THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Semantic inferentialism as (a form of) active externalism

OPEN ACCESS

Semantic inferentialism as (a Form of) active externalism

Adam Carter[1] · James H. Collin[1] · Orestis Palermos[1]

Abstract Within contemporary philosophy of mind, it is taken for granted that externalist accounts of meaning and mental content are, in principle, orthogonal to the matter of whether cognition *itself* is bound within the biological brain or whether it can constitutively include parts of the world. Accordingly, Clark and Chalmers (Analysis 58(1):7–19, 1998) distinguish these varieties of externalism as 'passive' and 'active' respectively. The aim here is to suggest that we should resist the received way of thinking about these dividing lines. With reference to Brandom's (1994, 2000, Inquiry 47:236–253, 2008) broad semantic inferentialism, we show that a theory of meaning can be at the same time a variety of active externalism. While we grant that supporters of other varieties of content externalism (e.g., Putnam 1975 and Burge (Philosophical Review 95:3–45, 1986) can deny active externalism, this is not an option for semantic inferentialists: On this latter view, the role of the environment (both in its social and natural form) is not 'passive' in the sense assumed by the alternative approaches to content externalism.

Keywords Active externalism · Extended cognition · Semantic inferentialism

1 Passive and active externalism

Suppose Hilary is presently thinking a thought he would naturally articulate with the expression 'Water is wet'. Externalists about mental content (e.g. Kripke (1980),

---

✉ Adam Carter
j.adam.carter@ed.ac.uk

James H. Collin
james.collin@ed.ac.uk

Orestis Palermos
s.o.palermos@ed.ac.uk

[1] School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Dugald Stewart Building, Edinburgh EH8 9AD, UK

🕿 Springer

Putnam (1975), Burge (1986)) insist that what Hilary counts as thinking *about*—that is, the content of the thought Hilary is currently having—is at least *partly* a matter of Hilary's physical or social environment, and thus not entirely a matter of how things stand internal to Hilary.[1]

For example, according to Putnam (1975), when one utters that 'Water is wet', one means that H_2O *is wet* if all of one's previous interactions with what looked, felt and tasted like water were in fact interactions with H_2O. Accordingly, the content expressed by sentence-token utterances containing natural kind terms is individuated externally by one's physical environment (and *mutatis mutandis* for a variety of non-natural kind terms such as 'sofa', 'brisket' and 'red'). According to Burge (1986), however, there is no reason to believe that this external process of individuating mental content is necessarily restricted only to varying features of one's physical environment. Instead, we can come up with a version of social externalism by running roughly the same argument with respect to varying features of one's socio-linguistic environment. For example, *pants* in the States means trousers whereas, in UK, underwear. Thanks to a number of such and similar thought experiments against the suggestion that mental content supervenes solely on internal states, content externalism is nowadays widely taken for granted.[2]

Not so for the comparatively more radical thesis that Clark and Chalmers (1998) have labelled *active externalism*. According to active externalism, parts of the world can do much more than individuate mental contents.[3] Parts of the world—for example, a notebook—can partly *constitute* cognition. This is, of course, a striking thesis. A commonly discussed variety of active externalism is the hypothesis of *extended cognition*, according to which *cognitive processes* are realizable by brain, body and *world*. To make the idea of extended cognition more concrete, consider the following widely referenced case from Clark and Chalmers (1998):

> *Otto*: Otto suffers from Alzheimer's disease, and like many Alzheimer's patients, he relies on information in the environment to help structure his life. Otto carries a notebook around with him everywhere he goes. When he learns new information, he writes it down. When he needs some old information, he looks it up. For Otto, his notebook plays the role usually played by a biological memory.[4]

Of course, a common knee-jerk reaction to the case of Otto is to insist that while Otto relies on his notebook in a way that is *functionally akin* to the way that an agent

---

[1] An associated commitment of content externalism will be that Hilary and Hilary's intrinsic duplicate, Hilary*, can be thinking *about* very different things. See, for example, Dretske (1996) and Lycan (2001) for a stronger kind of mental content externalism according to which conscious mental states (with phenomenal properties) are externally individuated. This thesis is referred to as phenomenal externalism or qualia externalism.

[2] *Cf.* Bach (1987) and Crane (1991) for some notable exceptions.

[3] Active externalism comes in three central forms—roughly in order of perceived radicality: the hypothesis of extended cognition, the extended mind thesis and the distributed cognition thesis. For more details on the distincion between these three forms of active externalism see AuthorA, Author C & Co-Authors. Notable proponents of versions of active externalism include Clark and Chalmers (1998; Clark 2008; Hutchins 1995; Menary 2006a, b; Theiner 2011; Wilson 2000, 2004).

[4] It should be noted, however, that the case of Otto was originally meant to motivate the extended mind thesis (and not the extended cognition hypothesis).

with a normally functioning memory (call her Inga) relies on her biological memory, Inga's biological memory is part of the cognitive process she employs, but Otto's non-biological notebook can't be part of his.

Clark responds that such a denial exhibits an objectionable sort of 'bioprejudice'.[5] In order to guard against such bioprejudice, Clark & Chalmers insist that judgments about what counts as part of a cognitive process should be guided by what they call the *parity principle*:

> PARITY PRINCIPLE: If, as we confront some task, a part of the world functions as a process which, were it to go on in the head, we would have no hesitation in accepting as part of the cognitive process, then that part of the world is part of the cognitive process. (Clark and Chalmers 1998, p. 8)

With reference to the parity principle, notice that since Otto's notebook is functionally isomorphic with Inga's biological brain, we must include Otto's notebook as part of his cognitive process so long as we count Inga's biological brain as part of hers.

While active externalists debate amongst themselves *which* parts of the world can play a role in cognitive processes, and in virtue of which criteria, active externalists are unified in insisting that the role one's environment plays *vis-a-vis* cognition is not *merely*, as most content externalists tells us, a passive content-individuating role, operating in the distal, historical background. Clark & Chalmers have thus distinguished this kind of *active* externalism—as a striking thesis about the constituents of cognition—from what they call the *passive* externalism about content that has been inherited from Putnam- and Burge-style thought experiments.

Passive (content) externalism and active externalism thus seem to diverge in both the scope of the claim they are making, as well as in the kind of philosophical motivations undergirding each. It is therefore natural to think of these varieties of externalism as disconnected in the following respect: theses about meaning and mental content are supposed to be orthogonal to the matter of whether cognition is extended in the sense claimed by active externalism. More simply: semantic theses (e.g. Putnam 1975) are metaphsyically inert, in a way that what Clark & Chalmers are claiming, with reference to Otto and his notebook, is not.[6]

In this paper, our aim will be to challenge this standard way of thinking about semantic or content externalism as necessarily *passive*, thereby bringing into question the implicit assumption that content and active externalism are always distinct, in *any* of their existing formulations. In particular, we suggest that the arguments motivating the hypothesis of extended cognition may force one into viewing Brandom's semantic inferentialism as a variety of active externalism in the following way: The considerations in virtue of which external artefacts count as part of a cognitive process motivate, on the inferentialist program, a similar treatment of the parts of the world that feature (in the relevant ways) in the discursive practices that confer meaning on our words. If we are right, then, contrary to the received way of thinking about content and active externalism as simply orthogonal to each other, semantic theses can both have

---

[5] See, for example, Clark (2008).
[6] Cf. Chomsky (1995)'s suggestion that the results of the kinds of thought experiments that motivate content externalism are not useful in any scientific theory. See here also Lau and Deutsch (2014)

a strong impact on and be informed by the debate with regards to the *realisation* of cognition.

2 Extended cognition and CRC

To begin with, let us focus a bit more on cognitive extension, and specifically, a central problem it faces: on an implausibly lax construal of which parts of the world feature in cognitive processes, the proponent of extended cognition runs the risk of failing to save the differences between the cognitive processes themselves and the world external to them. In the literature on extended cognition, this general problem has been termed the 'cognitive bloat' objection to the hypothesis of extended cognition.[7] As Allen-Hermanson (2013, 792) puts it:

> If a notebook counts as part of one's mind, then why not the yellow pages, the internet, or even parts of the natural world that supply information and support cognition?

In order to avoid ruling-in too much of the world as constituting the cognitive process, then, the proponent of extended cognition needs a principled (and not too inclusive) explanation for what is ruled-in and ruled-out. A natural starting place here is to supplement the parity principle with Clark's common-sense functionalist 'glue and trust' criteria:

2.1 Clark's glue and trust conditions

1. 'That the resource be reliably available and typically invoked.'
2. 'That any information thus retrieved be more-or-less automatically endorsed. It should not usually be subject to critical scrutiny. […] It should be deemed about as trustworthy as something retrieved clearly from biological memory.'
3. 'That information contained in the resource should be easily accessible as and when required.'

Conditions (1–3) do seem to do the trick across a spectrum of cases.[8] A standard worry, however, is that they are ultimately still too inclusive. Regular access to a phonebook, after all, seems to, at least *prima facie*, satisfy (1–3). But to include ordinary agents' phonebooks as part of their *cognition* will be an unwanted inclusion for many who are happy to allow for Otto's notebook.[9]

---

[7] For variations of this objection see Adams and Aizawa (2001; Rupert 2004; Spaulding 2012). Cf. Kirchhoff (2013)

[8] Consider, for example, a case in which Tim uses an encyclopedia every now and then for reference purposes. Even though Otto's notebook passes (1–3), Tim's encyclopedia does not. For instance, the encyclopedia fails condition (1). While it is *accessible* to Tim, it not readily available in the sense that Otto's notebook is. While Otto treats his notebook as an extension of his agency, Tim treats the encyclopedia as something that (through his agency) can be helpfully consulted.

[9] See here Rupert (2004) and Palermos (2014).

Fortunately, however, for the proponent of extended cognition, there is an alternative way to motivate cognitive extension that is independent of the common-sense functionalism that underlies the parity principle and the 'glue and trust' criteria, and which seems to avoid the problem above. Specifically, it has been argued (Clark (2008), Chemero (2009), Froese et al. (2013) and (Palermos (2014)) that when an external element is non-linearly related to the agent's biological cognitive system, it can count as a constitutive part of an overall cognitive system that extends to include all the contributing parts.[10] The motivation for this is that, according to dynamical systems theory, these non-linear relations that arise out of continuously and reciprocally interacting parts on the basis of feedback loops from one part to the other give rise to an overall non-decomposable system that consists of all of them. Accordingly, when there are 'ongoing feedback loops' or 'continuous reciprocal causation' (CRC) between internal and external parts, the result is an extended cognitive system that consists of all of them.

Put another way, 'when we confront a recognizably cognitive process, running in some agent, that creates outputs (speech, gesture, expressive movements, written words) that recycled as inputs drive the process along', these inputs should count as constitutive parts of the overall cognitive process (Clark 2007, sec. 7).[11] Now, to see how this 'CRC criterion' may put the 'cognitive bloat' worry to rest, just consider that Otto's own inscriptions on his notebook continuously feedback to his bio-cognitive system. Otto has his notebook constantly open, going back and forth over its pages, looking up its contents as well as taking notes about almost every single detail that would normally be registered in his biological memory. Accordingly, his dense interaction with his notebook gives rise to a causal reciprocity that does not seem to be present in the phonebook (or any other similar) case.[12]

3 Semantic inferentialism and TOTE cycles

While keeping the above in mind (we will return soon), let us now turn to Brandom's analytic pragmatist project of explicating meaning in terms of *inference* (rather than representation). Whilst broadly Tarski (1944) theories explicate meaning in terms of reference or truth,[13] pragmatist theories of meaning, such as Brandom's semantic inferentialism, begin instead with the *practices* that institute these relations.[14]

---

[10] For more details, and the stronger claim that these non-linear relations are also necessary for cognitive extension, see Palermos (2014).

[11] As Wheeler (2001) notes, the causation at play in Clark's articulation of CRC 'involves multiple simultaneous interactions and complex dynamic feedback loops such that (i) the causal contribution of each component in the system partially determines and (ii) is partially determined by the causal contributions of large numbers of other components in the system…'

[12] For detailed discussions of the CRC criterion and how it helps with respect to the cognitive bloat worry and related objections to the hypothesis of cognitive extension see Palermos (2014).

[13] Donald Davidson is the most celebrated proponent of a Tarskian theory of meaning; Tarski himself did not attempt to put his axiomatic theory of truth to this use.

[14] As MacFarlane (2010, 81) puts it, 'One of the central themes of Brandom's work is that we should construct our semantic theories around material validity and incompatibility, rather than reference, truth, and satisfaction.' For some specific remarks on this point, see Brandom (1994, 91).

It should be no surprise that pragmatist approaches to meaning that are *inferentialist*, such as Brandom's, will—much more so than orthodox truth-conditional theories of meaning—rely on the notion of a *cognitive process*[15] in order to explicate meaning—and in particular, the cognitive process of *inferring*.[16] In order to appreciate this point, let's look at the basic contours of the inferentialist view.

According to semantic inferentialism, to grasp a proposition *p*, an interlocutor must be able to exercise a discriminative cognitive capacity[17]—*viz.*, the capacity to discriminate between:

(i). what other propositions one would be *committed* to, were one committed to *p*

(ii). what propositions one would be *entitled* to, were one committed and entitled to *p* and

(iii). what one would *not* be entitled to, were one committed[18] to *p* .[19]

Of course, the cognitive skill to navigate this network of inferential relations need not be either perfectly sound or perfectly complete, but unless an interlocutor can make the above kinds of discriminations at least to some significant extent, she cannot be said to grasp *p*. Furthermore, a commitment to *p* can be thought of as a 'node' in a wider inferential structure, and the *content* of *p* depends upon its position within the structure.

Now, two important points—the first with respect to the nature of these commitments/inferential nodes and the second with respect to their effects—are in order. Firstly, for Brandom, the connections of commitment and entitlement relevant to the content of a proposition involve not only inferential connections between other propositions but also connections to worldly noninferential *circumstances* and *consequences* of application.

To unpack this idea further, on one direction, empirical claims involve, in addition to connections between propositions, afferent connections with the world through perceptions: the input of appropriate stimuli induces reliable differential responses to the relevant kinds of objects. For example, the worldly presence of a rabbit commits one to *There is a rabbit*. Similarly, but to the opposite direction, some claims—paradigmatically, *normative* claims—involve, in addition to connections between propositions,

---

[15] We consider, in the next section—and in line with an anonymous referee's suggestion—whether there is any scope to deny that on Brandom's programme inferring is a cognitive process. At this point, however, our objective is just to put the view on the table.

[16] Brandom (1994, 134) remarks that the 'explanations of what it is in practice to treat inferences as correct as what ultimately license appeal to material proprieties of inference, which can then function as semantic primitives.'

[17] This exercise of a discriminative capacity required for grasping a proposition is part and parcel with Brandom's insistence that 'inferring is a kind of doing' (1994, 91).

[18] Brandom eschews the term 'belief', which he takes to be equivocal between a normative and a psychological reading, in favour of 'commitment', which he understands normatively.

[19] The inferential relation of commitment here is understood as an obligatory, deductive relation. So, for instance, a commitment to *Snow is white* involves a commitment to *There is something such that it is white*. Entitlement-preserving inferential relations are understood as non-obligatory, defeasible, inductive relations. Commitment and entitlement to *The ground is wet* involves *prima facie* entitlement to *It has been raining*, which can in turn be defeated by further entitled commitments such as *Someone has hosed the ground*. Incompatibility relations are, like commitment relations, monotonic; commitment to precludes entitlement to, e.g., a commitment to *Snow is white* precludes entitlement to *Snow is red* (assuming a commitment to snow's being monochromatic).

efferent connections with the world through action. Commitment to *Consuming meat is wrong* commits one to not eating the rabbit.

Second, and as the first point indicates, describing these activities—giving an account of what it is that a person is doing when she grasps a proposition or takes up a commitment—essentially involves both *normative* and *modal* vocabulary. On the one hand, normative vocabulary is required since agents are taking up *commitments* that bring with them *obligations* to take up other commitments. A person committed to *There is a rabbit* ought also to acknowledge a commitment to *There is an animal*, and is obliged to not assert that *There are no animals*.

'Ought', 'obliged' and other cognate expressions are distinctively *normative*. On the other hand, *modal* vocabulary is required to describe or make explicit the nature or content of the commitments that persons take up. Commitments are about the world, and worldly things have properties that are objectively incompatible with other properties. Being a negatively charged particle, for example, involves repelling other negatively charged particles. The content then of *That is an electron* can be made explicit by using modal claims such as *Were that object in a system with other negatively charged particles it would repel them*. In short, modal language must be used to describe the contents one is committed to—how one has taken the world to be; how one has represented the world—whereas normative language must be used to describe the *acts* one is committed to by holding that *p*. As Brandom puts it, normative and modal vocabularies respectively correspond to 'the subjective and objective poles of intentional relations' (Brandom 2008, 181).

Against this background, (and bearing in mind the discussion of CRC in §2), it will be helpful to make more precise the nature of these two kinds of commitments and how they relate to one another. Brandom starts by noting that, according to the pragmatist's guiding idea, the above commitments, which capture semantic relations between words and the world, are neither static nor passive, but are instead established on the basis of *practices*. Accordingly, Brandom advises, to understand meaning, 'don't look, to begin with, to the relation between representings and represented, but look to the nature of the *doing*, of the process, that institutes that relation[20]' (Brandom 2008, 177–8). This is because, on Brandom's pragmatist proposal, 'the most fundamental kind of intentionality (in the sense of directedness towards objects) is the *practical* involvement with objects exhibited by a sentient creature dealing skillfully with its world' (ibid., 178). It is this practical intentionality that is at the basis of, allows for the emergence of, and is in fact a generic case of, the more specific *semantic* intentionality exhibited in *discursive practices* of language-use.

Now, and with this point in view, we are finally in a position to see why semantic inferentialism might be a version of active externalism. According to Brandom, in order for a practice to qualify as a case of the above necessary form of practical intentionality, it must manifest a *Test-Operate-Test-Exit* (TOTE) cyclical structure of perception, performance, assessment of the results of the performance, and further performance. It must be, that is, 'a process or practice consisting of an open-ended sequence of feedback-governed performances' (ibid.).[21] And what this means—Brandom notes in a

---

[20] Our italics.
[21] For the original presentation of the TOTE cycle, see Miller et al. (1960). Brandom first applies the idea of TOTE cycles in Brandom 1994, 528–9, and expands this in a number of places in Brandom 2008.

manner that seems to directly echo the dynamical systems theory arguments for active externalism—is that such 'feedback-governed processes, practices, and abilities […] cannot be specified without reference to the changes in the world that are both produced by the system's responses and responded to within each loop in the TOTE cycle.' Accordingly:

> Feedback-governed practices are 'thick', in the sense of essentially involving objects, events, and worldly states of affairs. Bits of the world are *incorporated* in such practices, in the exercise of such abilities (ibid.).

4 Semantic inferentialism as (a Form of) active externalism

There's no denying the striking similarity between both the arguments for, and the conclusions of, semantic inferentialism and active externalism (the latter in the form of the dynamically motivated extended cognition hypothesis). In fact, given the points above, it is natural to view semantic inferentialism *as a form of* active externalism, according to which the inferential cognitive process in virtue of which words acquire their meaning extends to constitutively include aspects of the external world.

At this juncture a worry[22] might be raised about whether the semantic inferentialist of the Brandomian stripe can understand inferring as a *cognitive* process, as we have understood it here. The concern can be made out in the following way: Doesn't Brandom emphasise the role of the interpreter and the third person perspective in his account of meaning in such a way that the inferences we need to focus on in order to understand the meaning of our words are not really the cognitive processes of a cognitive system but merely a theoretical, interpretive device that philosophers of language have to use? On this view, our overall interpretation of a person or community's behaviour involves *attributing* inferences to them, but inference is not, in the end, a cognitive process.

We take it that understanding inferring as a cognitive process coheres well with Brandom's inferentialism. Though it is true that Brandom understands the concepts in normative terms and holds that norms are instituted by social practices, he steers clear of the view that conceptual contentfulness—and therefore inference, since contentfulness is understood in inferential terms—is *merely* a device of interpreting a person's overall behaviour. Brandom takes his account to be one of *original* (as opposed to *derived*) intentionality; an account of the sorts of activities that agents must engage in in order to *have* contentful mental states. Specifically, in comparing his view to accounts of derived intentionality, whereby "conceptual contents might be attached to states and performances […] by the stipulation of some interpreter who is independently able to deploy the concepts involved," (Brandom 2009, 177) he writes:

> 'This approach construes conceptual contents as consisting in the significance states and performances have within the behavioral economy of a system because of their interactions with each other and with the environment. According to this

---

[22] Thanks to an anonymous referee for raising this point.

sort of account, it is the activity of the system itself that establishes the conceptual contentfulness of the states it exhibits, the performances it produces, and the expressions it uses. In this way it is distinguished from the first sort of account, according to which conceptual contents are conferred by the activity of an external interpreter. [This] picture does not require that conceptual contentfulness derive from the antecedently conceptually articulated intentions and beliefs of another. Rather, in some sense the system itself implicitly takes or treats its own states as contentful, and thereby makes them so.'[23] (Brandom 2009, 178)

Making inferences, on Brandom's account, is therefore a *bona fide* cognitive process; indeed *the* cognitive process *par excellence*, since inference-making is the cognitive process constitutive of cognition. Grasping or understanding a concept, for Brandom, *consists in* having a practical mastery of the inferences it is involved in.[24] Using the example adverted to previously, a person with a mastery of the concept RABBIT will be disposed to make the inference from *X is a rabbit* to *X is an animal, X is not a prime* number, and so on. Though her practical mastery need not be perfect, possessing the concept RABBIT requires making sufficiently many and sufficiently accurate inferences of this sort, in appropriate circumstances. Inference-making on this picture is an activity that subjects engage in, and not merely something we attribute to people in order to interpret them.

A potential objection is worth considering at this point. The anticipated objection begins by cautioning that we should not be so fast in assuming that the inferences Brandom has in mind are cognitive inferences after all. Perhaps Brandom is not focusing merely on the inferences that take place within an individual's head, but instead on the *inferential practices* that take place within the relevant individual's community. In the above quoted passage, for example, Brandom includes a very interesting footnote (no. 1, p. 178) in order to point out "the possibility that the system whose doings one must consider as conferring genuine content may have to be a whole linguistic community rather than a single agent." If that's the system whose inferential practices we have to focus on, then it might seem dubitable that the relevant meaning-conferring inferential practices actually amount to cognitive processes that could, on the basis of CRC, be extended. In other words, if we have no reason to assume that the relevant inferential practices are *cognitive* practices, then the extended cognition hypothesis cannot apply to them.

While this may sound like a *prima facie* worrying objection, we contend that on closer consideration, it is not. We in fact have no good reasons to assume that the target inferential cognitive practices are not cognitive. Specifically, within cognitive science, there is no 'litmus test' on the basis of which researchers choose their subject matter. Instead, cognitive scientists focus on cognitive tasks that most people would intuitively like to call cognitive ones and then infer that whatever processes are constitutively involved in the systems that give rise to the relevant tasks are cognitive processes. According to Wilson (2001), for example, we may 'fix our ideas' on the basis of the

---

[23] See also (e.g.) Brandom (1994, 61): 'If the practices attributed to the community by the theorist have the right structure, then according to that interpretation, the intentional contentfulness of their states and performances is the product of their own activity, not that of the theorist interpreting that activity.'

[24] Brandom (1994, 89) gives a clear statement of this.

following incomplete but suggestive list of what may count as a psychological process or ability: "perception, memory, imagination (classical Faculties); attention, motivation, consciousness, decision-making, problem-solving (processes or abilities that are the focus of much contemporary work in the cognitive sciences); and believing, desiring, intending, trying, willing, fearing, and hoping (common, folk psychological states)" (Wilson 2001, p. 266). The only alternative to this approach would be to put forward a plausible 'mark of the cognitive' (Adam and Aizawa 2001, 2008, 2010) and then test whether it is borne by the processes cognitive scientists are interested in and class them accordingly. As a long-standing and ongoing debate demonstrates (Clark 2008, 2010; Menary 2006; Adams and Aizawa 2001, 2008, 2010; Ross and Ladyman 2010), however, the problem with this alternative is that there is no commonly accepted and non-problematic mark of the cognitive on offer. Therefore, following the current methodology within cognitive science, and since inferring is a paradigmatically cognitive process, it is by no means obvious how one could deny that cognitive practices are non-cognitive, even if they partly refer to inter- rather than intra-personal activities.

Accordingly, it is not implausible to assume that, under the appropriate conditions, meaning conferring practices can be extended along the lines suggested by the extended cognition hypothesis. Before we move on to one of Brandom's own examples in order to illustrate concretely how the meaning-conferring inferential process extends in practice, here is semantic inferentialism, in summary, for one last time. In Brandom's words (ibid., 183),

> a system counts as excersising *discursive* abilities, or engaging in discursive practices, hence as exhibiting *discursive* intentionality, insofar as the differential responsiveness of the system to the results of its own performances is essentially mediated by states whose functional role in the feedback process can be understood only by taking them to be *propositionally* contentful, that is, by specifying them in an *intentional* vocabulary—paradigmatically as involving the claim, belief, preference, or intention that *p*.

Moreover, according to this view, the TOTE nature of the underlying practical intentionality allows *parts of the world* to be essentially involved in the commitments that discursive creatures can undertake and which are, for the broad inferentialist, constitutive of meaning.

Brandom illustrates this with the following case. On the basis of previous TOTE cycles against the physical world and with each other, a group of language users employs the term 'acid*'. Accordingly, for them, the circumstances of application for acid* are when an item tastes sour, and its consequences of application are that an acid* will turn phenolphthalein blue—*viz.*, if an item tastes sour, these language users are entitled and committed to applying the term acid* to it, and given any item that has been termed an acid*, these language users are entitled and committed to its turning phenolphthalein blue. Acid* then is characterized by two conditionals:

(circ): If an item is sour then it is an acid*.
(cons): If an item is an acid* then it turns phenolphthalein blue.

And in employing the term 'acid*', this community of language users are implicitly endorsing these inferential commitments. Accordingly, if a member of the community finds an item that tastes sour but which turns phenolphthalein red, then she is faced with mutually incompatible commitments: she infers from the item's sour taste that it is an acid*, and from its being an acid* to its propensity to turn phenolphthalein blue, yet by exercising her capacities of perception she finds that the phenolphthalein is red. The cycle of perception, inference, action and perception has resulted in this incompatibility. Faced with the incompatibility, the agent is required to alter at least one of her commitments; she must—the force of this 'must' is normative–either give up the belief that the item is sour, the belief that the phenolphthalein is red, or else amend her concept of acidity* to accommodate her perceptions. This may take the form of restricting the circumstances of application, perhaps to *liquids* that are sour or liquids of a certain colour, or the consequences of application, perhaps to turning phenolphthalein blue only under particular circumstances, or even rejecting the concept acid* altogether as defective.

As this example demonstrates, then, there are at least three distinct ways in which interactions with the world are able to structure the content of our beliefs.

1. In a world-to-word afferent direction, we can expand our observations by drawing commitment- and entitlement-preserving inferential conclusions.
2. Again in a world-to-word direction, we can register any resulting materially incompatible commitments. And,
3. in a word-to-world efferent direction, which nevertheless constantly feeds back to both (1) and (2), we can repair any incompatibilities by modifying or relinquishing our commitments, or the concepts that link them inferentially (Brandom 2008, 185).

When this cyclical process is repeated enough times and we are sufficiently responsive to the way the world reacts to the application of our (discursive) practices by modifying them appropriately, the normative structure of our beliefs—the network of commitments and entitlements, what we are entitled or committed to infer from what—tracks the modal or nomic structure of the world. In other words, TOTEs allow the world to become incorporated to our inferential processes, to gradually infiltrate the meaning of our words, by affecting their future use and the shape of our discursive norms in a way that mimics some of the world's modal qualities. Even though this is not to say that our minds become a *mirror* of nature, this process allows them to embody a (pragmatic?) *understanding* of the world by tracking (some of) its modal or nomic properties.[25] Moreover, contrary to what Putnam and Burge assume, even when enough TOTE cycles between an individual and both her physical and social environment have been completed, such that her use of the relevant concept seems to be in no need for further amendments, the environment and the overall meaning-conferring process do not thereby become idle. On the contrary, TOTE cycles between an individual intentional agent and her (physical as well as social environment) are always in operation and have an effect on the meaning of our concepts, even when those cycles

---

[25] As well as cohering with active externalism, this sort of broad inferentialist account of meaning may have anti-sceptical benefits; cf. Collin (2013).

reliably result in success. In cases of success, these cycles have the imperceptible effect of further strengthening the meaning of our concepts, or they may further expand their meaning, if the relevant concepts are used in unprecedented (physical or discursive) contexts.

So how does this analysis compare with Putnam and Burge's passive externalism? Obviously, from the point of view of any and all individuals together, the way one's community responds to one's application of a word and how one calibrates one's use of that word in return, and so on, indicates that the social environment plays a much more active role in contributing what is meant by a given word than Burge's social externalism would predict. And contrary to Putnam-style externalism, what is meant by acid* does not depend merely on the role one's physical environment plays *in the distal historical background*—on what, in that environment, *happens* to be the microstructure (or the essential nature) of anything that the language user refers to by the word "acid*". Instead, on semantic inferentialism, one's physical environment is very much in the foreground, actively *driving*—responding to and restraining—the inferential process that confers meaning to a given word, and to our discursive practice as a whole.

Having thus clarified the differences between Putnam and Burge's passive externalism and the more active form of Brandom's content externalism, it is finally worth considering a worry concerning a possible clash between semantic inferentialism and active externalism. Specifically, a possible objection against our argument is to note that active externalism is a form of what is known as 'vehicle externalism'—according to which the vehicles of cognitive processes and states such as beliefs might be located outside of the organismic agent—whereas Brandom (2004) refers to the contents of beliefs—and more specifically what he calls 'doxastic commitments'—as being "vehicleless". Therefore there seems to be a discrepancy between active externalism and semantic inferentialism in the sense that the former is concerned with the vehicles of cognition whereas the latter assumes that cognitive states have no vehicles.[26] Even though this may seem like a worrying objection, we think it can be easily overcome if we, following Brandom, disambiguate between the different senses of 'beliefs' and their corresponding 'contents'. Here is where the distinction comes into play:

> [There are] two ways in which we can talk about *belief*. In one sense, one believes just what one takes oneself to believe, what one is prepared to avow or assert. In another sense, one believes, willy-nilly, the consequences of one's beliefs. […] The sense of belief in which one is taken actually to believe what one ideally *ought* to believe (at least given what else one believes), call it *ideal* or *rational* belief, can conflict with the sense of belief for which avowal is authoritative. (Brandom 1994, 195)[27]

Brandom avoids this ambiguity by using a univocal technical term 'doxastic commitment' to refer to all the commitments one is prepared to acknowledge *and* all the commitments that follow from these (acknowledged or unacknowledged). When

---

[26] We are thankful to an anonymous referee for pressing this objection. Addressing this worry has been particularly helpful in further spelling out the sense in which we take Brandom's semantic inferentialism to be a form of active externalism.

[27] See also Brandom 2000, 174.

speaking "officially" Brandom avoids the term 'belief' altogether in favour of unambiguous talk of doxastic commitment. So when Brandom talks of intentional states as being vehicleless, he is actually talking of doxastic commitments, both acknowledged and unacknowledged; *viz*. 'belief' in the *ideal* sense. The motivation for thinking of these as vehicleless is clear: what we are committed to—the consequences of the commitments we acknowledge—often outruns anything we might be disposed to acknowledge. Those who acknowledge a commitment to the axioms of Peano arithmetic will not (unless omniscient) acknowledge the infinitely many consequences of those axioms. These commitments are not grounded in any actual, concrete and worldly vehicles. It's a different story when it comes to 'belief' in the more down-to-earth, non-ideal sense: the sense in which avowal is authoritative, and one only 'believes' what one is prepared or disposed to avow. 'Beliefs' in this sense *are* vehicled, for the simple reason that there has to *be* something which *provides* the disposition to avow (even if this something is a system which extends beyond the boundaries of an individual organism).[28] It is this sense of 'belief' with which active externalists such as Clark are working; so the way in which 'beliefs' (non-ideal/dispositions to acknowledge commitments) are vehicled in 'vehicle externalism' does not clash with the way in which 'beliefs' (ideal/commitments both acknowledged and unacknowledged) are not vehicled in Brandom's picture.[29,30]

---

[28] See chapters 1 and 5 of Brandom 2015 for an account of his own commitment to a broadly Aristotelian metaphysics, and to dispositionalism.

[29] We should not be misled by Brandom's claim that he does not 'officially believe in beliefs' (Brandom 2000, 174). He does not mean that he does not believe in there being creatures with dispositions to acknowledge (some of) their commitments, which is what the foregoing requires. Indeed, he *cannot* mean that, since there being creatures with dispositions to acknowledge (some of) their commitments is part of his own story.

[30] How does this play out with respect to Brandom 2004, where Brandom claims that beliefs are vehicleless? There Brandom has in mind beliefs in the ideal, normative sense (2004, 240) and his argument that beliefs are non-vehicled hinges precisely on this way of understanding beliefs. The crux of the argument involves the need to understand why acknowledging a commitment regarding one's *own* beliefs (e.g. 'I believe that Kant lived in Köningsberg') involves undertaking commitments regarding things in the wider world (e.g. that Kant lived in Köningsberg). As Brandom puts it there, we require: "an explanation of why some event's befalling a non-intentionally specified vehicle – say, an inscription being placed in the 'belief-box' – should *amount to endorsing* a claim about how things are, to *undertaking* a *commitment* to things outside the believer being thus-and-so. To take this question seriously is to ask in what sense of 'inner' beliefs can coherently be understood to be inner, given that believing includes adopting a normative stance or attitude towards the external world – in particular, a commitment to the truth of a proposition (2004, 244)" If by '*amount to endorsing*' Brandom means *is the same as endorsing*, then Brandom must be in error; and for reasons that he himself gives. The claimables 'I believe that Kant lived in Königsberg' and 'Kant lived in Königsberg' are not equivalent (as Brandom points out). The first is entailed by 'Everyone believes that Kant lived in Königsberg' and the second it not (ibid.). Endorsing a particular claim p is not the same as endorsing a distinct claim not equivalent to p. For this reason, we strongly suspect this is *not* what Brandom means. Yet, it is certainly the case that endorsing 'I believe that Kant lived in Königsberg' involves *undertaking a commitment* to the claimable 'Kant lived in Königsberg'. There's nothing especially mysterious about this: endorsing a claim of the form 'I believe p' involves *attributing to oneself* a commitment to p – and that is enough to become committed to p (and all the subsequent commitments that follow from it). What all this shows, according to Brandom, is that we can't understand beliefs merely in terms of what goes on in the skull; beliefs have to be understood in terms of commitments to how things stand outwith the skull. It is in this sense that beliefs are vehicleless, as our earlier example of Peano arithmetic illustrated. But notice that all this turns on the *normative* conception of belief: Brandom is talking specifically about *undertaking commitments* to how things stand with the outside world. To reiterate then, the sense in which Brandom insists beliefs are 'vehicleless' simply is not at odds with the sense in which active externalists such as Clark are 'vehicle externalists'.

Though it is not essential to our claims here, we suspect that even doxastic commitments in the ideal sense are not *in principle* vehicleless but only contingently so. Specifically, the difference between non-ideal beliefs and ideal 'doxastic commitments' seems to be that the former have been acknowledged on the basis of existing (past or ongoing) TOTE cycles run by some agent or linguistic community. Therefore they are vehicled. On the contrary, doxastic commitments are ideal in the sense that even though there are possible TOTE cycles they could arise from, no such cycles have so far been run and this absence renders them vehicleless. Such vehiclelessness, however, is only a contingent matter of fact: 'Doxastic commitments' are only vehicleless, in the sense that there are no *actual* TOTE cycles to bear them as their vehicles, but there still are *potential* (i.e., ideal) TOTE cycles that *would* carry them as their vehicles, were those commitments actual and not ideal ones. So even though there are commitments that are vehicleless, this is not because they couldn't have vehicles were they to be actual, but only because—as a matter of fact—they have not been made actual *yet*.

5 Concluding remarks

To sum up, within contemporary philosophy of mind there is a common presupposition that all varieties of externalism about meaning commit one at most to a kind of passive externalism that can be distinguished from the more radical kind of active externalism embraced by Clark & Chalmers. In other words, content and active externalism are supposed to be orthogonal to one another, in *any* of their possible formulations. In this paper we reject this claim.

If one accepts active externalism in the form of the dynamically motivated extended cognition hypothesis, then broad semantic inferentialism is itself an active externalist account of meaning. Of course, the very idea of an active externalist account of meaning stands in conflict with the standard way of thinking about the distinction between meaning (or content) and active externalism: Contrary to active externalism, semantic theses are supposed to have absolutely no effect in the way we think about the *constituents* of cognition. But the argument, as we saw, is not hard to make.

Starting with the assumption that one is sympathetic to the extended cognition hypothesis, one will offer an *explanation* of what counts as part of a cognitive process, with reference to something like the CRC (on the basis of feedback loops) condition as a way to fend off cognitive bloat.[31] But, then, by turning to semantic inferentialism, one will notice that some of the feedback loops exhibited in TOTE cycles (in virtue of which parts of the world get incorporated in the meaning conferring inferential process) are going to *de facto* satisfy—at least with some considerable overlap—the CRC condition required for cognitive extension.

Note that we are not here claiming that CRC and TOTE are equivalent. Rather, the claim is just that, as we saw in §4, there can be considerable overlap. And for our argument to go through, we need only that there be *some* overlap. This is sufficient to establish that there is a clear sense in which (at least for the proponent of extended

---

[31] Notice that the argument we are here summarizing works even if the CRC condition is taken to be only sufficient for cognitive extension.

cognition), a thesis about meaning will itself be a kind of *active* externalism. And, of course, if that's right, then we need to rethink our explanations for why, precisely, meaning externalism has so far been supposed to be tantamount to passive externalism or, conversely, why meaning and active externalism are always supposed to come apart.[32]

References

Adams, F., & Aizawa, K. (2001). The bounds of cognition. *Philosophical Psychology, 14*(1), 43–64.

Adams, F., & Aizawa, K. (2008). *The bounds of cognition*. Blackwell Publishing Ltd.

Adams, F., & Aizawa, K. (2010). Defending the bounds of cognition. In Menary (Ed.), *The extended mind*. Cambridge: MIT Press.

Allen-Hermanson, S. (2013). Superdupersizing the mind: extended cognition and the persistence of cognitive bloat. *Philosophical Studies*. Springer, 1–16.

Bach, K. (1987). *Thought and reference*. Clarendon Press Oxford.

Brandom, R.B. (1994). *Making it explicit: reasoning, representing, and discursive commitment*. Harvard University Press.

Brandom, R.B. (2000). *Articulating reasons: an introduction to inferentialism*. Harvard University Press.

Brandom, R. B. (2004). From a critique of cognitive internalism to a conception of objective spirit: reflections on Descombe's anthropological holism. *Inquiry, 47*, 236–253.

Brandom, R. B. (2008). *Between saying and doing: towards an analytic pragmatism*. Oxford: Oxford University Press.

Brandom, R.B. (2009). *Reason in philosophy*. Harvard University Press.

Brandom, R.B. (2015). *From empiricism to expressivism: Brandom Reads sellars*. Harvard University Press.

Burge, T. (1986). Individualism and psychology. *Philosophical Review, 95*, 3–45.

Chemero, A. (2009). *Radical embodied cognitive science*. MIT press Cambridge.

Chomsky, N. (1995). Language and nature. *Mind, 104*(413), 1–61.

Clark, A. (2007). Curing cognitive hiccups: A defense of the extended mind. *The Journal of Philosophy, 104*, 163–192.

Clark, A. (2008). *Supersizing the mind: embodiment, action, and cognitive extension: embodiment, action, and cognitive extension*. Oxford University Press.

Clark, A. (2010). Coupling, constitution, and the cognitive kind: A reply to Adams and Aizawa. In Menary (Ed.), *The extended mind*. Cambridge: MIT Press.

Clark, A., & Chalmers, D. (1998). The extended mind. *Analysis, 58*(1), 7–19.

Collin, J.H. (2013). "Semantic Inferentialism and the Evolutionary Argument Against Naturalism." *Philosophy Compass, 8*(9), 846–856.

Crane, T. (1991). All the difference in the world. *The Philosophical Quarterly, 41*((162). JSTOR), 1–25.

Dretske, F. (1996). Phenomenal externalism or if meanings Ain't in the head, Where Are Qualia? *Philosophical Issues, 7*(JSTOR), 143–58.

Froese, T., Gershenson, C., & Rosenblueth, D.A. (2013). The dynamically extended mind. Retrieved http://arxiv.org/abs/1305.1958.

Hutchins, E. (1995). *Cognition in the wild*. MIT Press.

Kirchhoff, M.D. (2013). Extended cognition & the causalconstitutive fallacy: in search for a diachronic and dynamical conception of constitution. *Philosophy and Phenomenological Research, 87* (2).

Kripke, S.A. (1980). *Naming and necessity*. Harvard University Press.

Lau, J, & Deutsch, M. (2014). Externalism about mental content. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*, 2014. http://plato.stanford.edu/entries/content-externalism/.

Lycan, W.G. (2001). The case for phenomenal externalism. *Noûs* 35 (s15). Wiley Online Library: 17–35.

MacFarlane, J. (2010). Pragmatism and Inferentialism. *Reading Brandom: On "Making It Explicit"*, 81–95.

Menary, R. (2006). Attacking the Bounds of Cognition. *Philosophical Psychology, 19*(3). Taylor & Francis: 329–44.

Menary, R. (2006b). Attacking the bounds of cognition. *Philosophical Psychology, 19*(3), 329–344.

Miller, G.A., Galanter, E., & Pribram, K.H. (1960). *Plans and the structure of behavior. Holt, Rinehart and Winston*. Inc., New York.

Palermos, S. O. (2014). Loops, constitution, and cognitive extension. *Cognitive Systems Research, 27*, 25–41.

Putnam, H. (1975). The Meaning of 'Meaning'. *Minnesota Studies in the Philosophy of Science, 7*. University of Minnesota Press: 131–93.

Ross, D., & Ladyman, J. (2010). The alleged coupling-constitution fallacy and the mature sciences. In Menary (Ed.), *The extended mind*. Cambridge: MIT Press.

Rupert, R. D. (2004). Challenges to the hypothesis of extended cognition. *The Journal of Philosophy, 101*((8). JSTOR), 389–428.

Spaulding, S. (2012). Overextended Cognition. *Philosophical Psychology, 25*(4). Taylor & Francis: 469–90.

Tarski, A. (1944). The semantic conception of truth: and the foundations of semantics. *Philosophy and Phenomenological Research, 4*(3), 341–76.

Theiner, G. (2011). *Res cogitans extensa: a philosophical defense of the extended mind thesis*. Peter Lang.

Wheeler, M. (2001). Two threats to representation. *Synthese, 129*(2). Springer: 211–31.

Wilson, R. A. (2000). *The mind beyond itself*. New York: Oxford University Press.

Wilson, R. (2001). Group-level cognition. *Philosophy of Science, 68*(3), 262–273.

Wilson, R.A. (2004). *Boundaries of the mind: the individual in the fragile sciences-cognition*. Cambridge University Press.