Berkeley) reject classes in favour of instances, it ultimately rejects instances in favour of classes: to wit, any concept taken as a specific instance of something can contrastingly be understood as a class of yet more specific instances: so e.g. dog is a particular animal and, at the same time, a class of various breeds; while my dog Fella is a particular dog and, at the same time, a class of all my experiences and interactions with Fella.

One need not go so far as Fodor’s radical nativism [9,11,23], by which most concepts are innate, to allow that something innate is needed to kickstart

---

18 Fodor’s target is lexical concepts, which, for Fodor, means most concepts.

19 What precisely Fodor means by ‘innate’ is a subject of some controversy; for a good account of what he probably means, see [43, p. 230].
the endless cycle of concept acquisition and application (see Section 3.1). What the unified conceptual space theory offers – rather more modestly than radical nativism! – is a small set of innate protoconcepts (or, if you will, protoconceptual abilities), along with the conjunctive and disjunctive connectors needed to bind them together or progressively partition them into sub-concepts. These protoconcepts are suggested to consist of proto-object, proto-property, and proto-action/event, corresponding roughly to the grammatical categories, in English, of noun, adjective/adverb, and verb. They are not true concepts because they fail to meet all the usual desiderata of concepts as listed in the introduction: in particular, they are too few in number to be, of themselves, productive; and, being innate and therefore passively given to the agent, they are not under the agent’s endogenous control. By one means or another, we seem predisposed to encounter the world and structure our understandings of it in terms of (concrete and abstract) objects, (concrete and abstract) happenings, and the properties of both – to the extent that it seems impossible to imagine encountering the world any other way. Given the appropriate environment and the appropriate sensorimotor engagements with that environment, these protoconcepts and connectors can, or so the unified conceptual space theory claims, give rise to the most richly structured of conceptual frameworks. In the language of conceptual spaces theory, they do so through the progressive partitioning of an initially minimally partitioned protoconceptual space (see the left-most illustration in Figure 1) – by extracting from perception patterns, patterns of patterns, and patterns of patterns of patterns, whilst disregarding or discarding others.

5 Conclusions

Like the original 2001 version and many of the subsequent formulations of sensorimotor theory, and in keeping with the enactive tradition, sensorimotor++ attempts to resolve the seeming explanatory gap between mind and body, subjective experience and objective world; but it does so in a different way from the 2001 paper. What it comes down to, according to sensorimotor++ and the unified conceptual space theory of which it is a part, is two different perspectives we move constantly – and for the most part non-reflectively – between. In the one, the observer is front and centre; in the other, the observer is pushed into the background or, seemingly, eliminated altogether. As the discussion has meant to imply, we cannot resolve the two perspectives into one, unified perspective because of our position within the explanatory loop; resolving the tension between perspectives would require stepping outside the loop.

20 For more on protoconcepts, as described within the unified conceptual space theory, see [39,36,40]. A yet more detailed account of protoconcepts is intended for an upcoming paper.

21 I can, and do, choose to remain agnostic about what precisely is meant by ‘innate’: whether these protoconcepts are directly or indirectly specified in the genes or develop in the womb or something else again.
Sensorimotor engagement – as it is commonly understood, in terms of the so-called 'external' senses – is necessary but not sufficient for understanding either cognition in general or (my interest) conceptual cognition in particular. Incorporating interoception, proprioception, and emotion helps resist the extremes of externalism without falling into the trap of internalism. It strikes what I believe to be the ideal – if highly tensioned – balance between the two whilst suggesting how an account of salience might be grounded. Adding what I call a representationally disposed cognitive mechanism – described, in the unified conceptual space theory, by an algorithm for concept formation and evolution – actualizes that account. Representation – and therefore, pace Hutto, mental content – is a necessary part of any account of concepts and conceptual cognition that in any way tries to be complete, if only because, whenever we stop to reflect on our concepts as concepts – representations are what we find.

References


29. Morse, A., Ziemke, T.: The somatic sensory hypothesis (2010), unpublished manuscript


38. Parthemore, J.: Representations, symbols, icons, concepts... and why there are no mental representations. Proceedings of the Seventh Conference of the Nordic Association for Semiotic Studies 6-8 May 2011 (2013), forthcoming


