Theoria, 2026, 41(3), 396-433
https://doi.org/10.1387/theoria.27610

THEORIA
ISSN 0495-4548 – eISSN 2171-679X

Illusionism as a theory of mental content

(El ilusionismo como teoría del contenido mental)

Marco Facchin*
University of Antwerp
https://orcid.org/0000-0001-5753-9873

Keywords

Mental content
Illusionism
Mind-body problem
Eliminativism
Intentionality
Radical enactivism

ABSTRACT: Analytic philosophers of mind seemingly take the hard problem of consciousness to be the only aspect of the mind-body problem that has not been in principle solved. But phenomenal consciousness is not the only puzzling and mysterious mental phenomena. The capacity of our minds to have contents — to "be about" or "of" objects, and to represent them — is also puzzling and mysterious. And whilst, from the '80s on, it was largely assumed that mental contents could have been easily naturalized, the repeated and systematic failures of naturalistic theories of content invite us to search for an alternative approach. This paper aims to introduce one such alternative approach, which I dub content illusionism. Content illusionism revolves around two main claims. One is an eliminative claim, according to which mental contents do not exist. The other is the illusion claim, according to which there is an illusion of mental content: contents appear to exist and be "carried around" by our mental states. The paper offers a general presentation of content illusionism, explaining its core claims, showing that it is a compelling philosophical view of content, and defending it from some intuitive objections.

Palabras clave

Contenido mental
Ilusionismo
Problema mente-cuerpo
Eliminativismo
Intencionalidad
Enactivismo radical

RESUMEN: Los filósofos analíticos de la mente parecen dar por sentado que el problema difícil de la conciencia es el único aspecto del problema mente-cuerpo que aún no ha sido resuelto. Sin embargo, la conciencia fenoménica no es el único fenómeno mental desconcertante y misterioso. La capacidad de nuestra mente para tener contenidos —para «versar sobre» o «ser acerca de» objetos, y para representarlos— resulta igualmente enigmática. Y si bien desde los años ochenta del siglo XX se dio por hecho que los contenidos mentales podrían naturalizarse sin mayor dificultad, los fracasos repetidos y sistemáticos de las teorías naturalistas del contenido nos invitan a buscar un enfoque alternativo. El presente artículo se propone presentar precisamente uno de esos enfoques alternativos, al que denomino ilusionismo del contenido. Esta posición se articula en torno a dos tesis principales. La primera es de carácter eliminativista y sostiene que los contenidos mentales, en rigor, no existen. La segunda es la tesis de la ilusión, según la cual existe una ilusión de contenido mental: los contenidos parecen existir y ser «transportados» por nuestros estados mentales. El artículo ofrece una presentación general del ilusionismo del contenido, explica sus tesis centrales, muestra que se trata de una posición filosóficamente atractiva sobre el contenido, y la defiende frente a algunas objeciones de corte intuitivo.

Gako-hitzak

Eduki mentala
Ilusionismoa
Gogo-gorputz auzia
Eliminatibismoa
Intentzionalitatea
Enaktibismo erradikala

LABURPENA: Dirudienez, gogoaren filosofo analitikoentzat, kontzientziaren auzi zaila da printzipioz konpondu gabe dagoen gogo-gorputz auziaren alderdi bakarra. Alabaina, kontzientzia fenomenikoa ez da fenomeno mental harrigarri eta misteriotsu bakarra. Gure gogamenak *edukiak izateko* —alegia, «objektuei buruz» edo «objektuez aritzeko», eta haiek irudikatzeko— duen gaitasuna ere harrigarri eta misteriotsua da. Eta 1980ko hamarkadatik aurrera neurri handi batean ontzat eman bazen ere eduki mentalak erraz naturaliza zitezkeela, edukiaren teoria naturalisten porrot errepikatu eta sistematikoek ikuspegi alternatibo bat bilatzera bultzatzen gaituzte. Artikulu honetan, halako ikuspegi alternatibo bat aurkeztu nahi dut, nik *edukiaren ilusionismo* izendatu dudana. Edukiaren ilusionismoa bi tesi nagusiren inguruan egituratzen da. Lehenengoak izaera baztertzailea du, eta, horren arabera, ez dago eduki mental hertsirik. Bigarrena ilusioaren tesia da, eta, horren arabera, eduki mentalaren ilusio bat dago: badirudi edukiak existitzen direla eta gure egoera mentalek "garraiatzen" dituztela. Artikuluan, edukiaren ilusionismoaren aurkezpen orokorra egiten da; haren tesi nagusiak azaltzen dira, edukiari buruzko ikuspuntu filosofiko sinesgarria dela erakusten da, eta objekzio intuitibo batzuen aurrean defendatzen da.

* ***Correspondence to:*** Marco Facchin. Department of Philosophy and Letters, Faculty of Philosophy, Centre for Philosophical Psychology, Universiteit Antwerpen, Building R, Rodestraat 14, Antwerpen 2000, Belgium. – marco.facchin.marco.facchin@gmail.com / marco.facchin@uantwerp.be – https://orcid.org/0000-0001-5753-9873

How to cite: Facchin, Marco (2026). «Illusionism as a theory of mental content»; *Theoria. An International Journal for Theory, History and Foundations of Science,* 41(3), 396-433. (https://doi.org/10.1387/theoria.27610).

Received: 11 June, 2025; Final version: 27 June, 2026.
ISSN 0495-4548 - eISSN 2171-679X / © 2026 UPV/EHU Press

This work is licensed under a
Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

1. A manifesto

Most philosophers take the mind-body problem to have been reduced to the hard problem of consciousness — the problem of explaining why any subjective, *experiential feeling* accompanies our neural activities. The rest of the mind has been figured out — *at least in principle* (Chalmers, 1997).

Yet, phenomenal consciousness is not the only surprising, unique and "weird" feature of the mind. The ability of minds to have mental contents — their capacity to be about, "directed at", or represent things — is quite puzzling too. Think that 2+2=4. How did your mind manage to grasp an abstract mathematical truth, presumably outside time and space? Look at these words. How can your perceptual state "contain" them, making you aware of them and their meaning? Recall any dream of yours: how could such a weird and complex scenario have been shown to your mind's eye?

From the 80's on, a group of philosophers systematically tackled these questions. Their objective was to naturalize mental content — that is, explaining *why* mental states carry the contents they carry by resorting only to naturalistically acceptable, ontologically unsuspicious ingredients (e.g. Millikan, 1984; Fodor, 1987; Dretske, 1988). 40 years later, the problem is still open. The inability of naturalistic approaches to deal with the problem motivates a novel exploration of our conceptual landscape, and the search for an alternative approach to content (see Kriegel, 2011; Egan, 2020).

Here, I sketch *content illusionism* — one such alternative approach. This approach is indebted to Dennett (1991b; 2005) Frankish's (2016; 2023; 2026) and Kammerer (2021) illusionism about phenomenal consciousness, and revolves around the idea that content is an illusion: something that *merely appears* to exist, without *really* existing. If content illusionism is correct, then, our minds are not "about" nor "represent" anything — our mental states come before the mind's eye empty handed, so to speak. Yet, they don't *seem* empty handed. We won't *naturally*, *naively* describe them as such. On the contrary, they appear to "be about" or "represent" many things, sometimes very complicated ones.

Note that, just like its phenomenal counterpart, content illusionism isn't a *single* view, but a *family* of possible views with varying degrees of kinship and similarity (Frankish, 2016; 2023). So, in sketching content illusionism, I am sketching a *motif* (Raja, 2024): a set of usefully vague and open-ended ideas that guide a specific set of epistemic practices. Here, I wish to sketch such a *motif*, introducing it to the literature and showing that it is an attractive position. So, I won't articulate a full-fledged philosophical theory, nor offer a direct argument to the effect that content is an illusion. For such an argument, see (Facchin, 2025).

I shall thus begin by clarifying the terms of the debate, fixing the reference of "mental content" (§2.1); and presenting the core claims of content illusionism (§2.2). I shall then motivate content illusionism, showing that it is a compelling alternative to traditional views of content (§3), and deflect some obvious objections it faces (§4). I will then quickly substantiate the motif, offering a how-possibly explanation of the illusion of content (§5), and then conclude this manifesto by indicating some directions for future research (§6).

2. Presenting content illusionism

2.1. On "content" and its features

Content illusionism claims that mental content — henceforth, just "content" — is an illusion. But what is this content that's supposed to be an illusion? What does "content" mean?[1]

[1] Owing to space limitations, and the programmatic nature of the paper, the discussion of these matters will be highly schematic. See Facchin, 2025) for a more in-depth discussion.

As now customary (e.g. Shea, 2018; Mendelovici, 2018) I fix the reference of "content" by *ostension*. Looking at paradigmatic and uncontested[2] examples of contentful and contentless mental states, it should be possible to identify some natural and salient property states of the first type all have and states of the second type all lack, which is thus what "content" designates (cf. Schwitzgebel, 2016; 2024, ch.8).

As a first example of paradigmatic contentful mental states, consider *perceptual states* such as seeing an apple, hearing an ambulance passing, or smelling the dinner burning in the oven. Notice how, in all these cases, there is some object[3] being perceived — an apple, the passing ambulance, or the dinner turning into charcoal. Other perceptual-like mental states, such as hallucinations, misperceptions, or bouts of imagings share the same feature too, and are indeed paradigmatically contentful.

Consider now *occurrent thoughts* — the kind of sentences we "say aloud" in our heads when we think. Right now, you are "saying aloud" this very sentence — if you are not distracted. So, there is something that you are "innerly saying aloud", something that you are thinking. And the same holds for other explicit acts of thought like judging, deciding, evaluating, considering and inferring.

Consider lastly propositional attitudes — *the* paradigmatic example of contentful mental states (e.g. Fodor, 1987). Believing that p, desiring that q, predicting that r and intending to ϕ all have some sort of object that is believed, desired, predicted or intended introduced by the "that" or "to".

Consider now some paradigmatic *contentless* states. A first paradigmatic case is offered by *non-mental* states, like having a cavity, a broken bone, or a cold. Whilst one might be conscious of these states in virtue of their symptoms, none of these states is considered to carry mental content, on account of being non-mental states.

As a second paradigmatic case, consider now some psychophysical states, such as being hungover, being horny, being thirsty/hungry, being stressed, or being burned out. These states have a rich and complex phenomenology, and are thus mental states in a prominent and clear sense of the term. Yet they lack any object — when you are perceiving, there is something perceived. But when you are hungover, there is nothing *hungovered*, so to speak.

Consider, lastly, moods such as elation, depression, nervousness, or discomfort. These states are clearly mental and are recognized as such by everyone. They have a rich phenomenology, and a complex impact on our behavior. And yet, there does not seem to be anything *elated* in elation, *depressed* in depression or "nervoused" in nervousness; at least, not in the way anything is perceived in perception, or believed in belief.[4]

What sort of visible and salient property is universally present among contentful states and universally absent among contentless states? An *essential link with an object*. All contentful states are partially identified by it. Change the type

---

[2] Almost uncontested - content eliminativists may disagree (e.g. Rosenberg, 2019). But they too would concede that, if content existent, these states would be contentful.

[3] I'm using the term "object" a bit loosely here - the relevant object could be a state of affair, a property, or something else. Still, there is *something* being perceived.

[4] As the reviewers noticed, such a view is not universally held by philosophers, and, popularity-wise, it might even be on the back foot. Crane (1998), for example, called upon the phenomenological tradition (especially Sartre, 1971) to argue that the intentional object of such moods is *the world*. Roughly, just as my (and your) current perceptual state has *this sentence* as its object, or my fear of spiders is about spiders, the depression of the depressed person is "about" the world, or has the world as its object. I must confess I don't really see how depression "is about" the world, or how it has the world as its object in a way that is identical (or even roughly analogous) to the way perceptual states and emotions have objects or "are about" objects. So, I must confess that I cannot but regard that claim as false. I also suspect that the claim is generated by an overextension of a *different* claim coming from the phenomenological tradition, namely that our emotional states (moods included) are partially constitutive of our "openness" to the world; that is, of our *intentional* contact with the world (Heidegger, 1962; see also De Haan, *et al.* 2013 for an updated, "neurophenomenological" rendition of that claim). I find this second claim much more plausible - and indeed I accept the view that (at least some) emotions appear to be contentful. My fear of spiders is indeed *about* spiders, and has spiders as its objects.

of link, and you turn a belief into a desire or a perception into a fantasy. Change the object, and you turn a belief that p into a belief that q, the perception of an orange in a perception of an apple. Nothing of the sort holds for contentless states. Since there is nothing *depressed*, there is nothing whose change turns a depression of a into a depression of *b* — because depressions are not "of" nor "about". And since there is no such object, there can be no linkage with it whose change turns depression into elation, or burnout into hungover.[5]

Note that, strictly speaking, the examples above are examples of contentful and contentless *mental states*, rather than *content*. So, the above fixes the reference of "contentful" and "contentless" mental states. From these, it is easy to fix the reference of "having content" — to have content is to have the linkage with an object contentful states paradigmatically display. But once the reference of "having content" is given, the reference of "content" is too — "content" refers to the relevant objects contentful mental states are essentially linked to.

Is this usage of "content" idiosyncratic? Quite the contrary. Consider:

> "When we introspectively notice intentional states [...] we also notice something we are tempted to describe as what our mental states are "directed at" or what they "say"; *this is their (intentional) content*." (Mendelovici, 2018, p. 8; italics added)
>
> "There is some need for clarification in the literature regarding how we should understand the term 'content' and what a theory of content is about. The term 'content' is commonly used to *refer to the intentional object of a representation* – the thing (or property, abstract entity, etc.) represented." (Ramsey, 2023, p. 2 fn 1; italics added)

Mendelovici and Ramsey have very different philosophical affiliations: the former comes from analytic, "armchair" philosophy of mind, and the latter from (post?)analytic, empirically informed, philosophy of cognitive science. Yet they use "content" identically. Brentano reminds us this usage of "content" is way older than analytic philosophy, and can be traced at least back to the Scholastics[6]:

> "Every mental phenomenon is characterized by what the Scholastics of the Middle Ages called the intentional (or mental) inexistence of an object, and what we might call, though not wholly unambiguously, *reference to a content*, direction toward an object (which is not to be understood here as meaning a thing), or immanent objectivity." (Brentano 2012, p. 88, italics added)

This usage of "content", then, is common, and hides no sleight of hand. Indeed, it is so common that it "leaked" onto the discussion of sub-personal representational posits, to designate what these posits are about (e.g. Neander, 2017; Williams, 2019; Piccinini, 2022).[7] I won't challenge this usage here: after all, even subpersonal representational posits seem to have that kind of essential tie to an object.

Still, this usage of "content" is ambiguous. At times, the world seems to pick up o*rdinary objects*, broadly understood: mind-independent "middle sized dry goods" (cf. Clark, 2009). For example, when I see an apple, think of the colosseum or remember my hotel room in Nuremberg, the content of these states seems to be some ordinary physical object. Yet, other mental states suggest otherwise. When, for example, one hallucinates the proverbial pink elephant, or misperceives a dog turd as a man[8] there is no "middle sized dry good" corresponding to what one misperceived. It is natural, in such cases, to think of content as "mental things" of sort — as *sui generis* objects "in" the mind and "shown" to the

---

5 It is however plausible such changes can be executed by changing the relevant phenomenal features of these states.

6 A reviewer (whom I thank) informed me that it is possible to find even earlier discussions of intentionality, such as in the Nyāya school of philosophy; see (Das, 2021).

7 Presumably, the relevant sub-personal representations also include neural representations, even if they are supposed to be observables rather than posits (Thomason & Piccinini, 2018). For further discussion, see (Facchin, 2023; Drayson, 2025).

8 This actually happened; see Clark (2016, p. 8).

mind's eye; and indeed our folk conception of content does at times conceive contents in this way (cf. Crane & French, 2021).

So, are contents *ordinary objects* or not — or, better, given the eliminativistic view the paper will soon introduce: is "content" *supposed to refer* to ordinary objects or not? I am inclined to give a negative answer — based on the (hopefully uncontested) claim that contents have features *very unlike* the features of ordinary objects. Thus notice, for example, that contents are "satisfiable" (Facchin 2025, p. 206): they are (or "bear"; the literature attests both usages[9]) conditions of satisfaction. And they likely do so essentially, as indeed in many cases contentful states must be either satisfied (e.g. accurate, truthful, or fulfilled) or not. Ordinary objects surely are not satisfiable in this way. Contents are also "apparently worldly" (*ibidem*), in that they present themselves as objects existing over and above the mental state "carrying them" — and yet, in many cases (especially when the content is not satisfied) they seem to be internal or immanent to the state itself. Again, ordinary objects don't seem to possess this feature.

I don't think that this conclusion is (or, should be) controversial. Most philosophers don't take contents to be ordinary objects. This is presumably why philosophers strive to *naturalize* them — ordinary objects stand in no need of naturalization. And notice that one doesn't *need* to buy into some Meinongian theory of non-actual but existent objects to hold the view that contents are not ordinary objects. Taking contents to be *propositions* (e.g. Frege, 1918; Fodor, 1987) or possible worlds (e.g. Wittgenstein, 2001; Lewis, 1973; Kiefer & Howhy, 2018) is enough to take contents *not* to be ordinary objects.

So, "content" designates the objects our contentful mental states (and sub-personal representational states) are essentially linked to, as weird and "non-ordinary" as these objects may be. What does content illusionism say about *them*?

2.2. The core claims of content illusionism

Content illusionism is a (family of) theory (or theories) of content. But "theory of content" is ambiguous between two readings (Ramsey, 2023):

> **Reading (1):** Theories of content are meta-semantic theories, detailing the factors in virtue of which a contentful state carries the content it carries.
>
> **Reading (2):** Theories of content are metaphysical theories, accounting for the ultimate nature of contents.

Reading (1) today prevails: mainstream naturalistic theories of content are theories of content in this sense (e.g. Millikan, 1984; Shea, 2018).[10] And so are their non-naturalistic alternatives (e.g. Bourget & Mendelovici, 2019). Reading (2) is, in contrast, very *ageé*: it calls to mind "Meinong's object theory" (Meinong, 1960) or James's view of contents as physical particulars we bodily relate to (James, 1911).

Shamefully, content illusionism contributes to the confusion concerning the meaning of "theory of content" because it is not a theory of content in any of these two senses. Indeed, it consists of two main claims, the first of which is that *content should be eliminated* — a claim that uneasily sits with both reading (1) and (2).

How should this eliminative claim be understood? Eliminativism comes in two basic variants — entity eliminativism and discourse eliminativism (Sprevak & Irvine, 2020). Entity eliminativism claims that something does not exist — it makes a negative existential claim. Discourse eliminativism makes no negative existential claim (or at least, it's not com-

---

[9] See Zahnoun (2023).

[10] This doesn't, however, mean that such theories are metaphysically agnostic about content. As noticed above, they often take (or end up taking) contents to be propositions or possible worlds.

pelled to make any). Rather, it makes a *linguistic* claim; namely, that we should opt out of certain discursive practices involving the to-be-eliminated word. The eliminative claim advanced by content illusionism should be read as advancing a form of *entity* eliminativism:

> **Eliminative claim:** There are no mental contents.

The eliminative claim entails that the reference of "content" is empty (since there are no contents). And because of this, the eliminative claim *entails* (but is *not* identical to) a weak form of discourse eliminativism, at least for scientific realists.

This is because contents (and the representational states "carrying" them[11]) allegedly play a number of relevant roles in the modern mind sciences. For instance, they are called upon to explain specific patterns of successes and failures of an agent (e.g. Godfrey-Smith, 2006; Shea, 2018) and in identifying common, abstract patterns of response or functioning that are exhibited despite the physical difference of the stimulus or input conditions (e.g. Dennett, 1987; 1991). Now, if one is a scientific realist, one should also discursively eliminate content from these explanations (Ramsey, 2020). The reason is simple: realists take these statements to be ontologically committing. But if the eliminative claim is correct, these commitments are wrongheaded — hence, *they shouldn't be made*. So, the linguistic practices making them should be avoided.

Notice however that the form of discourse eliminativism entailed by the eliminative claim concerns only the *scientific* discourse about the mind.[12] Other relevant forms of content-invoking talk do not bear the relevant ontological commitments, and thus are left untouched. Thus, it does not affect our use of mental contents for *making sense, predicting and rationalizing* the behavior of ourselves and others (e.g. Dennett, 1987; Toon, 2023; Facchin *et al.*, 2025), as these discursive usages of content are entirely possible even without making any ontological commitment to the reality of content. For the same reason, it does not affect usages of content in narrative practices — including self-narration, behavior shaping practices and narrative practices regulating social coordination (Dennett, 1992; Demeter, 2022; Hutto, 2022). Thus, the form of discourse eliminativism entailed by the eliminative claim is remarkably moderate, far removed from Churchlandesque fantasies of all-encompassing linguistic reforms.[13] Indeed, as I shall explain, certain everyday usages of content *cannot* be coherently eliminated by the illusionist, on the pain of contradiction. This is because they are *constitutive* of the illusion of content (see § 5).

The *what*? Well, the illusion of content is what tells apart content illusionism from content eliminativism and fictionalism about mental content. Content eliminativists only say that content does not exist[14] (e.g. Rosenberg, 2019), leaving unexplained why we *think* it exists and behave as if it does. And fictionalism is *primarily* a claim concerning how we should interpret the semantics and/or pragmatics of our content-involving discourse — it's a piece of philosophy of language that, whilst of course sympathetic to anti-realism, neither entails nor is entailed by it (Gan, 2024). *Strictly speaking*, fictionalism has no ontological strings attached.

In contrast, content illusionism *has* ontological strings attached, and not merely negative ones. For, the second main claim of content illusionism is a positive existential claim, the illusion claim:

> **Illusion claim:** There is an illusion of content: whilst there are no mental contents, there seem to be mental contents "given" before our mind's eye.[15]

---

[11] Throughout the paper, I shall assume that representational states *constitutively* involve content. This seems an innocent assumption, given the notion of content at play.

[12] And only on the assumption that some form of scientific realism is correct. Fictionalists won't be troubled by this (see Sprevak, 2013).

[13] Fantasies that, to be fair to the Churchlands, they now have wisely abandoned (e.g. Churchland, 2012).

[14] Or, more often, that *some* contents in *some* mental states or processes do not exist (e.g. Churchland, 1992; Hutto & Myin, 2013).

[15] Or equivalently: content is *seemingly given* to our minds.

What is this illusion, and how should it be characterised? And why is it an illusion, and not just a simple mistake, or an hallucination?

Let's start with the second question. Hallucinations and mistakes are *subjective* and *asystematic*. They are subjective, because they are factually not shared by subjects. My hallucinations, my perceptual mistakes are *uniquely* mine, and it would be very odd if we were *both* hallucinating a pink elephant, or *both* mistaking an apple for a watermelon. Also, hallucinations and mistakes are asystematic. There is no stimulus that reliably causes me to hallucinate or to misperceive. In contrast, illusions are *intersubjectively concordant* and *systematic* — and that's what distinguishes them from simple hallucinations and perceptual errors (Macpherson, 2013; Mandik, 2016). Seeing the two lines of the Müller-Lyer makes *everyone* see two lines of different length. Looking at the complex arrangements invented by Ames makes everyone see rotating windows as oscillating (the so-called "Ames window" illusion), or people changing size when walking into a room (the "Ames room", see Ittelson, 1952).[16] These illusions are readily produced by "consuming" an appropriate stimulus, and are shared by *anyone* consuming it.[17]

Mental contents display the same intersubjective concordance and systematicity — hence they are aptly called an illusion. At least in normal conditions, when gazing at a perceptual scene, we will all see the same thing – i.e., be in a perceptual state with the same content. Similarly, by hearing the premises $\forall(x)\ f(x) \rightarrow g(x)$ and $f(a)$, we will all conclude that $g(a)$; that is, we will all formulate a judgement with the same content. By hearing "Napoleon" we will all think of a French general turned emperor, and so forth. Sure, there are contextual and agential variation: if one does not know modern formalized logic, reading $\forall(x)\ f(x) \rightarrow g(x)$ and $f(a)$ won't let them conclude that $g(a)$, and if we are at the weekly meeting of our *Animal Farm* reading group, hearing "Napoleon" will make us think of a fictional character. But such changes are not random — they are lawfully based on the context, knowledge, and ability of the subject. Even mental contents of states that are *not* tokened as a direct response to any stimulus are largely identical across subjects: for example, we tend to dream anxiety-inducing scenes (Revonsuo, 2000); and subjects affected by schizophrenia tend to hallucinate voices (e.g. Sartorius *et al.*, 1986). This uniformity is quite remarkable: of *all* the things we could dream (or hallucinate) we dream (or hallucinate) only a very limited, and intersubjectively concordant, portion.

How, however, should this illusion be conceived? This is a pivotal question to answer. After all, our intuitive conception of an illusion is that of a state that *misrepresents* something (e.g. Macpherson, 2013) — that is, we conceive illusions of contentful states. Clearly, if this were the relevant sense of illusion, content illusionism would be incoherent.[18]

We can, however, think of illusions in other ways. One such way is *behavioral*.[19] On this view, to be illuded is to be "pulled" to adopt a certain behavioral pattern by a stimulus. When, for example, one is in the grip of the Müller-Lyer illusion, one will behave in certain characteristic ways: they will say that one line is longer, indicate the line with inbound arrowheads as longer, they will show surprise after measuring the two lines, *etc*. And, in a way, the subject *can't but* behave in this way: it is compelled by the stimulus to do so.[20] Similarly, the content illusionist suggests, we can't but behave as if we were constantly shown contents being "carried around" by our mental states. For example, we cannot but make *sincere, but false* reports of mental contents, and we cannot but *sincerely* — that is, without explicit dissimulation — treat others (and ourselves) as if their mental states hosted contents. As I shall sometimes say for the sake of simplicity, we *fit the behavioral stereotype* of a persona believing that mental content is real (cf. Schwitzgebel, 2002; 2013), and

---

16 Both illusions are better experienced in video format: see https://www.youtube.com/watch?v=3t5qQB-FuJg for the Ames room and https://www.youtube.com/watch?v=cVepIZLepVc for the Ames window.

17 Including peoples from very different, "non-western" cultures and non-human animals (see Amir & Firestone, 2025).

18 This is a version of the "cognitive collapse" argument against eliminativism (see Baker, 1987; Sprevak, 2013). This is why the cognitive collapse (and related objections) are not addressed in §4.

19 Another way to think about the illusion will be introduced in §4.

20 Of course, this shouldn't be read *too* rigidly - a person who *knows* about the illusion will typically be able to suppress these behavioral patterns and exhibit different, "not-illusion-oriented" ones. But the pull of "illusory" behavioral patterns remains real - after all, one line *really seems* longer then the other!

thus we believe in a way such that the ascription of such a belief is (given our relevant folk psychological ascriptive practices) *false, but warranted*.[21]

An example to clarify. John wakes up tired and covered in sweat. As he gets to the kitchen, his partner looks at him and asks if everything is ok. John confirms that he just had a terrible night. He says something like:

> "I dreamed that my high school crush was investigating a murder, and I knew they wanted to frame you. So, I was driving you away to hide you, but I realized I wasn't really driving, I was cycling, and when I looked at the road again, I saw I was speeding on a sky-jumping trampoline. So, I tried to brake but I couldn't find the brake on the bike, and so I took this enormous jump and then woke up at 2:45 AM. And I was so upset and anxious I couldn't fall asleep until about 6:00."

What are John and his partner doing? John is presumably explaining why he looks awful, and he does so by means of a strange and incoherent narration. What matters is not if the narration is true or a true report of an experience (the illusionist of course claims that it isn't, but this isn't the focus now). What matters is how John and his partner *treat* it. John is presumably aware that the piece of narration is literally false (nothing of that really happened), and of its poor aesthetic qualities. Yet he is *compelled* to produce that narration — at least, if he's a collaborative speaker.[22] He cannot, for example, recount a different event of his life (it would not *explain* why he looks terrible) or recount a better, more scary story to justify his lack of sleep (e.g. a passage from Stephen King's novel *IT*). Indeed, if his partner noticed that John is doing something like that, they would assume that John is either abandoning the conversation; or trying to make a fool of them. Similarly, John's partner cannot, if they're being a collaborative speaker, remark on the lack of aesthetic qualities or coherence of John's narration. They're both naturally inclined (and compelled by social norms) to treat John as if during the night he was the spectator of the "mental film" he recounted — a mental film so realistic, engaging and anxiety inducing to have disrupted John's sleep pattern. They both fit the *behavioral stereotype* of people believing that John had such and such a dream, and should thus be ascribed the relevant belief (given our folk psychological practices). Most importantly, John and the partner *can't but* fit that stereotype. They can't but *behave as if* John witnessed some sort of contentful "inner movie" while sleeping, justifying the (for the illusionist false) ascription of the corresponding belief.[23]

Aren't John and his partner just pretending, as Toon (2023) recently argued? No. Even if John and his partner behave only as *if* John witnessed an "inner movie" while sleeping, their behavior is not pretense — not even if pretense were *just* a specific, enculturated behavioral pattern, as Toon (2023, p. 72-75) seems to hold. For, even if pretending to ϕ were just a behavior, the behavioral profile of pretending to ϕ would differ, no matter subtly, from *actually* ϕing — indeed, this is what essentially tells apart pretense behavior from non-pretense behavior (see Facchin & Rucinska, 2024). We can thus always contrast pretending to ϕ with really ϕing. But, in the case at hand, we have *no* such contrast. There is no "more serious" version of what John and his partner are doing. There is only a "less serious" version — *pretending* to have had a nightmare. In that case John's behavior will be different — if John only *pretends* to be tired, he won't, for ex-

---

[21] Won't the identification of behavioral stereotypes *presuppose* content? After all, behavioral stereotypes are typically identified in terms of that-clauses (cf. above: "the stereotype of the person *believing that content is real*"). Not really: we can think of behavioral stereotypes as "real patterns" (Dennett, 1991) to which our ascriptive practices "latch onto" - and crucially the existence of ascriptive practices does not *entail* the existence of the contents ascribed (just like the practice of exorcizing demons does not entail the existence of demons, no matter how successful that practice seems). Thanks to Erik Myin for this suggestion. See also Hutto & Satne (2015) for a similar idea defended in more detail.

[22] Why *that* narration rather than any other narration? I don't know. But that's not a problem unique to content illusionism. I know of no theory able to predict and explain why we dream *exactly* what we dream, rather than any other thing we could dream.

[23] What about any phenomenal state John may have undergone in his sleep? Is the content illusionist *forced* to deny their existence? Strictly speaking, the answer is negative. The content illusionist has only to deny that such states carry the content John reported (and any other mental content). This of course goes against *some* ways of conceiving of phenomenal states as essentially contentful (e.g. Mendelovici, 2018), but not against the arguably standard characterization of phenomenal states as non-contentful (Dennett, 1988; Chalmers, 1997).

ample, yawn naturally (though he may *pretend* to yawn, exaggerating his movements) and his movements won't be sluggish in the way the movements of a tired person are. He may even recount a piece of Stephen King's *IT* without triggering any bad reaction in his partner or stepping away from the relevant discourse commitments. So, John's ordinary case can be contrasted with a pretense case, but there is no *non-pretense* case with which to contrast *it*. Thus, it's no pretense.[24]

But how else can we make sense of "behaving as if", if not in terms of pretense? We can do it in terms of (systematic, intersubjective) *mistakes*. Consider this analogy. On a remote island, there is a remote tribe. The tribe's people live mostly on vegetables, because most of the meat they possess serves to stave off the Great Spirit of the Forest — which, if not appeased, would pillage the tribe's garden and cause a famine, or even eat the village people. To appease the Great Spirit, a delegate of the tribe offers it a daily sacrifice. At dusk, she goes near the Sacred Mound near the Forest where the Great Spirit resides. She leaves the corpse of a butchered animal there, and picks up the bones of the one offered the day before. There is, of course, no Great Spirit of the Forest, and the Sacred Mound is just a gigantic ant nest. The ants belong to an undiscovered, nocturnal and mostly carnivorous species. By offering this sacrifice, the tribe is basically ensuring the ants will be occupied throughout the night, and won't eat the vegetables from the village garden (or the village people). Still, villagers know nothing about these ants: it's taboo to get to the sacred mound at any other time, the taboo is strictly observed, and the ants are not yet active at dusk.

What are the tribe's people doing? Notice, the question is not what they *think* they are doing — the question is what they are *really* doing. We can exclude that they are *really* appeasing the Great Spirit, since there is no Great Spirit to appease. But notice that they are not just pretending to appease the Great Spirit. They are *not* pantomiming the prescribed ritual, but actually carrying it out. They are thus behaving as if there was a Great Spirit — thereby fitting the relevant behavioral stereotype - even if there is none, *simply because they are wrong*. Still, their behavior is pragmatically useful (it *really* keeps the carnivorous ants at bay) and able to sustain the belief that *there really* is a Great Spirit: who else would be devouring the sacrificed animals? In the same way, the illusionist suggests, by treating each other as having mental contents we are wrong, but in a useful and epistemically self-sustaining way.

Now, by endorsing the eliminative claim and the illusion claim, illusionism re-arranges our explanatory targets. Rather than explaining content (which does not exist), we should try to explain the *illusion* of content — which is real, and whose origins and grip are at present mysterious. The explanation of it could take a number of different forms, depending on the mechanism called upon to explain the illusion and its grip.[25] It's possible to envisage different forms of content illusionism based on the kind of mechanism invoked to explain the illusion of content. One option could be that of invoking a purely cerebral mechanism — the selective pressure to socially coordinate, for example, may have made us evolve a convenient system to communicate our relevant internal states so as to coordinate with each other. On the opposite side of the spectrum, the mechanism may be entirely "external" and "social", and dependent, for example, on our narrative practices. And between these two extremes, there is a vast sea of possibilities I cannot explore here. In §5, I will sketch a minimal mechanism generating the illusion which calls upon both neural and social (really sociocultural) resources. I do believe that such a "hybrid" view is the correct one. However, in line with the programmatic objectives of the paper, I won't defend it here. Indeed, in line with these objectives I wish to highlight that the "hybrid" view I will later sketch is not the only possible one, and that the space of possible mechanisms generating the illusion is quite wide.

Notice that the above means that, *strictly speaking*, content illusionism is not a theory of content (in any of the two senses seen above), but a theory of the *illusion of content.* Still, theories of content (in any of the two senses) and content illusionism have an important *explanandum* in common: our content-invoking introspective reports. For the illusionist, such reports are a constitutive part of the illusion, and should be explained by the relevant mechanism responsible

[24] I make this argument at length in a paper currently under review.

[25] If this sounds preposterous, keep in mind that philosophers as diverse as Clark (2015), Lee (2019) and Egan (2020) make an analogous claim about content *itself*. In comparison, the claim that a mechanism can explain the *illusion* of content seems fairly mild.

for it — indeed, such a mechanism would provide an error *theory* accounting for the production of these (false) reports. For the realist, such reports are genuine, truthful reports (at least in most cases), and their production should be explained by an appropriate introspective mechanism. Moreover, illusionism and theories of content (in any of the strict senses seen above) are clearly mutually exclusive, in the sense that the truth of content illusionism implies the falsity of any theory of content and *vice versa*. As such, it makes practical sense to continue to talk of content illusionism as an *alternative* to realism.

This concludes my general presentation of content illusionism. This is, admittedly, a very impressionistic, general description of content illusionism. It doesn't substantiate — or even suggest — any particular illusionistic doctrine. This is by design: recall that here I wish only to present the general *motif* of content illusionism, rather than any specific illusionist view. But why should one bother with illusionism?

3. Motivating content illusionism

Here, I wish to give some very general reasons as to why content illusionism is an attractive position, which we should prefer to the alternatives. But to do so, I must first clarify what the relevant alternatives are. To do so, I follow (Frankish, 2016), and take the relevant alternative to be only *realism*, either conservative or revolutionary.[26]

Conservative realism, like illusionism, adopts an austere ontology, wherein the only admissible entities and properties are the ones accepted by our *normal* (in the sense of Kuhn 1970) natural sciences. Unlike illusionism, however, the approach is committed to the existence of content — it is a form of realism, after all. Theories of content that identify contents with regular, mundane objects — such as James's (1911) account wherein contents are concrete particulars — are one straightforward example of conservative realism. Yet, conservative realism is nowadays more represented by a specific, "reductive" reading of naturalistic metasemantic theories, which tries to *reduce* or even *identify* content with the relevant naturalistic features mentioned in the relevant meta-semantic theories (see e.g., Lee, 2019; 2021 for a modern articulation of this view).[27]

Revolutionary realism retains the realistic commitments of conservative realism, but it doesn't try to squeeze content in the mold of normal natural sciences. On the contrary, it argues that content cannot be so squeezed. To account for content, one must either operate *outside* of the framework offered by natural sciences; or revolutionize that framework introducing some novel ontological ingredients. Phenomenologists who doubt that the mind can be naturalized are a valid representative of the first tendency (Husserl, 1970; Gallagher, 2017), whereas semantic primitivists taking semantic features as a basic ingredient of reality (Burge 2010) are a good example of the second tendency.

Now, why should one prefer content illusionism to these forms of realism?

---

[26] What about all non-realistic views of content? My answer is that, when properly analyzed, these are either not alternatives to illusionism, or *extremely* unattractive ones. As seen above, fictionalism in its various incarnations is a theory about the semantics and/or pragmatics of our content-involving talk. It's not a theory of how that talk is produced and why we believe in it (i.e., behave as if mental contents are real). So, illusionism and fictionalism are not alternatives (though they can still clash on the semantics and/or pragmatics of the content-involving talk). Similar considerations also hold for *instrumentalism*. Eliminativism about content is a truncated and unconvincing form of illusionism: it lacks the illusion claim, and so it misses an obvious *explanandum*, namely the illusion of content. Since I cannot see any other non-realistic alternative, I conclude that realism, in its various forms, is the only palatable alternative to illusionism.

[27] Notice that in this case the conservative realist would be *weakly* illusionist, in that they would claim that content exists, but are really different from what they appear (they just are certain relevant naturalistic ingredients called upon by theories of content). For the relationship between reductionism and "weak" forms of illusionism, see Frankish (2016).

3.1. Motivation #1: Parsimony

A first reason to prefer content illusionism is that (ceteris paribus[28]) it's more *ontologically parsimonious* than the alternatives — it denies the existence of contents, whereas the alternatives are committed to it. In the case of conservative realism, the parsimony is only *quantitative*: both views agree on the *kinds* of things admissible in our ontology, but disagree on their *number*: according to the illusionist, their number does not include contents. According to the conservative realist, it does. In the case of revolutionary realism, the parsimony is also *qualitative*: illusionist and revolutionary realist disagree also on the *kinds* of things our ontology admits. For the revolutionary realist is committed not only to the kinds that figure in our normal sciences, but also to some other kind pertaining to our revolutionary sciences. That's why it's *revolutionary* realism, after all.

What about *mechanistic* (and, relatedly, *explanatorily*) simplicity? One may worry that the mechanism generating the illusion of content *must* be pretty complicated, generating pretty complex explanations. This would compare them unfavorably to the mechanisms — and the explanations — at the realist's avail. After all, the realist could call upon a simple detection mechanism, which would enable a straightforward explanation of why we *behave as if* our minds were populated by contents (namely, we do so *because we detect them*). I'll deal with this issue in §5, showing that the mechanisms responsible for the illusion of content can be remarkably simple, and that a *satisfactory* realistic explanation of our reports and behavior must be *at least as complex* as the illusionist's one.

Now, if *parsimony* is a reason to prefer a theory over the alternatives, we have a reason to prefer illusionism over realism. Yet, parsimony offers such a reason only *everything else being equal*, and the realist would claim that everything else is *not* equal. I agree — I think it's even better for illusionism; which brings me to the second reason to prefer it.

3.2. Motivation #2: Better fit with the mind sciences

Representations and contents were the central explanatory tools of "classical" cognitive science.[29] But, pretty unsurprisingly, the landscape of the mind sciences has changed from the '50s — and at present, it's far from clear that representations and content play any role within the sciences of the mind (cf. Ramsey, 2007; Favela & Machery, 2024). Sure, the term "representation" is used, thereby invoking contents. Yet, empirical surveys suggest that the term "representation" is currently used to refer to any relevant internal state (and "content" anything that caused the tokening of the state), regardless of whether the state satisfies any definition of representation. Indeed, often the usage of "representation" outstrips the definitions of the term that practitioners of the mind science accept (Villaroya, 2017; Favela & Machery, 2023).

Examples of this over-liberal usage of representations are not hard to find. For example, Kaplan & Zimmer (2020) speak of brain-wide representations of current behavior — in spite of the fact that representations are typically taken to be causally decouplable from what they represent (Haugeland, 1991; Baker *et al.*, 2022), unlike brains from current behaviors.[30] Brette (2019) talks of cell assemblies *representing* models, where he clearly means *implementing* models which (in turn) represent external objects (see also Jakobson, 2013 for a sample of other such cases).[31] In the "predictive processing" literature it is commonplace to treat brains, and even entire organisms, as models of their niches (e.g., Friston,

---

[28] That is, keeping constant other, independent ontological commitments (e.g. to mathematical entities).

[29] Notice, however, that this is something even very "classical" cognitivists deny (Stich, 1983; Chomsky, 1995).

[30] Aren't brain-in-a-vat scenarios proof that brains *can* in fact be causally decoupled from current behaviors? No, they aren't. At best, they prove that the brain can be *very wrong* about how reality is. As many philosophers have already argued at length (Hurley, 1998, Ch.8; Thompson & Cosmelli, 2011; Hohwy, 2017), even an envatted brain must be immersed in the causal regularities provided by an active (i.e., behaving) body. Indeed, the vat is precisely providing *that coupling in* the absence of an actual behaving body.

[31] The difference matters: the expression "MOSAIC" *represents* a pretty famous model of motor control (see Haruno *et al.*, 2001). But "MOSAIC" does not compute a thing. For that, you need the model *implemented*, not *represented*.

2011) — where "modelling" boils down just to a non-trivial form of matching (e.g., Conant & Ashby, 1970; Clark, 2017). In their *Principles of Neural Sciences*, Kandell *et al.* (2012, p. 370) talk of the job of neural sciences being that of discovering how the brain represents mental states, but they clearly mean something like *constitute or realize* — and examples could be easily multiplied.

Now, perhaps this unruly usage of "representation" plays some relevant role in current cognitive science (e.g., it may play some communicative function, see Favela & Machery, 2024). But, as philosophers of the mind sciences have repeatedly argued, it does currently play no *referential* role — at least, if the term "representation" is supposed to refer to *contentful* representations. For, in many — indeed, *most* — sectors of current cognitive scientific theorizing, the word "representation" (and its cognates) fails to refer to any *genuinely representational* posit or observable.[32] This point has been convincingly made for components of cognitive mechanism in general (Kohar, 2023), various neurological structures (Anderson & Champion, 2022; Facchin, 2023; Krakauer & Ramsey, 2026), different kinds of — currently extremely popular — "predictive" models of the mind (Orlandi, 2014; 2016; 2018; Facchin, 2021a, b) and, most importantly, artificial neural networks — from relatively shallow classifiers to deep auto-encoders (Ramsey, 2007; Davies Barton *et al.*, 2024). This covers the majority of cognitive models and theories.

This doesn't (of course) mean that the word "representation" fails to refer *altogether* — it is used to refer to many different *non-representational* posits or observables. It means, however, that the word is *misused* so that its usage is no longer diagnostic of any genuine representational commitment. Quite literally *any* other word could do the same job — even ridiculous ones such as "demon". Yet, in such a case, "demon" would fail to refer to demons, and so presumably no one would rely on such a usage of "demon" to argue that the cognitive sciences are ontologically committed to demons, or in favor of a *demonic theory of mind*. Indeed, no one invoked the *pandemonium model* (Selfridge, 1958) to defend these views, even if it *explicitly* used the term "demon" and quantified over it!

What about *information-carrying states*? Surely at times the word "representation" refers to those, and such states have some sort of affinity with representations. Aren't they enough to justify the claim that cognitive science is still committed to representations and content? No, they aren't. Even leaving aside the — quite substantial — problems concerning whether cognitive science is *really* committed physical states "carrying" information (see Hutto & Myin, 2013; Ramsey, 2020; Zahnoun, 2025; Sganzerla *et al.*, 2025) and whether such a commitment is actually beneficial or warranted (Nizami, 2019; Brette, 2019, 2026), cognitive science remains committed only to a Shannon-inspired[33], *correlational* notion of information; and such a notion of information has no essential tie to representations and contents. At the most basic level, such a notion of information captures only *correlations* between states, such that observing that one state is the case makes the occurrence of the correlated state more likely. Formalizing a bit, a state S carries information about a state Z just in case $p(Z|S) > p(Z)$ (see Shea 2018, p. 76, see also Dretske, 1998; Hutto & Myin, 2013; Neander, 2017; Krakauer & Ramsey, 2026). Of course, such a purely correlational notion of information is not semantic, representational, and contentful in any way.[34] The fact that the street is wet raises the probability that it has just rained, but streets are in no sense contentful representations. The fact that the tide is high raises the probability that the moon is in a certain position in respect to the body of water, but bodies of water do not represent the moon. Thus, the fact that modern cognitive science is committed to information-bearing states neither entails nor supports the view that it is committed to representations.

---

[32] Notice that such claim holds for *most* sectors of current cognitive science. As a reviewer correctly noticed, it may be that in some sectors of cognitive science (e.g. formal semantics) the notion of content cannot be eliminated. It is worth noticing, however, that even in this case there are theories that eliminate meanings and contents in the relevant sense. Distributional semantics (Harris 1954) could be an example.

[33] I say "Shannon-inspired" because Shannon's original formalism has aspects that are often "glossed over" by neuroscientific practice and related philosophical discussion. For example, little attention is given to channel conditions (for more on this point, see Martínez, 2025).

[34] This is also why it makes a good fit for naturalistic theories of content - that is, theories that try to ground content in natural (non semantic, non mental) factors.

All in all, it appears that the current mind sciences are largely doing fine without representations. Hence, we should not be committed to them — which favors illusionism over realism. But the fit between illusionism and the current mind sciences goes even deeper.

3.3. Motivation #3: Better explanations of behavior

The fact that the mind sciences have "silently" turned non-representational suggests that eliminating content won't hinder their explanatory power. Indeed, as I shall now illustrate, the opposite is true: *explicitly* eliminating content and representation puts us *more* in contact with how the mind works, and afford us *better* explanations of behavior.

For example, we intuitively tend to think of visual perception as a sort of "mental photo" of our surroundings. Iconic, depiction-like conceptions of perception are still alive in some sectors of (classical) cognitive psychology.[35] But we now have enough evidence to be sure that perception *cannot* consist in the creation of such a mental photo — and thus that research projects aiming to explain perception in these terms are misguided. For, our perceptual system is nothing like a camera for mental photos — it's rather a system made to retrieve the relevant environmental information just in time to guide our online action (see Churchland *et al.*, 1994; O'Regan, 1992; 2011; Noe, 2004).[36]

Wouldn't recent work on mental imagery (e.g., Nanay, 2023) rescue the "mental photo" view of perception, as a reviewer suggests? I don't think so. Indeed, consider *the predictive processing* account of perception, which I take to be a prominent neuroscientific account of the continuity between mental imagery and perception (cf. Clark, 2013; 2014; 2016; see also Van Leeuwen, 2025; Nanay, 2025 for further discussion). Without getting into too much detail on the account (I did so in, Facchin, 2022) the idea is that the machinery of perception does not passively wait for perceptual inputs, but rather tries to *guess* what the perceptual input is. Crucially, such guesses take the form of the *endogenous generation of the incoming input*. When it comes to perception, this means that the guessed and the actual input can be easily compared, generating an error signal that can be used to "revise" the guesses and keep it in contact with reality. The endogenously generated input, however, can also be used "offline", and kept free from the control of the error signal, generating various kinds of imaging phenomena (Clark, 2012; 2016). Perception, then, is just a form of reality-controlled imagery, and consists in the production of a reality-tested "mental image" of the actual visual scene — or so, at least, predictive processing suggests.

Notice that predictive processing is taken to be a strongly representationalist theory, indeed it may be "as representational as cognitive scientific theories get" (Gładziejewski, 2016, p. 562), to use the words of a prime defender of the theory's representational reading. Appealing to it, then, should put significant pressure on my own view.

Yet, the opposite is true. Not only does predictive processing *support* — and indeed it *mandates* — an active account of vision (see Friston *et al.*, 2012), wherein exploratory movements (such as, for example, saccadic eye movements) are constantly performed to continuously provide the relevant error signal and keep perception anchored to reality, but the theory also *presupposes* mastery of sensorimotor contingency (Seth, 2014; Facchin, 2021a; 2022). The reason is trivial: to economically generate the right inputs, the brain must be able at least to anticipate how our movements will impact them. And such a mastery need not have any representational basis, it can simply be realized by direct neuronal connec-

---

[35] See for example Block (2023).

[36] Isn't "retrieving the relevant information" something like operating contentful structures? No, at least not given the sense talk of information has for philosophers and cognitive scientists. Information is (as recalled above) understood as just a correlation between types of states - but surely correlations are not in and by itself contentful. Indeed, this is precisely what makes informational theories of content (e.g. Dretske, 1988; Usher, 2001; Heemskerk, 2024) *naturalistic* theories. The correlational nature of information means that talk of "retrieving information" is metaphorical (Hutto & Myin, 2014; Ramsey, 2020). To speak literally, we should - as Orlandi (2014; 2016; 2018) correctly does - talk about the visual system re-organizing itself in response to environmental perturbations. But that would make the paper a bit too cumbersome to read.

tions allowing motor neurons to influence the neurons endogenously generating the incoming input.[37] Additionally, as it has been argued time and time again (Orlandi, 2014; 2016; 2018; Facchin, 2021a,b, 2022; Downey, 2016; 2018; Ramstead *et al.*, 2020, Kirchhoff & Robertson, 2019; Robertson & Kirchhoff, 2019; Van Es & Myin, 2020) the machinery of predictive processing is, upon closer analysis, markedly representation-free. Indeed, predictive processing has been one important wake up call for philosophers, who noticed — and criticized[38] — its abuse of representational language (e.g., Friston, 2011).

So, looking at predictive processing — to repeat, what is commonly taken to be a *strong* representationalist theory — and the way it depicts the role of mental imagery in perception does not support a "mental photo" view of perception. On the contrary, it ends up supporting an active view of perception based on our sensorimotor mastery. Indeed, in the predictive processing view, it is precisely because perception is *active* (and thus actively generating the relevant feedback signal) that perception remains controlled and "in touch" with reality.

The same holds for other mental states that seem to heavily revolve around the manipulation of content. For example, insight, imagination and problem solving are not the result of the mingling of mental contents, but rather a product of an agent's intelligent and controlled engagements with a number of epistemically supportive environmental props (Gedenryd, 1998; Van Leuween *et al.*, 1999; Van Leuween, 2011; Vallée-Tourangeau *et al.*, 2016; Carvalho & Nolfi, 2016; Pezzulo & Nolfi, 2019; Bocanegra *et al.*, 2019).[39] Getting away from content, then, bolsters our theories in the mind sciences, affording us a better grip on our very mental life.[40]

But what about *folk psychological* explanations? Surely these are often content-based explanations. Won't eliminating mental content deprive us of them? No, it will not. As said above, content illusionism entails only a very limited form of discourse eliminativism, limited to our *scientific* explanations.[41] Of course, if one were to think (a la Fodor, 1987) that folk psychology offers a literal, representational causal-explanatory theory of how our mind works one would have to eliminate contents from folk psychology too. But I reckon that few, even amongst realists, hold such a view of folk psychology today. After all, we do know that such a view makes hefty, almost certainly false, assumptions about our cognitive architectures (Churchland, 1992; Ramsey *et al.*, 2013; Schwitzgebel, 2026), and that we can use folk psychology to causally explain our behavior even without making such hefty assumptions (e.g., Schwitzgebel, 2013; Langland-Hassan, 2022). So, I take it that folk psychology remains basically untouched by the elimination of content — which happens painlessly, without any significant loss of explanatory power.

3.4. Motivation #4: Getting a new, easier explanatory target

Talking of explanatory power, it's worth noting that endorsing content illusionism allows one to trade away a metaphysically hard explanatory target (mental content) for a metaphysically easy one (the illusion of content).

Of course, the mechanism explaining the illusion can be very complicated, and will likely mention various features of our nervous systems as well as the way they interact among them and with our environment (including, quite possibly, our *social* environment). Still, there is nothing in principle *metaphysically* complex in explaining our behaviors recurring

---

37 See Bovet (2007) for a simple model, Leinweber *et al.* (2017) for analogous neurological findings, and Facchin (2021a) for a philosophically oriented discussion of them.

38 For such criticism, see for example Van Es (2021), Baltieri *et al.* (2020). Notice also that such a criticism was shared by philosophers who defend a representationalist position - such as Clark (2017).

39 It's worth noticing that such a reliance on external props to solve complex cognitive tasks seems a quite basic feature of minds - it can be seen even in fairly "lowly" animal minds, like the minds of wasps and arachnids, which presumably don't trade in any content whatsoever (see Facchin & Leonetti 2024, for examples).

40 Anderson (2014) makes the same point concerning the explanation of our brain and its organization.

41 Assuming that some form of scientific realism is correct. See Ramsey (2020) for an argument to this effect tailored to representations.

to neural and environmental features — cyberneticists do so routinely (Braitenberg, 1986; Brooks, 1999; Nolfi, 2002; Carvalho & Nolfi, 2016). The same cannot be said for conservative and revolutionary realism. Revolutionary realism acknowledges that accounting for content from a naturalistic perspective is *metaphysically hard*, calling for a revolutionary revision of our naturalistic worldview. Conservative realism disagrees, and tries to account for contents in terms of familiar ingredients taken from our normal natural sciences. Such attempts, however, have thus far only produced failures, and the whole program seems to be stalling, if not degenerating (for this quite grim verdict, see Kriegel, 2011; Egan, 2020; Lycan, 2022). This strongly suggests that the problem of content is metaphysically hard — even if perhaps not as hard as the hard problem of consciousness (cf. Hutto & Myin, 2013; Mendelovici & Bourget, 2023). Thus, abandoning realisms in favor of content illusionism is an *advantageous* trade, one that we should be willing to make. Hence, we should prefer content illusionism over both forms of realism.

3.5. Motivation #5: Avoiding the problems plaguing realism

Relatedly, content illusionism avoids certain prominent pitfalls and problems that plague realistic approaches. For example, conservative realism is plagued by the *determinacy* problem: no matter how we try to refine our naturalistic metasemantic theories, such theories fail to ascribe determinate contents to the relevant mental states.[42] Another problem is that these theories overgeneralize: the relevant naturalistic "ingredients" that, according to these theories, ground contents are almost everywhere — and so too should be content, if these theories were correct. But this result seems wrong, at least insofar it ascribes mental contents to presumably mindless systems like smoke detectors and rifles (the examples are from Ramsey, 2007; Orlandi, 2014).[43] As for revolutionary realism, it typically insists on the role of *introspection* in the study and characterization of mental contents. But this leaves them open to a potent methodological objection — introspection is an extremely unreliable source of information about our mental life (e.g., Schwitzgebel, 2008). Indeed, even "rigorously trained" introspectors end up *disagreeing* on the specific structures of their mental contents (as attested by the failure of the introspectionist program in early scientific psychology). Moreover, in many (arguably most) cases our introspective reports are not the result of *introspection*, but rather confabulation (see Chater 2018, for a number of examples). And notice that we are unable to discern between introspection and confabulation — indeed, it's often shocking to discover that many of our introspective reports are groundless, and invented "on the spot". So revolutionary realism seems methodologically flawed, and thus an unreliable guide to our mental reality.

Content illusionism, of course, avoids these problems. By denying the existence of content, it avoids the problems of indeterminacy and overgeneralization that plague conservative realism. And, since it claims that there are no contents, it has an easy time to account for why introspective reports appear to be systematically confabulated on the spot — they are because there is nothing "inside" to report![44]

3.6. Motivation #6: Content illusionism accounts for the features of content better than any form of realism

A last reason to prefer content illusionism is that it accounts for certain salient features of content better than the realistic alternatives.[45] What are these features? I have already mentioned two — systematicity and intersubjective

---

[42] This problems changes in its articulation depending on the theory of content under scrutiny, see Artiga & Sebastián (2018) for informational accounts, Cummins (1989) for causal accounts, Fodor (1990) for teleological theories, Segundo-Ortin & Hutto (2021) for accounts based on abstract similarities, and Egan (2020) for a review.

[43] Importantly, even some revolutionary realistic theories of content suffer from the same problem (Nes, 2008; Raimondi, 2021).

[44] Well, what about the (seemingly) non-confabulatory cases of introspection? I have an account of these too. The important point to notice is that they are not reports of independently existing mental contents, but ways in which an agent makes sense of themself and his situation (Facchin et al., 2025).

[45] I make this point in full detail in Facchin (2025).

concordance (§2). Another important feature is that illusions are persistent in the face of opposing evidence. No matter how many times we measure the two lines of the Müller-Lyer illusion, they will always look different (Fodor, 1984).

Now, I have argued that content is systematic and intersubjectively concordant above, and I won't rehearse the point here. And, to be persuaded that content is persistent, is sufficient to notice that content eliminativists like Rosenberg or Hutto & Myin never — despite their sincere conviction of the inexistence of content — reported mental content vanishing. To the contrary, current content eliminativists clearly voice the persistence of content, by wondering how it is possible, if the mind really is a non-representational system, that something nevertheless seems to stand in for environmentally absent objects and features (Downey, 2024). So, content quite trivially has these features. But these features are only weak indicators of illusion. In fact, whilst they discriminate illusions from errors and hallucinations (which are unsystematic, intersubjectively discordant and, in the case of errors, non-persistent), they do not discriminate between illusions and reality. After all, the appearance of real objects, properties and events is subjectively concordant, systematic and persistent too.

Illusions, however, have other features too, which real objects typically lack. One is that they are *bound to a single mode of access*, and trying to access them in other ways, or to sensorily explore or manipulate them, typically reveals their illusory nature. Spectators of illusions, has been noted, are extremely passive (Dennett, 1991b; Hohwy, 2013), as exploring illusions "spoils the trick". Many illusions discovered by Ames, for example, force the spectator to look with just one eye through a peeping hole, minimizing the spectator's ability to explore the scene (Ittelson, 1952), and the appearance of diversity in the Müller-Lyer illusion can be dispelled by connecting the edges of segments via parallel lines. Illusions also show something that fails to fit in *the causal order of the world*. By looking at the Müller-Lyer lines, one could have the impression that the arrowheads caused the lines to take different lengths — something we know can't be the case. Or, by looking in the Ames room, one has the impression that moving causes people to grow or shrink in size, something that we know is not (and cannot) be the case.

Notice, importantly, that these features are *highly diagnostic* of illusions. Natural objects are typically accessible from multiple channels and freely explorable. Moreover, they are tightly knit in the causal order of the universe. Compare: if I were to show you my Ferrari, but force you to see it only from a *single* point of view and insisting that you cannot see it from any other angle, or perceptually explore it, you would likely think I really have no Ferrari, but just a Ferrari-shaped cardboard façade with which I'm trying to fool you. And this impression would likely be reinforced if my Ferrari fails to "fit" in the causal order of the world as a Ferrari would (e.g., if it's impalpable, or cannot be damaged).

Content, however, displays these two features. For one thing, the only *direct access* we have to our mental contents is by means of introspection. Mental states, after all, are *private*. Of course, we can *indirectly* access them and their contents (for example, through inference from behavioral data[46]) — but the same holds for illusions: one can indirectly access the Ames room illusion by reading about it. But to access it directly – that is, for the illusion to be *given* – one can only look through the peeping-hole. And the same goes for content — for it to be given, it must be introspected. Indeed, attempts to access contents from a different perspective — for example, by using sophisticated techniques of feature visualization, which should make transparent to us the contents represented by neurocomputational models — do not uncover things that seem even remotely mental contents. Consider, for example, the features visualized in **figure 1**:

[46] More recently, "mind-reading" techniques have allowed to ascribe contents by inference from *neural* data - but such technologies remain inferential through and through, requiring sophisticated probabilistic and information theoretic assumption in order to work (e.g. Kriegeskorte & Douglas, 2019, pp. 169-170).

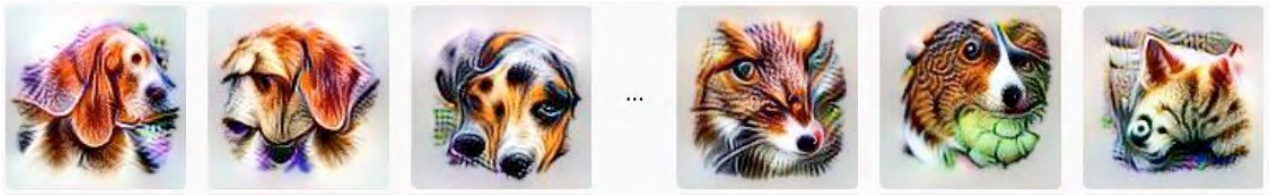


Figure 1

Several "mental contents" of a neurocomputational model, respectively labeled as "Tiger Cat" (right) and "Labrador Retriever" (left). A visual inspection of such things seems sufficient to confirm that states presenting these things would not in fact be presenting any tiger cat or Labrador retriever to any subject. (Reprinted from Olah et al., 2018, reused under CC-BY)

The features visualized are *labelled* as "beagles" and "tigers". But they are *nothing like* the contents of our mental states representing tigers and beagles. If when you try to think of a Beagle you were to end up thinking *these things*, you'd likely see a neurologist. And this is no isolated case — generally the features "visualized" by such techniques don't even remotely resemble our mental contents.[47]

Content also exhibits apparent causality.[48] We "see" contents causing our behaviors all the time: my desire for another beer makes me walk towards the bar, and my belief that drinking and driving is unsafe makes me leave my car keys home. But we also know that content itself can hardly be the cause of my movements and behaviors. All the causal work seems to be made by its "vehicle" — the physical (neural) state carrying content around. It is that state, the particular firing of some of my neurons, that makes me order another beer, or that makes me leave the car keys on the table, back home. Subtract these physical states from the vignette, and you will subtract my behavior. Subtract the contents, and I will move just the same. You can, for example, imagine that I have recently died (just before hitting the bar), and that a mad scientist is using some electrodes in my brain to make me move, Galvani-style.

Notice that the causal inertia of content is an upshot of the fact that (most) philosophers endorse some form of externalism about content: content, they hold, is grounded in factors that are *external* to the subject and thus extrinsic to the subject's state. This leaves the door open to atom-by-atom duplicates of a subject that are nevertheless embedded in an environment wherein the relevant external factors do not obtain and are thus contentless. Being atom-by-atom duplicates of a normal subject, these duplicates *will duplicate any causal relation the normal subject has with the environment*: the subject and the duplicate will behave identically, react in identical ways to identical stimuli, and even produce identical reports.[49] Yet, the duplicate *simply has no content*, strongly suggesting that content is *not at all causally involved* in the production of our behavior. Davidson's (2016) swampman scenario offers the most famous example of such duplicates — and the one that most often lets content realists concede the point (e.g., Millikan, 1993; Dretske, 1988; Shea, 2018). But similar thought experiments could be replicated for *other* theories of content, leading to a similar conclusion.[50]

Realists, whether conservative or revolutionary, have a hard time accounting for these features of content. Conservative realists take contents to be regular parts of the natural world. But such parts are typically accessible from a number of different perspectives, and causally efficacious. Why shouldn't content be? Conservative realists, as far as I can see, have

---

47 For another example, see https://ai.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html.

48 In the efficient sense of causality. Content may still be causal in some other sense.

49 Such a duplicate would thus be, in a way, an intentional zombie - the analog of a philosophical zombie in regards to mental content (rather than phenomenal consciousness).

50 If at all needed: many theories of content nowadays are some sort of variation on teleosemantic themes (e.g. Neander, 2017; Shea, 2018; Heemskerk, 2024), and are thus vulnerable to the Swampman counterexample.

never offered a successful answer to this question. Revolutionary realists could, in principle, sidestep these problems by positing novel, special objects (contents) that are accessible only through introspection and that are causally insulated from the rest of the world. But I take that most revolutionary realists would like not to. After all, semantic primitivists such as Burge (2010) give contents a role in the mind sciences, and thus in the causal explanation of our behavior. Presumably, such a role would contrast with content lacking any causal powers. Similarly, phenomenologists seem to take our most basic experiences to be grounded in our bodily interactions with the world (e.g., Husserl, 1970; Heidegger, 1926/1962; Merleau-Ponty, 1963), which suggests that our mental contents are causally influenced by bodily goings on — a view that, again, clashes with the causal insulation of content. In contrast, content illusionists have a straightforward time in accounting for them. After all, the illusionist claims that contents *are illusions*. The fact that continents have the prototypical feature of illusions, then, should be quite expected.

3.7. Summing up

To recap, I have offered six general reasons to prefer content illusionism over realism. Content illusionism (1) is more parsimonious than realism, it (2) better fits with the mind sciences, (3) affords better explanations of our behavior, (4) allows us to trade away a metaphysically hard explanatory target for an easier one, (5) avoids the puzzles and pitfalls of realism, and (6) better accounts for certain relevant features of content. Taken individually, no such reason is powerful enough to motivate the adoption of content illusionism. Nor are they intended to. Even taking them together, I don't want to suggest that these six reasons offer a *knockdown argument* against realism — they clearly don't. Still, together, these six considerations indicate that content illusionism has significant attraction over realism, and should thus be preferred to it.

4. Defending content illusionism

I shall now try to defend content illusionism from some of the most obvious objections to it. I will here consider three objections.

4.1. Objection #1: Content illusionism denies the data

The first objection contends that content illusionism *denies the data*. According to the objector, mental contents are simply and undeniably given to our awareness. Any denial of this is extremely implausible, and methodologically unwise. Surely our first-personal, introspective knowledge of our mind is an important — if not *primary* — source of data concerning our mind.[51]

The illusionist's reply is threefold: (i) it is false that content illusionism denies the data; (ii) even if contents were given, the objection would fail; (iii) no methodologically pernicious move is being made. Let me expand.

(i) Content illusionism does not deny any *data*, if "data" means what it usually means: the result of third-personal, public and repeatable observation or measurement of *something*. In the case at hand, the data are introspective reports concerning mental contents, and seemingly content-involving behaviors (see Dennett, 1991b), whose existence the illusionist does not deny.[52] To the contrary, they are central to the illusionist — for they are a constitutive part of the il-

[51] This objection is inspired by an analogous one leveled against illusionism about phenomenal consciousness (e.g. Searle, 1992, 1995; Strawson, 2004).

[52] What if by "data" one means something else – like a private apprehension of something mental (cf. Searle 1995)? Well, these are not data in the common, *unproblematic* sense of the term. On the contrary, these "data" are highly contentious entities whose existence is *explicitly*

lusion of content, which is the illusionist explanatory target. Of course, the illusionist denies *the truthfulness* of these reports, as it entails the existence of content (which the illusionist negates). But to deny this is not to deny the data. Compare: no one would claim that the humanist Lorenzo Valla denied the data when he exposed Constantine's Donation as a forgery. He only accounted for the data in a different way.

(ii) Allow — *only* for the sake of discussion — essentially private data to be relevant in the discussion. What does it mean that these data are *given*? More precisely, should "being given" be read as a success (factive) term? This is a yes/no question that admits only two answers — none of which puts the desired pressure on illusionism.

If the answer is no, then "being given" works roughly like "being believed": something can be given, or be believed, without being real. But then the fact that contents are given (in this sense) poses no challenge to illusionism. Indeed, the illusionist *could* reasonably say that the illusion of content consists in contents seemingly being given to awareness without these contents being real!

If the answer is yes, then "being given" works roughly like "being known": something can be given, or be known, only if that thing is real. That's why from the *knowledge* of p we can infer that p is the case. So, knowing p *presupposes* that p. Hence, were content given in this factive sense, its existence would be presupposed. But to presuppose the existence of content just is to presuppose the falsity of content illusionism. Hence, were "being given" factive, the objection would be circular.[53]

(iii) It is also false that our "first-person" knowledge of our mind is an important, or even *primary*, source of data in the study of the mind. Psychological enterprises relying on such knowledge historically *failed*, and the study of the mind has progressed mainly through *third-personal*, indirect measures of mental activity. This holds even for the subjective phenomenon *par excellence*, namely phenomenal consciousness (see Seth, 2021).

So, content illusionism does not deny the data – it only accounts for them in a different way. At this point, it is easy to imagine that the realist would still argue that a problem persists: the realist can immediately (and seemingly unproblematically) call for *an abductive inference to the best explanation* to infer the existence of content — what else could explain the production of the relevant report? Surely the illusionist must concoct some sort of complex, highly *ad-hoc* account to explain it. And such a complex explanation can't be better than the simpler realist one.[54] But it can, as I will argue in §5.

4.2. Objection #2: The "no appearance/reality distinction" objection

The second objection argues that content illusionism is wrong because, just like phenomenal consciousness, content leaves no room for the appearance/reality distinction. If it seems to me that I am thinking of a green apple, then I'm really thinking of a green apple. Of course, I may be wrong — the apple I was thinking about turned out to be the tiniest of watermelons. But even in this case, my thought would be wrong only because its content was *apple* rather than *watermelon*. So, if it seems to one that one is thinking about x, one is really, *infallibly and eo ipso* thinking about x. There simply is no room for error.[55]

I must confess I'm really baffled by this line of reasoning. It's quite easy to imagine a case of error in regards to content.[56] Imagine, for example, a super-neurosurgeon, able to mess with a person's introspective mechanism in the follow-

---

denied by illusionism. The objection, then, would be a re-statement of the eliminative claim (§2.2), and so it would not be an *objection* at all – at least, not until an independent argument for the existence of these "data" is provided.

53 Identical considerations, it seems to me, hold if "givenness" is substituted with "acquaintance" (e.g. Balog, 2023) - or any other "direct, unmediated" epistemic relation.

54 Thanks to a reviewer for pressing this point.

55 For an analog point concerning phenomenal consciousness, see for example Kripke (1980).

56 And phenomenal consciousness: see Dennett's (1988) "alternative neurosurgery" scenario, from which I'm loosely drawing here.

ing way: whenever the person is in a state whose content is related to apples, the mechanism introspects a state whose content is related to oranges, and vice versa. So, when a person who underwent this surgery will think about *apfelstrudel*, she will token a state whose content is a particular apple pie typical of the alpine region — but this person will introspect (and report) that they are thinking about an orange pie. And when this person sees a glass of orange juice, the person will introspect seeing a glass of *apple* juice — and so forth. Of course, such fine-grained interventions are (currently) not real — but they obviously are physically possible. It's thus quite easy to imagine a case where a person introspects one content when they actually have another. There is a space for the appearance/reality distinction.

The realist may reply that the illusionist is not entitled to use such a scenario to defend their position, as the scenario seems to presuppose the existence of certain contents, including the contents of introspective states. The point is well taken, but not fatal to the illusionist, for at least two reasons. First, when articulating the point, the illusionist can simply grant the existence of content for the sake of discussion – and then proceed to show that the realist is not entitled to the objection, given that they must allow for an appearance/reality distinction given their own theoretical commitments. Secondly, it's relatively easy to modify the scenario sketched above in a way that does not presuppose the existence of any content. Indeed, it is sufficient to change what the introspective mechanism monitors (and does) to make the scenario compatible with content illusionism. Thus, imagine that rather than detecting contentful mental states, the introspective mechanism is sensible to the — surely enormously complex — *dispositions to behave* towards apples and oranges.[57] Imagine further that the introspective mechanism does not produce contentful states, but "reroutes" the agent behavior; that is, when the introspective mechanism tracks the activation of apple-related behavioral dispositions (including the disposition to report apple-related contents) it deactivates those and activates orange-related dispositions (and vice versa). Thus, the mechanism sketched above can easily be made compatible with content illusionism, and so there's nothing *problematic* in the illusionist appealing to it.

4.3. Objection #3: Content illusionism entails semantic nihilism

The third objection argues that content illusionism cannot be *true*, because, if the world were like content illusionism describes there would be no content, and thus nothing that can be true or false. After all, it seems that the meaning — and thus the truth — of what I am writing hinges upon our (shared) interpretation of these signs; that is, it seems to hinge on (presumably contentful) mental acts of interpretation. Eliminate those, and eliminate the meaning from the writing — and thus, the potential truth or falsity — of what I am writing.

The content illusionist can react to this challenge by noticing that it hinges on a common, but fairly substantial, presupposition — namely, that mental content is (directly or indirectly) the source of all content, meaning, and every semantic evaluable property (e.g., Grice, 1989; Aizawa & Adams, 2005; Mendelovici, 2026).

Yet, this assumption should not be granted — at least, not without argument. On the one hand, the assumption presupposes a specific, typically conservative realist (see references above) view of mental content as the ultimate ground for all content. But, as noticed above, conservative realism is not in good health, and so it cannot offer us a reason to espouse this assumption. Nor does the assumption seem particularly "natural" or "obvious" to make — the question of whether certain (non-mental) states have intrinsic, mind-independent content is fairly old and admits a variety of solutions (e.g., Plato, 1997). And, as weird as it may sound, some terms almost universally seem to be "natural names" for certain things, as attested by the kiki/bouba effect (see Cwiek *et al.*, 2022). But even without resurrecting an (admittedly quite cryptic) Cratylean view of "natural significance", the content illusionist can allow for semantic properties

[57] Wouldn't these dispositions be so multifaceted and unruly to make necessary for the introspection mechanism to transduce them into some sort of inner representational medium (for this point, see Clark & Toribio, 1994)? Not at all. Ordinary, non-representational devices can unproblematically yield the same output in response to a multifaceted and unruly input class. See (Ramsey, 2007, pp. 219-220; Facchin, 2026) for examples.

not grounded in mental content in at least two ways. Illusionism, in and by itself, is neutral in respect to both (it doesn't compel or favor one over the other).

First, the illusionist can join a broadly *pragmatist* tradition which takes semantic properties to originate from specific social, cultural practices (e.g., Wittgenstein, 2009; Haugeland, 1990; Brandom, 1994; Hutto & Satne, 2015). The illusionist can simply help themselves with one such account, and dodge the objection. And notice that, even on the quite dubious assumption that such an account is an account of the origins of contents in the relevant sense discussed here (i.e. content as "something" carried around by contentful states, see Myin & van den Herik, 2021), the account itself does not confer content to our mental states. After all, our cultural practices do not typically revolve around the neuronal states that should carry the relevant contents, and so would not assign contents to them.

The fact that the content illusionist has such a move to their avail has one important consequence I wish to highlight here: it allows the illusionist to also block a different intuitive realistic objection. The objection, roughly, contends that content illusionism cannot be coherently formulated, for illusions are *misrepresentations* (e.g., Macpherson, 2013): they are representations *whose content gets things wrong*. If this were the case, then, the illusion claim (i.e. mental content is an illusion) would presuppose exactly what the eliminative claim (i.e. there is no mental content) denies. Hence content illusionism would be internally inconsistent.

I already dealt with this objection in §2.2 when I elucidated the illusion of content in behavioral terms: to be under the spell of the illusion of content is to behave so as to satisfy the stereotype of the believer that one's mental contents are thus-and-so. Some may find these terms objectionable — for example, because they hold that the identification of such a behavioral stereotype still presupposes content. Such a burden (if at all present) can now be discharged by characterizing the illusion of content in terms of public representations (see Toon, 2023 for a similar idea). To be under the illusion of having some visual content, for example, is *to behave as if one has an "inner picture" of the external world*. To be under the illusion of having a mnemonic content is to behave *as if one was able to retrieve knowledge from a sort of "inner notebook"*. To be under the illusion of having a contentful imaginative state is *like painting a mental painting*. Dreams are like *inner movies*, and to think *is like talking to oneself*. Such metaphors offer the familiar basis through which we understand our own contentful mental states, and can aptly characterize the illusion of content in a way that *does not* presuppose mental content.[58]

Going back to semantic nihilims, there is a second move the illusionist can try to identify semantic, or at least proto-semantic, notions. After all, even in the absence of content, there can still be *functional success* and/or failure conditions that can be treated as at least proto-semantic under a broadly correspondence notion of semantics. A world inhabited only by paramecia would (most likely) be a world with no mental content, but the organelles inside the paramecia would still have functions, which they can still fail or succeed to perform. For example, ribosomes would still have the function of producing proteins by "reading" RNA sequences, and RNA bases would still have the function of "copying" DNA sequences. And these forms of functioning seem to presuppose a form of structural matching which can be usefully treated as a semantic or proto-semantic, natural form of "picturing" (see Sachs, 2019) — even if we are in no way tempted to think that RNA sequences somehow carry a "content" (a sort of special, *sui generis* object somehow representing or replicating the protein to be produced).

I have here defended content illusionism by three of the most obvious, most direct objections against it. At this point, content illusionism may appear as a genuine contender when it comes to dealing with mental content and its philosophical problems. Or, maybe, it would appear as a genuine contender if I had provided a positive story about illusionism — concerning, for example, how the illusion is generated. But, thus far, I haven't.

Time, then, to fill in this gap.

---

[58] Won't this characterization of the illusion clash with the behavioral one? I don't think so. To fit the behavioral stereotype is, in many cases, to behave as if one is guided by a number of mental analogs of external representations.

5. The illusion of content: Sketching an explanation

The aim of this section is to sketch an explanation to the illusion of content (an "error theory", as some philosophers might prefer to call it). As is common in the illusionist literature (see Dennett, 1991b, Frankish, 2026), the explanation will take the form of a mechanism.[59] The functioning of such a mechanism will, as I shall show, cause us to adopt the relevant behavioral stereotype of a person believing that they have mental contents thus-and-so.

Owing to space limitations and the programmatic character of the paper, I will be light on empirical details. For this reason, I will also try to sketch the *simplest possible* mechanism generating the illusion. As said above (§ 2.2) it is possible to call upon other, different and more complex mechanisms. For now, however, this simple one must do.

Notice that whilst here I aim at providing the simple possible mechanism, I don't think my explanation is a *how-possibly explanation*. Whilst surely simplified, the mechanism I am sketching is made up of two components that *are actually present* in our cognitive architecture. Let me introduce them.

5.1. Component #1: An internal controller

Some devices exhibit a behavior which is entirely controlled by environmental variables. So called "passive walkers", for example, are simple robots whose walking behavior is guided by the slope they are placed upon (McGeer, 1990).

Most animals are not devices of this kind. Whilst they remain open to environmental influences, they also have an inner *controller* enabling them to flexibly interact (and partially decouple from) environmental variables: the brain. More in detail, what such a controller must minimally do is to *keep track* of the organism's bodily variables (i.e. the organism's bodily state — think of it as the set of receptors and actuators "wired" to the brain), and keep them within precise bounds of viability. And notice that this "keep track" can simply take the form of detecting — and correcting for — deviations from a "default" state (Aikins, 1996; Bruineberg *et al.*, 2018). For example, if the appropriate receptors notice a deviation from the desired bodily temperature of 36.7°, the brain can correct for it inducing rhythmic muscle spasm (trembling), which will, in some circumstances, raise the temperature. Or, in the case of a more sophisticated control system, the correction can happen by operating on one's central heating system. The predictive processing theory of perception mentioned above (§3) offers a modern — and importantly *representation-free*[60] — rendition of this idea (see Clark, 2013, 2016; Facchin, 2021a,b; Seth, 2014b; Williams, 2018a,b; Bruineberg *et al.*, 2018). But the idea that the brain is such a control system predates it (e.g., Powers 1973; Ashby, 2013; Clark, 1997; Anderson, 2014) and seems to be certain beyond any reasonable doubt.[61] What is relevant to my purpose here is that a controller of this kind must exist (else animals would have a hard time keeping themselves alive) and that it must constantly keep track of the organism's physiological states: its bodily variables and the state of receptors and actuators.[62]

5.2. Component #2: Agent-detectable manipulative social signalling

Most animals must coordinate with other animals (of same or different species) to some end. This need for coordination gives rise to "signaling games" (Skyrms, 2010). Descriptively, the point is very familiar: green vervets emit specific

---

59 Really at the level of generality I will operate, it is a mechanistic sketch (Piccinini & Craver, 2011).

60 See references given in §3. The "modeling" talk that abounds in the predictive processing literature uses a sense of model that comes from the "good regulator theorem" (Conant & Ashby, 1970) and which is markedly non-representational. See Thobani (2024).

61 Of course this is not to claim that the brain performs only that role as a controller - a claim that runs into a vast number of psychological (Sun & Firestone, 2020) and neurological (Facchin & Viola, 2025) counterexamples.

62 Such architectures are also importantly entangled with illusionism about phenomenal consciousness, see (Dennett, 2013; Clark *et al.*, 2019; Frankish, 2026).

calls to "invite" other vervets to run away from specific predators (Seyfarath *et al.*, 1980); bees "dance" to signal to other bees the location of flowers to "harvest" (Von Frisch, 1993), *homo sapiens* emits sound waves to coordinate with each other in a range of situations, and so forth.[63]

Signalling games are at least at times open to *manipulation*: a participant of the game may emit a signal when the external circumstances *do not* call for its production, thereby forcing other participants to act inappropriately — which typically results in some (perhaps short-term) advantage for the signaler. A smart vervet may emit a predator call when no such predator is in sight, to force other vervets to run and stop a losing fight.[64] Humans lie. Some fireflies mimic the mating signals emitted by other species of fireflies — thereby luring their prey (Lloyd, 1965).[65]

Crucially, in at least some cases, such manipulations are *agent-detectable*: they have physiological markers the agent emitting them (or, its control system) can keep track of, and act upon. This makes good sense. Consider the vervet case above: the vervet emitting the call does so to "force" the fight to end, to regain energy and perhaps "steal" the good over which the vervets were fighting as other vervets run to hide. If that vervet were to run too, they would be very poorly served by their attempted manipulation! In the human case, such markers are many, and range from physiological changes in skin conductance to particular patterns of neural activity (Godert *et al.*, 2001, Meijer *et al.*, 2014; Mandal, 2014).[66]

5.3. The mechanism at work

To see how these two components interact to generate the illusion, consider the following — no doubt rather simplistic — scenario. Let S_1 and S_2 be two humans engaged in a common activity. We can imagine them being prehistoric humans hunting for a bison — this will help us keep their mind and signaling games simple, and keep S_1 and S_2 as culturally unsophisticated as a human can be. Thus, imagine that, with some reliability, S_1 spots bisons, signals their presence to S_2, and then the two proceed to hunt their prey.

Let me now vary their routine. Imagine that, by some accident (*which* accident is not very important), S_1 enters in the physical states of action readiness that follow the spotting of a bison *when no bison is actually around* (and so no bison can have been spotted). S_1's brain is therefore preparing and "launching" all the relevant bison-related (motoric and non-motoric) responses: S_1 is thus emitting the relevant signal (e.g., turning towards S_2 saying "Bison!") and starting the hunting routine (which includes a number of bodily movements and physiological changes, such as the release of adrenalin). Yet no bison is there. Thus S_1's internal state and behavior will fail to *evolve* in the usual way, and S_1 (or its brain) will register the fact. Roughly put, there is no bison-related sensory stimulus incoming, and S_1's brain will arrest the relevant motor routine and physiological changes just launched.

S_1 would, at this point, be quite perplexed. S_1 sees now that there is no bison out there: his brain is not tracking the presence of bison-related states at the sensory periphery, if you wish. Nor has his brain tracked any bison-moving-away related sensorimotor state — so S_1 knows the bison has not run away. So, S_1's behavior fits the stereotype of the believer that there is no bison out there (and indeed, S_1 is not trying to track the bison, to run away from it, to surprise it in an

---

[63] Aren't such signals contentful? Perhaps. But these signals would at best be external (non-mental, public) representations, which the illusionist can allow for (§4). Do signaling games presuppose mental contents - for example, communicative intentions, or representations of communicative goals? Arguably, no. Bacteria take part in - sometimes very complex - signaling games (Diggle *et al.*, 2007; Hasrem & Van Impe, 2022), and presumably bacteria do not mentally represent a thing.

[64] Anecdotically attested in Cheney & Seyfarth (1990, pp. 213-216).

[65] Again, fireflies do so presumably without having to internally represent anything, so again manipulative signaling does not presuppose content.

[66] Can elements outside the nervous system of an agent be components of the mechanisms responsible for the agent's mentality? The short answer is: yes (Kaplan, 2012). Components of mechanisms need not be spatially contiguous, or part of the same anatomically defined system.

ambush, etc.). But S_1 will also fit the stereotype of the believer *that they saw a bison* — their brain didn't register *any* of the markers of manipulative signaling. Thus, if asked, S_1 would defend the claim that they really saw a bison, that the signal they omitted was no joke, that they themself was ready to attack the bison, and so forth. More simply put: S_1's behavior is such that we are licensed to say that S_1 believes that a bison was seen, but there is no bison out there.

The obvious conclusion — that is, the behavioral stereotype S_1 is forced towards by the behaviors above — is that *the seen bison must have been "in here"*, in S_1's mind. Thus, S_1 will say that they saw a bison that nevertheless wasn't out there (but really *seemed* out there), that S_2 wasn't able to see such bison, and that when they looked at the spot again the bison simply wasn't there anymore. It's easy to see how a little bit of cultural evolution could lead S_1 to say he had a sort of "private mental picture" of the bison — and this is roughly our folk conception of perceptual content (Fischer *et al.*, 2026).[67] But even without such cultural scaffolds[68], it is clear that S_1 is now pushed towards the belief that there are certain mental "objects" that are not "out there" but rather inhabit our minds. That is, S_1 is pushed towards the behavioral stereotype of a believer in such objects: S_1 will talk about them, say that they are present — perhaps even given — to their mind, produce genuine introspective reports about those, and so forth.

I could add bells and whistles to this mechanism sketch. For example, adopting a "predictive processing" view and its idea of precision may explain why we are *particularly certain* that mental contents exist (cf. Clark *et al.*, 2019). Calling upon a specific socio-material embedding may explain why the illusion has, for us, the specific features it has (why, for example, we take perception to be some sort of picture, cf. Draisma, 2000; Toon, 2023). But these are, in an important sense, bells and whistles. The core of the mechanism consists in the non-manipulative production of signals that falsely express and report what's going on in our mind.[69]

5.4. Illusionism as the best explanation

In §3.1 I argued that illusionism is quantitatively and qualitatively more parsimonious than realism. I also asserted (without much of an argument) that it is also more mechanistically (and thus, explanatorily) parsimonious. Moreover, in §4.1 I mentioned that the realist could uphold their belief in content on the abductive grounds provided by an inference to the best explanation. And while above I anticipated something in the way of an answer, my points were left largely unsubstantiated. Effectively, I took an *argumentative loan* for the sake of exposition. It is now time to start paying that loan back.

Let me then start with the shape of the realist's alternative. In a nutshell, the realist's alternative is a simple *detection mechanism*. The mechanism has some part that immediately detects the relevant content, and an output channel whereby the detector mechanism can influence the production of the subject's reports concerning their mental contents. Such a mechanism surely seems simpler than the one I sketched above, and so allows us also to offer a better — simpler and more direct — explanation of our reports.

Isn't such a simple detection mechanism *obviously* mechanically simpler than the mechanism I sketched above? How can the illusionist argue that this mechanism is more *mechanistically complicated* than the one I sketched? I cannot deny that the simple detection mechanism looks simpler than the one I sketched. Once we try to think about the mechanism's *details*, however, the apparent simplicity rapidly disappears.

---

[67] Indeed, the picture is so popular that it's also shared by the non-folk. For example, Chris Frith is a neuroscientist having a sophisticated and complex theory of our minds, and yet still he subscribes to that picture of content (Frith, 2007, p.135).

[68] Their mention is suggestive of an "extended mind" (Clark, 2008) picture of the illusion, where the illusion is at least in part constituted by our external representational practices and our reflection on them (cf. Dennett, 2000). This is a fascinating theme I must pass under silence for the sake of brevity.

[69] This is also what John was doing in the example above (§2.2).

The first, and most obvious problem we run into is that the realist *must* commit to all the components of my story too. These components are so simple, undemanding — and I dare to say *obviously real* — that to deny their existence seems if not impossible, at least a super-herculean task. To deny that brains track and monitor bodily states, or that there are manipulative signals in signaling games one must have a pretty revolutionary theory of brains and communicative behaviors. And to motivate the adoption of such a theory *is on them*.

Noticing this is already sufficient to "cash in" the previous point concerning mechanistic simplicity. The realist is prima facie committed to the existence of all the components I have pointed out above, to which they add a commitment to the simple detection mechanism I outlined above.[70] The realist, then, puts *more mechanisms* in our head, giving us a *more mechanistically complicated* picture of our cognitive architecture.

And, once we think about *how* the simple detection mechanism gets "wired" to the ordinary circuitry of the brain, the complexity becomes quite dazzling. For, the simple detection mechanism is something like a "Cartesian theatre" (Dennett, 1991b): a mechanism *isolated from* the ordinary mechanism of the brain, to which all contents converge and are shown for our benefit. Contents must thus be forwarded to the detection mechanism (or the detection mechanism must be able to move around the brain), complexifying the (already quite complex) connection architecture of our brain — to the point that such a centralized detection mechanism may be *biologically* impossible (Mesulam, 2008). And notice that making such a mechanism a "virtual" mechanism, or part of the brain "software", does not solve the problem. According to all physical theories of computation (Piccinini, 2025) there must be an *isomorphism* between computational states and the substrate implementing them. Thus, were such a "virtual mechanism" implemented in the brain, *something in the brain* should go through the relevant state changes of the virtual mechanism. There would still be a piece of the brain wherein all content must be forwarded. Thus, the realist actually *complexifies* the overall picture of mental mechanisms, adding a component. In contrast, the illusionist does not call upon ad hoc mechanisms, and explains the illusion with "the regular" mechanisms of the mind and brain (cf. Dennett, 1991b, 2025; Frankish, 2026).[71]

But the above is only the first problem (or, to be more precise, set of problems) the realist must face. A second problem (or set thereof) concerns *content detection*. The realist must in fact offer an account of it, and yet it is very hard to see how such an account could go. Normally, detection happens in virtue of the fact that the detected state causally influences the detector. Some examples to clarify: litmus papers detect the Ph. of a substance because the substance Ph. causes them to change color. (Bulb) thermometers detect temperature because the temperature causes the fluid in the thermometer to expand or contract, going "up" and "down". Optic smoke detectors detect smoke (and thus, fires), because the smoke enters the inner chamber of the detector, causing the beam of light inside it to scatter (which causes the detector's photosensitive cell to ring the alarm).

As argued above, however, content seems to lack the causal powers needed to causally influence anything. This means that the realist either lets go of the idea that contents are detected — which would of course force us to reject the simple model sketched above — or comes up with some *non-causal* story about detection. I don't see what form such a model may take, and more generally this seems a significant complication on the simple story the detection mechanism allegedly affords.

Thirdly and lastly, the simple detection mechanism *can't be the whole story about our introspective reports*. It is well known that some introspective reports are confabulations, and thus they report contents that have not (indeed, cannot) have been detected. Some of the most famous cases include neurotypical subjects making up on the spot some "in-

---

[70] This inflation of commitments is of course also made more acute by the realist's commitment to the existence of content.

[71] These problems are also compounded by the fact that this mechanism also must interface and communicate with all other neural and cognitive mechanisms, in a way that can generate all sorts of conflicts between them. Consider again the "bison" example sketched above. Why hasn't the simple introspective mechanism vetoed the utterance of "bison", given that no corresponding visual content was introspectable (notice that, by construction of the case, S1 enters in the states that follow the perception of a bison, so no "visual" bison content can be introspected)? These kinds of conflicts are a problem for the realist to solve - and their solution is guaranteed to further complicate the realist's story. Dennett's (1991b) discussion of Orwellian and Stalinesque models offers a vivid illustration of (some) such problems.

trospected preferences" to justify their choices (Nisbett & Wilson, 1977), and the case of split brain patients confabulating a reasons for their actions when the stimulus motivating them was given to the right cerebral hemisphere, which is unable to produce verbal reports (Gazzaniga, 1995).[72]

Indeed, introspective reports are likely produced by many different mechanisms (Schwitzgebel, 2012), generating a multiplicity of reports and a variety of introspective practices. Some reports are confabulations, invented "on the spot" to justify our behaviors (Chater, 2018). Other reports conform to our socially shared narrative practices (Hutto, 2012), allowing us to partake in a social community.[73] Other reports are free acts of self-exploration, whereby we strive to interpret and make sense of how we physiologically feel (Facchin *et al.*, 2025). Other reports are optimistically biased to make us self-confident and keep us going, no matter how objectively bleak the situation is (e.g., Sharot, 2011). Reports may simply express how we feel and what is going on inside our minds (e.g., Wittgenstein, 2008), or the upshot of our decisions (Moran, 2001), or even be an attempt to mold us and our behavior into a form we find more congenial to us (Hutto, 2016; 2022). Our introspection, and the rich mental life it purportedly shows, is manifestly *not* just the effect of a mechanism that detects what's currently going on in our mind.

Sure, to acknowledge such a complexity — and such a variety of introspective practices — *does* force me to complexify the mechanism I sketched above too. But notice how such a complexification is not really damaging for the illusionist.

On the one hand, the basic story I sketched about the illusion can be replicated. Regardless of the introspective practice and the specific mechanism generating them, *if* the participant produces a genuine (non-manipulative, in the sense seen above) report, then the participant will end up *fitting the behavioral stereotype* of the believer that their mental content is thus-and-so. And, if there is no such content (as I have argued here), then the participant is under the illusion. The illusionist, then, can simply (roughly) repeat the same explanatory story for each and every introspective practice. The realist cannot. They need to posit different, additional mechanisms. As a result, their explanation will also be way less unified and homogeneous.

On the other hand, whilst such a plurality of mechanisms and introspective practices is surely compatible with realism (cf. Schwitzgebel, 2012), it is not really *predicted* by most realist accounts. Realists tend to have a monistic view of introspection, where introspection is the output of a single mechanism. This is indeed the basis of the various "-isms" competing to offer accounts of introspection. For example: expressivism explains introspection in terms of the expression of mental states (e.g., Finkelstein, 2008), whereas inferentialists explain introspection calling upon inferential mechanisms, perhaps operating on a theory of mind or folk psychology (e.g., Gopnik, 1993). Illusionists, in contrast, have historically expected such a plurality of mechanisms. Dennett's (2005, pp. 72-76) *tuned deck* story colorfully conveys the point. The story's protagonist is the conjurer Ralph Hull, apparently able to play a "tuned deck" trick no other conjurer was able to replicate, or even understand. Later, Hull revealed the trick: there simply was no "tuned deck" trick! Hull had at his avail a *plurality* of (already well known) different moves yielding the same output (e.g., guessing the card a member of the public chose, or whatever). When Hull announced he was going to perform the "tuned deck" trick, he simply performed one of those tricks — cycling through the moves made between each iteration of the "trick", which made the tuned deck trick impossible to track: when another conjurer were looking for a move A, Hull made a move B; when a conjurer were controlling for a move B, Hull performed a move C — and so forth.

By analogy, illusionists expect the illusion to be the product of multiple different mechanisms — which make the illusion hard to disciple, as there is no single mechanism always responsible for it we can always call upon to explain content away.

---

[72] See also Chater (2018) for several other cases. Also notice that such confabulations are not things that subjects willingly make up. They indeed reflect the subject's "honest" view of reality (see Schneider, 2003).

[73] Mesquita (2022 pp. 200-203) offers a striking example with the case of Ramla, a Muslim girl that, in virtue of his Muslim upbringing (and the emphasis on mutual social obligations such upbringing placed on her) reported feeling ashamed in many instances wherein a westerner would have rather felt sad.

The above, I take, makes a case for illusionism being *less* mechanistically complex than realism. The very same considerations help the illusionist respond to a different realist's counterargument; namely that the existence of contents can be abductively inferred by the *best explanation* of their apparent givenness.

The first step in this direction is that of noticing that the *existence* of contents isn't in and by itself sufficient to explain their *apparent givenness*.[74] Hence, appealing only to the existence of content does not explain (let alone best explains) their apparent givenness. To see why this is the case, imagine an agent having only non-introspectively detectable contentful states (for example, an agent having only "neural", sub-personal representational states). Were such an agent to exist, content would (by construction of the case) exist. And yet, the mere existence of content would not explain why content appears to be given to the agent. Indeed, in the scenario described, these constants wouldn't be given (apparently or not) to the agent *at all* — for the agent has no way to access them (again, by construction of the case).

So, in and by itself, the existence of contents *does not* explain their apparent givenness. The existence of content can of course be *part* of the explanation of why content seems given only if it is joined by something else; namely a mechanism that allows the agent to access their own content. And the detection mechanism I sketched above is the simplest mechanism at the realists' avail. And, as I argued above comparing it to its illusionistic alternative, that mechanism does not offer a *better* explanation of the apparent givenness of content.

The realist could of course reply that the simple detection mechanism I have put forth and examined is *too* simple, and that some more complex and realistic mechanism would indeed provide a better explanation. Indeed, the space of possible introspective mechanisms is quite large (see Kammerer & Frankish, 2023) — and whilst space limitations forbid its exploration here, I must confess I have no straightforward argument to *deny* that there is no *simpler* mechanism able to satisfy the realist.

Perhaps, then, it is convenient to put the point as a *challenge* to the realist: the realist has to provide a mechanism that is (a) simpler than the one I sketched, and (b) able to overcome the problems I have highlighted. I have a hard time imagining how the realist could do that — and it is on the realist to prove that I'm wrong.

6. Conclusions

I have introduced content illusionism to the literature. I have motivated it, defended against some obvious objections, and sketched a mechanism responsible for the illusion. Of course, this is just the first step. In an important sense, the real work to develop and defend content illusionism begins now — after the more programmatic points have been made.

In particular, it is necessary to:

(i) provide a fuller, more detailed mechanistic sketch of the mechanisms responsible for the illusion of content.
(ii) elaborate a richer and clearer picture of what the illusion of content amounts to and its origins, both developmentally (are we born with it or we acquire it?) and phylogenetically (presumably, it is not coextensive with life and it has "popped up" on *some* branch of the tree of life).
(iii) gauge the epistemological implications of such a view. If there are no contents, how should we understand belief and knowledge? How should we understand rationality and inference?

[74] Of course, were the realist able to demonstrate that contents exist, they would prove that illusionism is wrong. But, in the objection I'm considering, the realist is inferring (abductively) to the existence of content based on the fact that it best explains their apparent givenness. And I'm contesting that the existence of content alone offers such an explanation.

(iv) explore the relations between content illusionism and illusionism about phenomenal consciousness. It is intuitive to think that the two views must be coherent and mutually supportive — and yet illusionists about phenomenal consciousness heavily rely on representations in their explanatory practices (e.g., Kammerer, 2021).

There is, then, plenty of work for illusionists to do. I hope the present paper has persuaded you this is *valuable* work, and enticed you in developing your view of content in an illusionist direction (*especially* if you are a radical enactivist).

Acknowledgements

An early version of this paper (in late '22, early '23) was read and commented upon by (in random order) Bruno Cortesi, Valentina Martinis, François Kammerer, Angela Mendelovici, Thomas van Es, Zuzanna Rucinska, Erik Myin, Farid Zahnoun, Giulia di Rienzo, David Schoute and Wouter van Hooydonk. Thank you for your comments and questions. The paper in this version has been presented to the *Expaning Mind* seminars in Wollongong. I thank the whole crowd of the Gong, but a special thanks goes to Dan Hutto and Erik Myin for their engagement during the seminar.

This work has been financed with the generous FWO (*Fonds Wetenschappelijk Onderzoek*) grant "Towards a globally non-representational theory of the mind" (grant number 1202824N)

REFERENCES

Aikins, K. (1996). Of sensory systems and the aboutness of mental states. *Journal of Philosophy, 93*(7), 337-372.

Aizawa, K., & Adams, F. (2005). Defending non-derived content. *Philosophical Psychology, 18*(6), 661-669.

Amir, D., & Firestone, C. (2025). Is visual perception WEIRD? The Müller-Lyer illusion and the cultural byproduct hypothesis. *Psychological Review*. Advance online publication. https://doi.org/10.1037/rev0000549

Anderson, M. L. (2014). *After Phrenology*. The MIT Press.

Anderson, M. L., & Champion, H. (2022). Some dilemmas for an account of neural representation: A reply to Poldrack. *Synthese, 200*(2), 169.

Artiga, M., & Sebastián, M. Á. (2020). Informational theories of content and mental representation. *Review of Philosophy and Psychology, 11*(3), 613-627.

Ashby, W. (2013). *Design for a brain: The origin of adaptive behaviour*. Springer Science & Business Media.

Baker, L. R. (1987). *Saving belief: A critique of physicalism*. Princeton University Press.

Baker, B., *et al.* (2022). Three aspects of representation in neuroscience. *Trends in Cognitive Sciences, 22*(11), 942–958

Balog, K. (2023). The abolition of phenomena: A voyage among the zombies. *Klēsis Revue Philosophique* 55. https://www.revue-klesis.org/pdf/klesis-55-07-katalinbalog-the-abolition-of-phenomena-a-voyage-among-thezombies.pdf

Baltieri, M., Buckley, C. L., & Bruineberg, J. (2020, July). Predictions in the eye of the beholder: An active inference account of Watt governors. In J. Bongard, J. Lovato, L. Hebert-Defruésene, R. Dasari & L. Soros, *Artificial Life Conference Proceedings, 32* (pp. 121-129). The MIT Press. https://direct.mit.edu/isal/proceedings-pdf/isal2020/32/121/1908564/isal_a_00288.pdf

Block, N. (2023). *The border between seeing and thinking*. Oxford University Press.

Bocanegra, B. R., Poletiek, F. H., Ftitache, B., & Clark, A. (2019). Intelligent problem-solvers externalize cognitive operations. *Nature Human Behaviour*, 3(2), 136-142.

Bourget, D., & Mendelovici, A. (2019). Phenomenal intentionality. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2019 ed.). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/fall2019/entries/phenomenal-intentionality/

Bovet, S. (2007). *Robots with Self-Developing Brains*. (Doctoral dissertation). University of Zurich.

Braitenberg, V. (1986). *Vehicles: Experiments in synthetic psychology*. The MIT Press.

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

Brentano, F. (2012). *Psychology from an empirical standpoint*. Routledge.

Brette, R. (2019). Is coding a relevant metaphor for the brain? *Behavioral and Brain Sciences, 42*, e215. https://doi.org/10.1017/S0140525X19000049

Brette, R. (2026). *The Brain, in theory*. Princeton University Press.

Brooks, R. A. (1999). *Cambrian intelligence: The early history of the new* AI. The MIT Press.

Bruineberg, J., Kiverstein, J., & Rietveld, E. (2018). The anticipating brain is not a scientist: the free-energy principle from an ecological-enactive perspective. S*ynthese, 195*(6), 2417-2444.

Carvalho, J. T., & Nolfi, S. (2016). Cognitive offloading does not prevent but rather promotes cognitive development. *PloS One, 11*(8), e0160679. https://doi.org/10.1371/journal.pone.0160679

Chalmers, D. J. (1997). *The conscious mind: In search of a fundamental theory*. Oxford University Press

Chater, N. (2018). *Mind is flat: The remarkable shallowness of the improvising brain*. Yale University Press.

Cheney, D. L, and Seyfarth, R. M. (1990). *How Monkeys See the World*. University of Chicago Press

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

Churchland, P. M. (1981). Eliminative materialism and the propositional attitudes. *The Journal of Philosophy, 78*(2), 67-90.

Churchland, P. M. (1992). *A neurocomputational perspective*. The MIT Press.

Churchland, P. M. (2012). *Plato's camera*. The MIT Press.

Churchland, P. S., *et al.* (1994). A critique of pure vision. In C. Koch & J. Davies. *Large scale theories of the brain* (23-60). The MIT Press.

Clark, A: (1997). *Being there*. The MIT Press.

Clark, A. (2008). *Supersizing the mind*. Oxford University Press.

Clark, A. (2009). Spreading the joy? Why the machinery of consciousness is (probably) still in the head. *Mind, 118*(472), 963-993.

Clark, A. (2012). Dreaming the whole cat: Generative models, predictive processing, and the enactivist conception of perceptual experience. *Mind, 121*(483), 753-771.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and brain sciences, 36*(3), 181-204.

Clark, A. (2014). Perceiving as predicting. In D Stokes, M. Matthen & S. Biggs (Eds.) *Perception and its modalities* (pp. 23-43). Oxford University Press.

Clark, A. (2015). Embodied prediction. In T. Metzinger & W. J. Windt. (Eds.). *Open Mind, 7* (pp.1-21). MIND Group. https://doi.org/10.15502/9783958570115

Clark, A. (2016). *Surfing Uncertainty*. Oxford University Press.

Clark, A. (2017). How to knit your own Markov Blanket: Resisting the second law with metamorphic minds. In T. Metzinger, & W. Wiese (Eds.), *Philosophy and Predictive Processing, 3* (pp. 1-19). MIND Group. https://doi.org/10.15502/9783958573031

Clark, A., Friston, K., & Wilkinson, S. (2019). Bayesing qualia: Consciousness as inference, not raw datum. *Journal of Consciousness Studies, 26*(9-10), 19-33.

Clark, A., and J. Toribio. 1994. Doing without representing? *Synthese 101*(3), 401–431

Conant, R. C., & Ashby W. R. (1970). Every good regulator of a system must be a model of that system. *International Journal of Systemic Science, 1*(2), 89-97

Crane, T. (1998). Intentionality as the mark of the mental. *Royal Institute of Philosophy Supplements, 43*, 229-251.

Crane, T., & French, C. (2021). The problem of perception. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2021 ed.). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/fall2021/entries/perception-problem/

Cummins, R. (1989). *Meaning and mental representation*. The MIT Press.

Ćwiek, A., *et al.* (2022). The bouba/kiki effect is robust across cultures and writing systems. *Philosophical Transactions of the Royal Society B, 377*(1841), 20200390.

Das, N. (2021). Vācaspati on aboutness and decomposition 1. In C. Meyns (Ed.). *Information and the History of Philosophy* (pp. 79-106). Routledge.

Davidson, D. (2016). Knowing one's own mind. In A. Pessin & S. Goldberg (eds.). *The Twin Earth Chronicles* (pp. 323-341). Routledge.

Davies-Barton, T., Raja, V., Baggs, E., & Anderson, M. L. (2026). Debt-free intelligence: Ecological information in minds and machines. *Philosophical Psychology, 39*(2), 239-265.

De Haan, S., Rietveld, E., Stokhof, M., & Denys, D. (2013). The phenomenology of deep brain stimulation-induced changes in OCD: an enactive affordance-based model. *Frontiers in human neuroscience, 7*, 653. https://doi.org/10.3389/fnhum.2013.00653

Demeter, T. (2022). A mental fictionalism worthy of its name. In T. Demeter, T. Parent, & A. Toon (eds.). *Mental Fictionalism: Philosophical Explorations* (pp. 86-103). Routledge.

Dennett, D. C. (1987). *The Intentional Stance*. The MIT Press.

Dennett, D. C. (1988). Quining qualia. In A. J. Marcel & E. Bisiach (Eds.), *Consciousness in contemporary science* (pp. 42–77). Oxford University Press.

Dennett, D. C. (1991). Real patterns. *The Journal of Philosophy, 88*(1), 27-51.

Dennett, D. C. (1991b). *Consciousness Explained*. Little Brown

Dennet, D. C. (1992). The self as a center of narrative gravity. In F. Kessel, P. Cole, D. Johnson, M. Hakel (Eds.) *Self and Consciousness: Multiple Perspectives*. Erlbaum.

Dennett, D. C. (2000). Making Tools for Thinking. In D. Spreber (Ed.) *Metarepresentation: A Multidisciplinary Perspective* (17-29). Oxford University Press.

Dennet, D. C. (2005). *Sweet Dreams*. The MIT Press.

Dennett, D. C. (2013). Expecting ourselves to expect: The Bayesian brain as a projector. *Behavioral & Brain Sciences, 36*(3), 209-210.

Diggle, S. P., Gardner, A., West, S. A., & Griffin, A. S. (2007). Evolutionary theory of bacterial quorum sensing: when is a signal not a signal? *Philosophical Transactions of the Royal Society B: Biological Sciences, 362*(1483), 1241-1249.

Downey, A. (2016). Radical sensorimotor enactivism & predictive processing. In T. Metzinger, & W. Wiese (Eds.). *Philosophy and Predictive Processing*, 10 (pp.1-26). MIND Group. https://doi.org/10.15502/9783958573123

Downey, A. (2018). Predictive processing and the representation wars: A victory for the eliminativist (via fictionalism). *Synthese, 195*(12), 5115-5139.

Downey, A. (2024). Theatre Hunger: An Underestimated 'Scaling Up' Problem. *Erkenntnis, 90*(6), 2503-2521

Draaisma, D. (2000). *Metaphors of Memory: A History of Ideas about the Mind*. Cambridge University Press.

Drayson, Z. (2025). Representations are (still) theoretical posits. THEORIA. *An International Journal for Theory, History and Foundations of Science, 40*(3), 240–255. https://doi.org/10.1387/theoria.27209

Egan, F. (2020). A deflationary account of mental representations. In J. Smortchkova, K. Dolega, & T. Schlicht (Eds.), *What are Mental Representations?* (pp. 26–54). Oxford University Press.

Facchin, M. (2021a). Predictive processing and anti-representationalism. *Synthese, 199*(3), 11609-11642

Facchin, M. (2021b). Are generative models structural representations? *Minds and Machines, 31*(2), 277-303.

Facchin, M. (2022). *The metaphysics of Predictive Processing: A non-representational account*. Ph.D. (Doctoral dissertation). IUSS Pavia.

Facchin, M. (2023). Neural representations unobserved — or: A dilemma for the cognitive neuroscience revolution. *Synthese, 203*(1), 7.

Facchin, M. (2025). Is mental content an illusion? *Synthese, 206*(5), 1-29.

Facchin, M. (2026). Defusing the Representation-Hungry Challenge. *Review of Philosophy and Psychology*, 1-28. https://doi.org/10.1007/s13164-026-00822-2.

Facchin, M., & Leonetti, G. (2024). Extended animal cognition. *Synthese, 203*(5), 138.

Facchin, M., & Rucińska, Z. (2025). Public charades, or how the enactivist can tell apart pretense from non-pretense. *Erkenntnis, 90*(5), 2095-2117.

Facchin, M., Rucińska, Z., & Fondelli, T. (2025). Radically embodied introspection. *Topoi, 45*, 373-389.

Facchin, M., & Viola, M. (2025). Structure and function in the predictive brain. *Biology & Philosophy, 40*(6), 1-26.

Favela, L. H., & Machery, E. (2023). Investigating the concept of representation in the neural and psychological sciences. *Frontiers in Psychology, 14*, 1165622. https://doi.org/10.3389/fpsyg.2023.1165622

Favela, L. H., & Machery, E. (2025). The concept of representation in the brain sciences: The current status and ways forward. *Mind & Language, 40*(2), 215-225.

Finkelstein, D. H. (2008). *Expression and the Inner*. Harvard University Press.

Fischer, E., Allen, K., & Engelhardt, P. E. (2026). Scientific or naïve? Perceptions of direct and indirect realism, and why they matter. *Philosophy and Phenomenological Research, 112*(1), 102-129.

Fodor, J. (1984). Observation reconsidered. *Philosophy of science, 51*(1), 23-43.

Fodor, J. (1987). *Psychosemantics*. The MIT Press.

Fodor, J. (1990). *A theory of content and other essays*. The MIT Press.

Frankish, K. (2016). Illusionism as a theory of consciousness. *Journal of Consciousness Studies, 23*(11-12), 11-39.

Frankish, K. (2023). What is Illusionism? *Klēsis – Revue de philosophie, 55*. https://www.revue-klesis.org/pdf/klesis-55-04-keith-frankish-what-is-illusionism.pdf

Frege, G. 1918 [1997]. Thought. In M. Beaney (ed.), *The Frege Reader* (pp. 325-45). Blackwell.

Frankish, K. (2026). The Emancipation of the drone: A Dennettian thought experiment. *Philosophical Psychology, 39*(3), 910–940.

Friston, K. (2011). Embodied inference: Or "I think therefore I am, if I am what I think." In W. Tschacher & C. Bergomi (Eds). *Implications of Embodiment: Cognition and Communication*, (pp. 89–125). Imprint Academic.

Friston, K., Adams, R. A., Perrinet, L., & Breakspear, M. (2012). Perceptions as hypotheses: saccades as experiments. *Frontiers in psychology, 3*, 151. https://doi.org/10.3389/fpsyg.2012.00151

Frith, C. (2007). *Making Up the Mind*. Oxford University Press.

Gallagher, S. (2017). *Enactivist interventions.* Oxford University Press.

Gan, N. (2024). What Kind of Non-Realism is Fictionalism? *Ergo: An Open Access Journal of Philosophy, 15*. https://doi.org/10.3998/ergo.5720

Gazzaniga, M. S. (1995). Principles of human brain organization derived from split-brain studies. *Neuron, 14*(2), 217–228.

Gedenryd, H. (1998). *How designers work: Making sense of authentic cognitive activities*. (Doctoral dissertation). Lund University.

Gładziejewski, P. (2016). Predictive coding and representationalism. *Synthese, 193*(2), 559-582.

Gödert, H. W., Rill, H. G., & Vossel, G. (2001). Psychophysiological differentiation of deception: The effects of electrodermal lability and mode of responding on skin conductance and heart rate. *International Journal of Psychophysiology, 40*(1), 61-75.

Godfrey-Smith, P (2006). Mental representations, naturalism and teleosemantics. In G. MacDonald & D. Papineau (eds.) *Teleosemantics* (pp.42-68). Oxford University Press.

Gopnik, A. (1993). How we know our minds: The illusion of first-person knowledge of intentionality. *Behavioral and Brain sciences, 16*(1), 1-14.

Grice, H. P. (1989). *Studies in the Way of Words*. Harvard University Press.

Hall, M., & Robinson, D. (2021). Acoustic signalling in Orthoptera. *Advances in Insect Physiology 61*, 1-99.

Harris, Z. S. (1954). Distributional structure. *Word, 10*(2-3), 146-162.

Haruno, M., Wolpert, D. M., & Kawato, M. (2001). Mosaic model for sensorimotor learning and control. *Neural computation, 13*(10), 2201-2220.

Hashem, I., & Van Impe, J. F. (2022). Dishonest signaling in microbial conflicts. *Frontiers in microbiology, 13*, 812763. https://doi.org/10.3389/fmicb.2022.812763

Haugeland, J. (1990). The intentionality all-stars. *Philosophical perspectives, 4*, 383-427.

Haugeland, J. (1991). *Representational genera.* In *Having Thought* (pp.171-207). Harvard University Press.

Heemskerk, J. (2024). *How informational teleosemantics works: Towards a realist theory of content* (Doctoral dissertation). University of Warwick.

Heidegger, M. (1926/1962). *Being and Time*. Harper & Row.

Hohwy, J. (2013). *The predictive mind*. Oxford University Press.

Hohwy, J. (2017). How to entrain your evil demon. In T. Metzinger, & W. Wiese (Eds.), *Philosophy and Predictive Processing, 2* (pp. 1–15). MIND Group. https://doi.org/10.15502/9783958573048

Hurley, S. (1998). *Consciousness in Action*. Cambridge University Press.

Husserl, E. (1970). *The crisis of European sciences*. Northwestern University Press.

Hutto, D. D. (2012). *Folk psychological narratives: The sociocultural basis of understanding reasons*. The MIT Press.

Hutto, D. D. (2016). Narrative self-shaping: A modest proposal. *Phenomenology and the cognitive sciences, 15*(1), 21-41.

Hutto, D. (2022). Getting real about pretense: A radical enactivist proposal. *Phenomenology and the Cognitive Sciences, 21*(5), 1157-1175.

Hutto, D. D., & Myin, E. (2013). *Radicalizing Enactivism*. The MIT Press.

Hutto, D. D., & Satne, G. (2015). The natural origins of content. *Philosophia, 43*, 521-536.

Ittelson, W. H. (1952). *The Ames demonstrations in perception; a guide to their construction and use*. Princeton University Press.

Jacobson, A. J. (2013). *Keeping the world in mind: Mental representations and the sciences of the mind*. Palgrave MacMillan

James, W. (1911). The tigers in India. In *The meaning of Truth* (pp. 43-50). Longman Green & Co.

Kammerer, F. (2021). The illusion of conscious experience. *Synthese, 198*(1), 845-866.

Kammerer, F., & Frankish, K. (2023). What forms could introspective systems take? A research programme. *Journal of Consciousness Studies, 30*(9-10), 13-48.

Kandel, E. R., Schwartz, J. H., Jessell, T. M., Siegelbaum, S., Hudspeth, A. J., & Mack, S. (Eds.). (2012). *Principles of neural science* (Vol. 4, pp. 1227-1246). McGraw-hill Medical.

Kaplan, D. M. (2012). How to demarcate the boundaries of cognition. *Biology & Philosophy, 27*(4), 545-570.

Kaplan, H. S., & Zimmer, M. (2020). Brain-wide representations of ongoing behavior: A universal principle? *Current opinion in neurobiology, 64*, 60-69

Kiefer, A., & Hohwy, J. (2018). Content and misrepresentation in hierarchical generative models. *Synthese, 195*(6), 2387-2415.

Kirchhoff, M. D., & Robertson, I. (2018). Enactivism and predictive processing: A non-representational view. *Philosophical explorations, 21*(2), 264-281.

Kohár, M. (2023). *Neural machines: A defense of non-representationalism in cognitive neuroscience*. Springer.

Krakauer, J. W., & Ramsey, W. (2026). Mental representation without neural representation: Understanding the evidence. *Philosophy and the Mind Sciences, 7*(1). https://doi.org/10.33735/phimisci.2026.12204

Kriegel, U. (2011). *The sources of intentionality*. Oxford University Press

Kriegeskorte, N., & Douglas, P. K. (2019). Interpreting encoding and decoding models. *Current opinion in neurobiology, 55*, 167-179.

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

Langland-Hassan, P. (2022). Why pretense poses a problem for 4E cognition (and how to move forward). *Phenomenology and the Cognitive Sciences, 21*(5), 1003–1021

Lee, J. (2019). Structural representation and the two problems of content. *Mind & Language, 34*(5), 606-626.

Lee, J. (2021). Rise of the swamp creatures: Reflections on a mechanistic approach to content. *Philosophical Psychology, 34*(6), 805-828.

Leinweber, M., *et al.* (2017). A sensorimotor circuit in the mouse cortex for visual flow prediction. *Neuron, 95*(6), 1420–1432

Lewis, D. (1973). *Counterfactuals*. John Wiley & Sons.

Lloyd, J. E. (1965). Aggressive mimicry in Photuris: Firefly femmes fatales. *Science, 149*(3684), 653-654.

Lycan, W. G. (2022). A Rylean Mental Fictionalism. In T. Demeter, T. Parent, & A. Toon (eds.). *Mental Fictionalism: Philosophical Explorations* (pp. 70–85). Routledge

Macpherson F. (2013). The philosophy and psychology of hallucination: An introduction. In F. Macpherson, D. Platchias (eds.). *HALLUCINATION - philosophy and psychology*. (pp. 1-39). Oxford University Press.

Mandal, F. B. (2014). Nonverbal communication in humans. *Journal of Human Behavior in the Social Environment, 24*(4), 417–421.

Mandik, P. (2016). Meta-illusionism and qualia quietism. *Journal of Consciousness Studies, 23*(11-12), 140-148

Martínez, M. (2025). The information-processing perspective on representation. *Philosophy and the Mind Sciences, 6*. https://doi.org/10.33735/phimisci.2025.11321

McGeer, T. (1990). Passive dynamic walking. Int. J. Robotics Res., 9(2), 62-82.

Meijer, E. H., Selle, N. K., Elber, L., & Ben-Shakhar, G. (2014). Memory detection with the Concealed Information test: A meta analysis of skin conductance, respiration, heart rate, and P300 data. *Psychophysiology, 51*(9), 879-904.

Meinong, A. (1960), The theory of objects, in R. Chisholm (ed.). *Realism and the Background of Phenomenology*. Free Press.

Mendelovici, A. (2018). *The Phenomenal Basis of Intentionality*. Oxford University Press.

Mendelovici, A. (2026). Belief as self-verifying fictions. In J. Lewis-Jong & E. Schwitzgebel (eds.) *The Nature of Belief* (pp. 40-66). Oxford University Press.

Mendelovici, A., & Bourget, D. (2023). Facing up to the problem of intentionality. *Philosophical Perspectives, 37*(1), 228-247.

Merleau-Ponty, M. (1963). *The structure of behavior*. Beacon Press.

Mesquita, B. (2022). *Between Us: How Cultures Create Emotions*. WW Norton & Company.

Mesulam, M. (2008). Representation, inference, and transcendent encoding in neurocognitive networks of the human brain. *Annals of neurology, 64*(4), 367-378.

Millikan, R. G. (1984). *Language, Thought and other biological categories*. The MIT Press.

Millikan, R. G. (1993). *White queen psychology and other essays for Alice*. The MIT Press.

Moran, R. (2001). *Authority and estrangement: An essay on self-knowledge*. Princeton University Press.

Myin, E., & Van den Herik, J. C. (2021). A twofold tale of one mind: revisiting REC's multi-storey story. *Synthese, 198*(12), 12175-12193.

Nanay, B. (2023). *Mental imagery: Philosophy, psychology, neuroscience*. Oxford University Press.

Nanay, B. (2025). Mental imagery, predictive processing, aphantasia, and the interaction between philosophy and cognitive science: Responses to Amy Kind, Christian Scholz, and Neil Van Leeuwen. *Mind & Language, 40*(3), 333-340.

Neander, K. (2017). *A Mark of the Mental*. The MIT Press.

Nes, A. (2008). Are only mental phenomena intentional? *Analysis 68*, 205–15

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*(3), 231-259.

Nizami, L. (2019). Information theory is abused in neuroscience. *Cybernetics & Human Knowing, 26*(4), 47-97.

Noe, A. (2004). *Action in Perception*. The MIT Press.

Nolfi, S. (2002). Power and the limits of reactive agents. *Neurocomputing, 42*(1-4), 119-145.

Olah, C., Satyanarayan, A., Johnson, I., Carter, S., Schubert, L., Ye, K., & Mordvintsev, A. (2018). *The building blocks of interpretability. Distill, 3*(3), e10. https://doi.org/10.23915/distill.00010

Orlandi, N. (2014). *The Innocent Eye*. Oxford University Press.

Orlandi, N. (2016). Bayesian perception is ecological perception. *Philosophical Topics, 44*(2), 327-352.

Orlandi, N. (2018). Predictive perceptual systems. *Synthese, 195*(6), 2367-2386.

O'Regan, J. K. (1992). Solving the "real" mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology/Revue canadienne de psychologie, 46*(3), 461-488.

O'Regan, J. K. (2011). *Why red doesn't sound like a bell: Understanding the feel of consciousness*. Oxford University Press.

Pezzulo, G., & Nolfi, S. (2019). Making the environment an informative place: A conceptual analysis of epistemic policies and sensorimotor coordination. *Entropy, 21*(4), 350. https://doi.org/10.3390/e21040350

Piccinini, G. (2022). Situated neural representations: Solving the problems of content. *Frontiers in Neurorobotics, 16*, 846979. https://doi.org/10.3389/fnbot.2022.846979

Piccinini, G. (2025). Computation in physical systems. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Fall 2025 ed.). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/fall2025/entries/computation-physicalsystems/

Piccinini, G., & Craver, C. (2011). Integrating psychology and neuroscience: Functional analyses as mechanism sketches. *Synthese, 183*(3), 283-311.

Plato (1997). *Cratylus*. Hackett.

Powers, W. T. (1973). *Behavior: The control of perception*. Aldine Publishing.co.

Raimondi, A. (2021). Crane and the mark of the mental. *Analysis, 81*(4), 683-693.

Raja, V. (2024). The motifs of radical embodied neuroscience. *European Journal of Neuroscience*, 60(5), 4738-4755.

Ramsey W. (2007). *Representation Reconsidered*. Harvard University Press.

Ramsey, W., Stich, S. P., & Garon, J. (2013). Connectionism, elimiantivism, and the Future of Folk Psychology. In W. Ramsey, D. Rumehalrt & S. Stich (Eds.) *Philosophy and connectionist theory* (pp. 199-228). Routledge.

Ramsey, W. (2020). Defending representational realism. In J. Smortchkova, K. Dolega, T. Schlicht (eds). *What are mental representations?* (pp. 54-78). Oxford University Press.

Ramsey, W. (2025). The hard problem of content is neither. *Review of philosophy and psychology, 16*(1), 153-174.

Ramstead, M. J., Kirchhoff, M. D., & Friston, K. J. (2020). A tale of two densities: Active inference is enactive inference. Adaptive behavior, 28(4), 225-239.

Revonsuo, A. (2000). The reinterpretation of dreams: An evolutionary hypothesis of the function of dreaming. *Behavioral and brain sciences, 23*(6), 877-901.

Robertson, I., & Kirchoff, M. D. (2019). Anticipatory action: Active inference in embodied cognitive activity. *Journal of Consciousness Studies, 27*(3-4), 38-68.

Rosenberg, A. (2019). *How history gets things wrong: The neuroscience of our addiction to stories*. The MIT Press.

Sartorius, N., Jablensky, A., Korten, A., Ernberg, G., Anker, M., Cooper, J. E., *et al.* (1986). Early manifestations and first contact incidence of schizophrenia in different cultures. *Psychological Medicine, 16*, 909-928

Sartre, J. P. (1971). *Sketch for a Theory of the Emotions*. Routledge.

Schnider, A. (2003). Spontaneous confabulation and the adaptation of thought to ongoing reality. *Nature Reviews Neuroscience, 4*(8), 662-671.

Schwitzgebel, E. (2002) A phenomenal, dispositional account of belief. *Noûs, 36*, 249-275.

Schwitzgebel, E. (2008). The unreliability of naive introspection. *Philosophical Review, 117*(2), 245-273.

Schwitzgebel, E. (2012). Introspection, what?, In D. Smithies & D. Stoljar (Eds.), *Introspection and Consciousness* (pp. 29-48). Oxford University Press

Schwitzgebel E. (2013), A Dispositional Approach to Attitudes: Thinking Outside of the Belief Box. In N. Nottelman (ed.), *New Essays on Belief: Constitution, Content, and Structure* (pp. 75–99). Palgrave Macmillan.

Schwitzgebel, E. (2016). Phenomenal consciousness, defined and defended as innocently as I can manage. *Journal of Consciousness Studies, 23*(11-12), 224-235.

Schwitzgebel, E. (2024). *The weirdness of the world*. Princeton University Press.

Schwitzgebel, E. (2026). Dispositionalism, yay! Representationalism, Boo!. Belief as self-verifying fictions. In J. Lewis-Jong & E. Schwitzgebel (eds.) *The Nature of Belief* (pp. 286-308). Oxford University Press.

Searle, J. (1992). *The rediscovery of the mind*. The MIT Press.

Searle, J. (1995). John Searle Replies. In Dennett, D. & Searle, J. 'The Mystery of Consciousness: an exchange'. *The New York Review of Books*. https://www.nybooks.com/articles/1995/12/21/the-mystery-of-consciousness-an-exchange/

Segundo-Ortin, M., & Hutto, D. D. (2021). Similarity-based cognition: Radical enactivism meets cognitive neuroscience. *Synthese, 198*(Suppl 1), 5-23.

Selfridge, O. G. (1958). Pandemonium: A paradigm for learning in mechanisation of thought processes. Proceedings of a Symposium Held at the National Physical Laboratory (pp. 513–526). *HMSO*. https://gwern.net/doc/ai/nn/1959-selfridge.pdf

Seth, A. K. (2014). A predictive processing theory of sensorimotor contingencies: Explaining the puzzle of perceptual presence and its absence in synesthesia. *Cognitive neuroscience, 5*(2), 97-118.

Seth, A. K. (2014b). The cybernetic Bayesian brain: From interoceptive inference to embodied selfhood. In T. Metzinger & J. M. Windt (Eds.), *Open MIND, 35* (pp. 1–24). *MIND* Group. https://doi.org/10.15502/9783958570108

Seth, A. (2021). *Being you: A new science of consciousness*. Penguin Publishing.

Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980). Vervet monkey alarm calls: Semantic communication in a free-ranging primate. *Animal Behaviour, 2*8(4), 1070-1094.

Sharot, T. (2011). *The Optimism Bias*. Penguin Publishing.

Shea, N. (2018). *Representation in cognitive science.* Oxford University Press.

Sganzerla, L., Hutto, D. D., & Kirchhoff, M. D. (2025). Beyond the extended mind: New arguments for extensive enactivism. *Synthese, 205*(3), 118.

Skyrms, B. (2010). Signals: *Evolution, learning, and information*. Oxford University Press.

Sprevak, M. (2013). Fictionalism about neural representations. *The Monist, 96*(4), 539-560.

Sprevak, M., & Irvine, L. (2020). Eliminativism about consciousness. In U. Kriegel (ed). *The Oxford handbook of the philosophy of consciousness* (pp. 348-370). Oxford University Press.

Strawson, G. (2004). *Mental Reality (II ed)*. Oxford University Press

Stich, S. P. (1983). *From folk psychology to cognitive science: The case against belief*. The MIT Press.

Sun, Z., & Firestone, C. (2020). The dark room problem. *Trends in cognitive sciences, 24*(5), 346-348.

Thobani, I. (2024). A triviality worry for the internal model principle. *Synthese, 204*(1), 36.

Thompson, E., & Cosmelli, D. (2011). Brain in a vat or body in a world? Brainbound versus enactive views of experience. *Philosophical topics, 39*(1) 163-180.

Thomson, E., & Piccinini, G. (2018). Neural representations observed. *Minds and Machines, 28*(1), 191-235.

Toon, A. (2023). *Mind as Metaphor*. Oxford University Press.

Usher, M. (2001). A statistical referential theory of content: Using information theory to account for misrepresentation. *Mind and Language, 16*(3), 311-334.

Vallée-Tourangeau, F., *et al.* (2016). Insight with hands and things. *Acta psychologica, 170*, 195-205.

Van Es, T. (2021). Living models or life modelled? On the use of models in the free energy principle. *Adaptive Behavior, 29*(3), 315-329.

Van Es, T., & Myin, E. (2020). Predictive processing and representation: How less can be more. In D. Mendoça, M. Curdao & S. Grouveia (Eds.). *The philosophy and science of predictive processing* (p. 7-25). Bloomsbury.

Van Leeuwen, N. (2025). Three questions on imagery and perception: A comment on Nanay's Mental imagery. *Mind & Language, 40*(3), 325-332.

Von Frisch, K. (1993). *The dance language and orientation of bee*s. Harvard University Press.

Vilarroya, O. (2017). Neural representation. A survey-based analysis of the notion. *Frontiers in psychology, 8*, 1458. https://doi.org/10.3389/fpsyg.2017.01458

Williams, D. (2018a). Predictive minds and small-scale models: Kenneth Craik's contribution to cognitive science. *Philosophical Explorations, 21*(2), 245-263.

Williams, D. (2018b). Predictive processing and the representation wars. *Minds and Machines, 28*(1), 141-172.

Williams, J. R. (2019). *The Metaphysics of Representations*. Oxford University Press.

Wittgenstein, L. (2001). *Tractatus logico philosophicus*. Routledge.

Wittgenstein, L. (2009). *Philosophical investigations*. Wiley & Sons.

Zahnoun, F. (2023). Some inaccuracies about accuracy conditions. *Phenomenology and the Cognitive Sciences, 22*(2), 461-477.

Zahnoun, F. (2025). Information in the brain: From metaphor to truth. *Theoria, 91*(5), e70021. https://doi.org/10.1111/theo.70021

Marco Facchin is a postdoctoral fellow at the University of Antwerp (FWO grant number 1202824N). He works in the philosophy of mind and cognitive science, with a focus on mental content, extended cognition, and "radically embodied" views of cognition.

Address: Department of Philosophy and Letters, Faculty of Philosophy, Centre for Philosophical Psychology, Universiteit Antwerpen, Building R, Rodestraat 14, Antwerpen 2000, Belgium. E-mail: Marco.facchin.marco.facchin@gmail.com / Marco.facchin@uantwerp.be – ORCID: 0000-0001-5753-9873