**Trick or Treat - A Response to Commentaries on “The Markov Blanket Trick”**

Vicente Raja1,2, Edward Baggs3,4, Anthony Chemero5,6, and Michael L. Anderson2,7,8

1Department of Philosophy, University of Murcia, Spain

2Rotman Institute of Philosophy, Western University, Canada

3Department of Language and Communication, University of Southern Denmark, Denmark

4Danish Institute for Advanced Study, University of Southern Denmark, Denmark

5Department of Philosophy, University of Cincinnati, USA

6Department of Psychology, University of Cincinnati, USA

7Department of Philosophy, Western University, Canada

8Brain and Mind Institute, Western University, Canada

We somewhat provocatively titled our target review article “The Markov Blanket *Trick*” (Raja et al., 2021) and then went on to elaborate why we thought it *was* a trick at least as deployed in much of the free energy principle (FEP) literature. But perhaps that was unfair. Perhaps it is instead a *treat*, a delicious boost of energy that can do real work. In the commentaries on the target review article, we find support for both attitudes.

Those who are largely content with framing the FEP and active inference deployment of Markov blankets as a trick are, along with us, **Di Paolo** and **Aguilera**, and, maybe to a lesser extent, **van Es & Hipólito** and **Palacios**. Those who think of Markov blankets as a treat include **Albarracin & Pitilya**, **Friston**, **Parr**, and **Ramstead**. For these latter authors, this sort of modeling has the power to make otherwise informal ways of understanding cognition more formal and therefore more useful.

In the following sections we will tackle the main topics that appear in the commentaries, with the above framing in mind. Much of our discussion will consist of clarifications, but we will also expand and extend our views in light of the topics raised by the commentators. In particular, we will further establish why, and in what ways, we think the Markov blanket trick *is* a trick, and doesn’t do what the treat proponents hope.

**1. Preliminaries: Evidence and Perception**

First of all, we want to comment on what counts as evidence and, concretely, on whether simulations are evidence in favor of FEP and active inference. The reason is that two of the commentaries (**Albarracin & Pitilya**, 2022; **Friston**, 2022) seem to think that we are somehow opposed to simulations. Nothing could be further from the truth. A look at some of our own scientific research shows that we are decidedly pro-simulation (e.g., Nalepka et al., 2019, 2021). Indeed, cognitive science as a discipline, when it has worked well, has been a back and forth between computational simulations and laboratory research on human behavior. We also acknowledge that simulations of the FEP and active inference have been valuable. Our favorite example is work by Jun Tani and his collaborators (e.g., Tani, 2016; Wirkuttis and Tani, 2021) that uses FEP and active inference, among other techniques, to control robots in their interactions with humans. This sort of work is valuable as a proof of concept, showing that it is possible that a cognitive architecture inspired by FEP and active inference can control human-like behavior. Our claims about empirical evidence for the FEP and active inference in the target article were in no way meant to be critical of important work like this. Our point was that, after two decades of promises, there is not much experimental evidence for FEP and active inference beyond simulations like these. This is especially concerning given the accelerating increase of literature production in the field. The field continues to expand rapidly (as **Friston** points out), but the non-simulation experimental evidence for the framework remains scarce. Our claim that experimental evidence for FEP and active inference is thin on the ground is hardly a radical critique. Proponents of the research program agree. In a recent review, Walsh et al. (2020) identify four empirical hypotheses derived from predictive processing, which is often proposed as a process theory for FEP and active inference, and find little experimental evidence for any of them. Simulations are a crucial part of the scientific endeavor, but they are not a substitute for experimental evidence from actual living things.

On a different but related note, we claim “perception remains unexplained” by FEP and active inference in the target article. Several commentaries (**Albarracin & Pitilya**; **Fiston**; **Parr**, 2022; **Ramstead**, 2022) disagreed with this claim. Our intention was to point out that the core issue of perception—e.g., how to go from an ambiguous, quasi-arbitrary, and generally insufficient stimulus to a concrete, stable, meaningful perception (see Raja, 2022)—remains unaddressed by FEP and active inference. This is, in part, because FEP and active inference make assumptions about how perception works; these assumptions are not an explanation of perception (more of this in 4). This is related to the imbalance between simulations and experimental work described above. Simulations allow for perception to be idealized. The examples we provide in the target article show this: simulations of the mouse in the T-maze assume just a few concrete, stable, meaningful perceptual states, so they don’t need to deal with the ambiguous, quasi-arbitrary, and generally insufficient stimuli real organisms (are supposed to) deal with. For FEP and active inference to explain the mouse’s visual perception, what is required is a story about how the dynamic stimulation of the receptor cells of the moving mouse (i.e., its optic flow) enable it to perceive a stable world. Currently, FEP and active inference do not provide such a story. Perhaps proponents of these views will provide this sort of detailed story in the future. Until they do, we stand by our claim that FEP and active inference assume, but do not explain, perception.

**2. The Core of the Debate: What Is a Thing?**

Some of the core claims of the target review article concern the notion of *thing* used in the FEP literature and the problems we find in it—in the following, we’ll use “*thing*” to refer to FEP’s particular notion and “thing” to refer to the everyday sense of the word. Put simply, FEP defines *things* as weakly-mixed random dynamical systems with an attractor set and a Markov blanket. Once *things* are defined this way, their activities in terms of self-organization, homeostasis, and cognition, among others, can be described as a form of Bayesian inference. We take it to be an uncontroversial characterization of the framework and we think it is shared both by the members of the FEP community, like **Friston**, **Parr**, and **Ramstead**, and by the critics of FEP, like **Di Paolo** (2022) and **Aguilera** (2022). It is true, though, that very recent formulations of FEP seem to not require an attractor set as a condition for the principle to apply (Ramstead et al., 2022). We will briefly address them at the end of this section, although our main focus will be the formulation made explicit in the target review article, which still is the common one in the FEP literature and the one framing the commentaries as well.

First of all, we want to point out that the target review article concedes FEP can be–and maybe is–true if the conditions under which it is developed apply. Namely, the criticism and suggestions associated with the “Markov blanket trick” depend neither on the success of the framework once its deeper assumptions are accepted nor on the formal abilities of its developers. Our point is that, even granting FEP is impeccable in the formal sense, the use of Markov blankets to define *things* remains problematic insofar as both its application and scope remain unclear. The commentaries by Ezequiel **Di Paolo** and Miguel **Aguilera** provide further support to this claim. Concretely, **Di Paolo** offers a suggestive theoretical analysis of the problematic use of Markov blankets in the context of constitutive self-organization. One of the key aspects of systems that engage in such a form of self-organization (e.g., living organisms, cognitive systems) is the ambiguous relationship between what counts as their constitution and what counts as their interaction with the environment. In these kinds of systems, **Di Paolo** claims, “interaction and constitution, both ongoing and materially overlapping processes, are not only mutually related; they interpenetrate in dialectical fashion without wholly losing their difference” (p. 8). The use of Markov blankets presupposes an ontological order between these ideas: *things* are already constituted when they get to interact with their environment. Such ontological order is not adequate to account for constitutive self-organization. In a similar critical fashion, Miguel **Aguilera** complements and deepens the criticisms of the target review article by pointing out another problematic aspect of the use of Markov blankets to individuate some kinds of things. **Aguilera** stresses the asymmetries within the thing-environment relationship. He shows how those very asymmetries and higher-order dependencies between states of different systems threaten the existence of Markov blankets for a wide variety of complex systems.

The commentaries of both **Di Paolo** and **Aguilera** concur with the target review article’s claim that the use of Markov blankets in the FEP literature is problematic and, therefore, these commentators belong to the group of researchers that see such use as a trick. Researchers that see it as a treat, however, are not completely oblivious to the issue. For instance, Thomas **Parr** acknowledges the ambiguity in the identification of Markov blankets in some sorts of things, such as those related to “self-organized patterns in biology, of the sort that are often described by reaction-diffusion systems” (p. 17). He then goes on and proposes the preliminaries of a way to overcome such ambiguity. To do so, **Parr** suggests approaching these systems in a more coarse-grained way—e.g., in the case of candle flames, avoid focusing on the molecules and focus on the densities of molecules. As it is a preliminary proposal, we prefer not to provide a value judgment about its formal or empirical adequacy. It seems interesting and only further research will shed light on its implications. However, we want to point out that **Parr**’s proposal highlights, even if indirectly, one of the *leitmotivs* of the target review article: that some decisions regarding the ontology of the systems and the thing-ness of *things* fall beyond the scope of FEP. In the case of candle flames, for instance, FEP accommodates choosing both molecules and densities of molecules as the relevant states to identify the Markov blanket. Given this, the only reason to choose densities over molecules seems to be purely pragmatic and not dictated by the principle itself. That is, the selection of densities obeys to considerations about *things* other than FEP.

Other proponents of FEP do not see the relationship between Markov blankets and things as problematic as the target article does. This is the case of Maxwell **Ramstead** and Karl **Friston**. In his commentary, **Ramstead** takes the separability of *things* provided by the use of Markov blankets to be a truism in physics. We hesitate to accept that claim. We do agree that some degree of separation or individuation of systems and their environment is generally going on in physics. But, first, this is hardly specific to physics. Notions of separability or individuation of things are present in metaphysics since its inception (e.g., haecceity) and can also be traced in many other sciences. And second, separability or individuality is not a property exclusive of things. We can separate or individuate things, but also states, properties, relations, words, etc. Separating or individuating are activities we do, but not only with things. In this sense, separability is not a defining property of things nor is it specially connected to Markov blankets.

Another aspect of **Ramstead**’s commentary has to do with a recent mathematical proof of the high likelihood of existence of Markov blankets in all high-dimensional systems (Sakthivadivel, 2022)—more specifically, a mathematical proof of Fiston et al.’s sparse coupling conjecture (Friston et al., 2021). According to **Ramstead**, proving this conjecture undermines the claims that cast doubt on the use of Markov blankets in a wide variety of biological and cognitive systems. We are not convinced that’s the case though. First, it is true Sakthivadivel (2022) proves the high likelihood of the existence of *weak* Markov blankets in high-dimensional systems. But the fact that they are *weak* is important. Avoiding the technical details, weak Markov blankets are different from “classic” Markov blankets both in the particular and in the formal senses. As **Ramstead** puts it, and beyond the formal difference, “the construct of Markov blanket has been weakened to “soft blankets,” which are not entirely sealed off from the outside world” (p. 306). We see this an important change in the FEP literature as it allows for relatively porous blankets that do not seclude internal states from external ones in the way “classic blankets” did. It also changes the formalism to identify the blankets from the one used during the last decade or so. And second, granting the correctness of Sakthivadivel’s mathematical proof, we think **Ramstead** and colleagues end up with a concept of (weak) Markov blankets that is not as well-justified as they need it to be. Proving the existence of weak Markov blankets in high-dimensional systems does not entail they exist in the way the FEP literature describes them. In other words, that we can find *things* in any high-dimensional system does not entail those *things* are the things we need them to be. For instance, it is quite clear that brains are supposed to be *things* in the FEP literature and that the Markov blanket of brains is constituted by sensory receptors (sensory states) and motor effectors (active states). But how can we be sure that a brain is a *thing* and that its Markov blanket, weak or not, falls exactly on those states? Nothing in FEP nor in the mathematical proof of the existence of weak Markov blankets answers this question. That precisely the brain is a *thing* and precisely sensory receptors and motor effectors are its Markov blanket depend on other assumptions about the kind of system the brain is and its role within cognitive systems. And the same can be said of virtually all the other *things* and Markov blankets typically described in the FEP literature.

Finally, Karl **Friston** doubles down on the FEP notion of *thing*. He argues that FEP’s criteria to account for *thing*-ness is to take *things* to be, definitionally, bundles of states within a Markov blanket. And the interactions or couplings between *things* are taken to occur through those blankets. In this sense, FEP practitioners do not start with things in the colloquial sense—like chairs, trees, or candle flames—to then model them with Markov blankets. On the contrary, they just begin with *things* as defined within the framework to then discover whether this can be used to model things of interest. This approach is very unsatisfactory as soon as some content is added to it; that is, as soon as the question is about a particular *thing* or a particular thing. The first moment of dissatisfaction echoes the problem already pointed out in our discussion of the commentary by **Ramstead**: the identification of *things* with things (e.g., brains, cells, pendulums, trees, etc.) in the FEP literature is rampant, but little justification is provided on why, if the Markov blanket of the *thing* exists, it happens to be (or not) precisely the one needed to understand the thing of interest. And one of the points of the target review article is that, when the focus moves from the abstract “bundles of states” to concrete systems, such identification becomes very difficult to work out. To illustrate this point, the target review article uses the example of coupled pendulums. Under FEP, the blanket of coupled pendulums is a postulated shared beam, with the velocity and the position of each pendulum as sensory and active states, respectively (Kirchhoff et al., 2018). We then point out that this is a weird Markov blanket for coupled pendulums taken as ordinary things even if it is the right one when taken as *things*. It just doesn’t seem the same kind of blanket as those informally used in the FEP literature as canonical examples of Markov blankets: sensory receptors and motor effectors for brains or the cell membrane for cells, for instance. There is an ambiguity here. On the one hand, it is claimed that *things* are what they are by definition even if the boundaries do not really align with the intuitive ones for ordinary things. On the other hand, when FEP presents possible examples of *things*, like brains or cells, the Markov blanket always falls in the “right” place from the intuitive point of view. We do not think this ambiguity is unsolvable in principle, but we do think it follows from the other assumptions about thing-ness that, being independent from FEP, are thoroughly embodied in the FEP literature. FEP may prescribe the existence of a Markov blanket for every *thing*, but other assumptions held by FEP researchers make that blanket appear whenever and wherever is needed.

The way **Friston** deals with this criticism and with the example of the coupled pendulums is difficult to parse. Right after we point out the ambiguity just highlighted, he claims: “If one wants to understand or simulate how two things are coupled, then drawing a Markov blanket between two clocks is the way to go.” (p. 12). We assume he makes this statement in the concrete context of FEP, because otherwise it is plainly wrong. One can understand and simulate how two things are coupled without the need for Markov blankets at all. A chief example of this—already mentioned in the target review article—is the famous HKB model, which has successfully modeled coupled oscillators in general (including coupled pendulums) for the last 40 years without the need for Markov blankets, internal states, partitions, and the like—see Kelso (2021) for a retrospective on the HKB model. The burden of proof seems to be on the side of **Friston** and other proponents of FEP. They are the ones who need to justify the use of Markov blankets and other elements of the theoretical apparatus of the framework. We suggested in the target review article that trying to re-describe the HKB model in terms of FEP could help to understand what the benefits in endorsing the latter would be even if the re-description is not exempt of difficulties (see p. 60 of the target review article). Our suggestion remained unaddressed in the commentaries.

At the end of the day, what we found in the FEP literature when we wrote the target review article is reproduced in the commentaries of **Ramstead** and **Friston**: an ambiguous use of the notions of *thing* and thing. An indefinite back-and-forth between the formal and the informal. The relevant literature goes from *thing* to thing and back again without carefully distinguishing one from the other. **Friston** does the same in his commentary when, for instance, he argues that FEP is focused on cyclical interactions and not acyclical ones, as we defend. However, our use of “acyclical” is purely formal: as a matter of fact, when Markov blankets have been presented in the diagrams of the FEP literature, they have taken the form of the typical Markov blankets one can find in acyclical causal networks. Indeed, this fact permits the partitions needed for active inference to work: sensory states affect internal states but not the other way around (i.e., no cycles there), internal states affect active states but not the other way around (i.e., no cycles there either), etc. We think **Friston** agrees with this partition and, therefore, with the sense in which we use “acyclical”. Nevertheless, he still claims FEP is interested in “cyclical” interactions. Why? We think it is because he jumps from our formal use of “acyclical” to an informal use of “cyclical”, merging the different senses and muddying the waters. This ambiguity between the formal and the informal occurs regularly in the FEP literature. For instance, that “surprise” means surprisal (aka Shannon information) and “belief” means Bayesian belief is part of the formal aspects of FEP. However, too often these notions are used in their ordinary sense without caveats. The same has occurred with the notion of prediction (Anderson and Chemero, 2013) or more recently with the notion of sentience (Kagan et al., 2022), for instance. FEP proponents tend to attribute very concrete meanings to well-known concepts to then mix both their regular meaning and FEP meaning uncarefully, sometimes in the same article and even in the same paragraph. After reading some of the commentaries to the target review article, we think this issue remains at the core of the still open ambiguity between *things* and things.

**3.** **It Is Just a Model!**

One of the hot topics within the recent FEP literature has to do with its status as a scientific entity. Is FEP a physical principle or is it just a bit of mathematics? Should we understand FEP as a theory of everything, as the variational principle of least action, or as a modeling framework? Should we be realists or instrumentalists regarding FEP? These are open questions and a couple of commentaries delve into them. Concretely, the commentaries of Thomas **van Es** & Inês **Hipólito** (2022) and Patricia **Palacios** (2022) address the status of the FEP as a modeling framework.

We agree with **van Es & Hipólito** (as well as with **Palacios**) that FEP and its corollary, active inference, are just a modeling framework. We were motivated by similar concerns when we began writing the target review article. FEP was, and still is, often presented as a fundamental theory to solve all problems related to self-organization, life, and cognition. However, the more we learned about it, the more it looked like a trick to impose the data structure needed for Bayesian inference onto any system. Namely, the more we read about FEP, the more it seemed to us to be “just” a modeling framework among many others. And, like all the other modeling frameworks, we found it depended on a good number of modeling assumptions that were explicit neither in the formal nor the informal parts of the literature.

In this context, we cannot do anything other than agree with **van Es & Hipólito** on their instrumental interpretation of FEP for the same reasons they present in their commentary. That said, we want to point out that instrumentalism posits some issues for the usual FEP rhetoric. Classically, the move from using Shannon information to using variational free energy in the FEP literature is justified in terms of the system to which it applies. For instance, if an organism is said to minimize the variational free energy of its autonomous states with respect to its environment, variational free energy is used just because it is an upper bound on Shannon information, so by minimizing the former, the latter is also minimized. But why minimize variational free energy and not Shannon information directly? The usual reason one can find in the FEP literature is that variational free energy is available to the organism (given a generative model) while Shannon information is not. But this answer assumes a realist take on FEP: variational free energy is not seen as “just” an instrument researchers can use to model the behavior of organisms, but as a quantity organisms themselves have access to. Why use variational free energy instead of using Shannon information directly if one defends an instrumentalist take on the FEP though? The only reason would be that variational free energy is easier to calculate than Shannon information, but if that’s the case, it is never explicit in the FEP literature. This literature almost always assumes the realist position.

In her commentary, Patricia **Palacios** also regards FEP as a modeling framework. Our impression is that **Palacios** generally agrees with some of the main criticisms presented in the target review article. She rightly points out, however, that we do not provide ways for FEP to escape from those criticisms. We are in full agreement. The target review article presents challenges to FEP that, we think, the proponents of FEP should address and solve if they find them relevant. It is our impression some of these challenges are indeed relevant as they are shared by other scholars and, moreover, some proponents of FEP have begun to take them seriously. However, we are not proponents of FEP and active inference, so we are not in the position to be the ones addressing our own criticisms of the framework. We adopt a critical stance toward FEP and active inference because we think that in current form they do not account for some of the scientific issues we are interested in—e.g., interactions, perception, affordances, etc. This is also why we use a different framework for our own research. In our opinion, researchers that endorse and believe in FEP must be the ones fixing it.

An old adage has it that while all models are wrong, some models are useful. We are happy to acknowledge that FEP modelers have successfully fulfilled the first part of this adage. We are less convinced that they have fulfilled the second part, particularly as a research program in neuroscience and psychology.

**4.** **One world or two?**

We direct our final comments toward **Friston**. Friston offers a spirited commentary, which we enjoyed reading. He opens by telling us that “philosophy is quintessentially adversarial” and he engages with our comments in correspondingly sporting fashion. We reply here in a similar spirit. The subtext of Friston’s commentary seems to be that philosophical disagreement is inessential, or it is merely decorative. The real work of a theory, he thinks, is done by its mathematics. As philosophers, we are bound to disagree with this. We think that, when it comes to trying to understand cognitive systems, philosophical considerations are unavoidable. Were this not the case, we would be out of a job.

Friston obliges us, in his commentary, by providing several helpful illustrative examples of the importance of philosophy. Friston makes a series of assertions that he takes to be uncontroversial, but that are in fact philosophically non-innocent. Consider the following. At one point, Friston instructs us to “remember the candle is in the philosopher’s head just in the same way as the consequences of her action. ... Put simply, in active (planning as) inference, affordances have to be recognised or inferred before plans can be realised through action.”

Friston understands affordance perception as an act of mental categorization. This should not surprise us. It follows from Friston’s general Helmholtzian worldview. According to Helmholtz, meaningful perception always requires an act of categorization. This is, of course, one of those instances in which assumptions about what perception is go beyond what is dictated by FEP.

The deeper problem, however, is that Friston here mischaracterizes what the concept of affordances is all about. The whole point of the affordance hypothesis is that meaning can be perceived directly, which is to say that meaning can be selected from the perceptual flux. Gibson’s hypothesis is that affordances are physical facts that project perceptually detectable structure (1979/2014). Gibson invented the word “affordance” precisely in order to avoid a puzzle that arises within the context of the classical theory of perception. The classical theory of perception posits that meaning exists in two places: meaning exists in the external world, where it gives rise to perceptual cues, and it exists again in some sense, in duplicate form, in the mental world. In Friston's words, in addition to existing as an object in the world, “the candle is in the philosopher’s head.”

Gibson never bought this story. Gibson’s entire career can be read as an attempt to formulate an alternative to this classical two-world theory of perception. The affordance hypothesis is an attempt to collapse the two-world hypothesis of meaningful perception into a one-world hypothesis. The affordance hypothesis says that meaning is detected directly. If meaning is detected directly, then it is possible to understand meaning as existing in only one place, namely in the relation between perceiver and world.

Friston's reading of the notion of affordances is, to be fair, the only reading available to him, given his Helmholtzian worldview. But it turns the affordance hypothesis into something less than what it is. It turns affordances from something relational and dynamic and continuous into something discrete, categorical, and *thing*-like. It turns affordances from direct percepts into indirectly inferred concepts. It turns them from a native construct within an ecological theory of perception into a retrofitted construct within the classical, Helmholtzian theory.

Friston wants FEP to be a totalizing theory for the cognitive sciences and neurosciences. He wants to be able to subsume all existing discussion of perception, action, and cognition under the all-consuming notion of the blanket. The Markov blanket trick, when applied to cognitive systems, turns out to be a very old trick indeed. It is the trick of presenting cognition as the creation of a mental replica of an external world that can only ever be inferred from imperfect cues. Some of us would prefer to leave that trick in the past.

Friston’s comments on affordances play a useful clarifying role. Friston does not think that FEP is compatible with direct perception. We agree with him on that. The question, for the rest of us in the cognitive sciences, is to decide whether it is worth investing our time in trying to keep up with the developments within FEP. Is FEP producing enough treats to make it worth our time, or are we being tricked?

**Funding**

V.R.’s research is funded by the Juan de la Cierva-Incorporación Research Program and the project PID2021-127294NA-I00 of the Spanish Ministry of Science. MA was supported by a Canada Research Chairs Program award, grant number SSHRC 950-231929. AC was supported by the Charles Phelps Taft Research Center at University of Cincinnati (USA).

**References**

Aguilera, M. (2022). The nonequilibrium boundaries of living systems: Comment on “The Markov blanket trick: On the scope of the free energy principle and active inference” by V. Raja et al. *Physics of Life Reviews*, 43, 23-25.

Albarracin, M., & Pitliya, R. J. (2022). Situated models and the modeler: A comment on “The Markov blanket trick: On the scope of the free energy principle and active inference” by Raja, Valluri, Baggs, Chemero and Anderson. *Physics of Life Reviews*, 43, 4-6.

Anderson, M., & Chemero, T. (2013). The problem with brain GUTs: Conflation of different senses of “prediction” threatens metaphysical disaster. *Behavioral and Brain Sciences,* *36*(3), 204-205. doi:10.1017/S0140525X1200221X

Di Paolo, E. A. (2022). Markov border crossings: Comment on “The Markov blanket trick: On the scope of the free energy principle and active inference” by Vicente Raja et al. *Physics of Life Reviews*, 43, 7-9.

Friston, K. (2022). The ultimate trick? Comment on: “The Markov blanket trick: On the scope of the free energy principle and active inference” by Raja et al. *Physics of Life Reviews*, 43, 10-16.

Friston, K., Heins, C., Ueltzhöffer, K., Da Costa, L., & Parr, T. (2021). Stochastic chaos and markov blankets. *Entropy*, 23(9), 1220.

Gibson, J. J. (1979/2014). *The ecological approach to visual perception: classic edition*. Psychology press.

Kagan, B. J., Kitchen, A. C., Tran, N. T., Habibollahi, F., Khajehnejad, M., Parker, B. J, ... & Friston, K. (2022). In vitro neurons learn and exhibit sentience when embodied in a simulated game-world. *Neuron*. <https://doi.org/10.1016/j.neuron.2022.09.001>

Kelso, J.A.S. (2021). The Haken–Kelso–Bunz (HKB) model: from matter to movement to mind. *Biological Cybernetics*, 115, 305–322. <https://doi.org/10.1007/s00422-021-00890-w>

Kirchhoff, M., Parr, T., Palacios, E., Friston, K., & Kiverstein, J. (2018). The Markov blankets of life: autonomy, active inference and the free energy principle. *Journal of the Royal Society: Interface*, 15, 20170792. <http://dx.doi.org/10.1098/rsif.2017.0792>

Nalepka, P., Lamb, M., Kallen, R. W., Shockley, K., Chemero, A., Saltzman, E., & Richardson, M. J. (2019). Human social motor solutions for human–machine interaction in dynamical task contexts. *Proceedings of the National Academy of Sciences*, *116*(4), 1437-1446.

Nalepka, P., Silva, P. L., Kallen, R. W., Shockley, K., Chemero, A., Saltzman, E., & Richardson, M. J. (2021). Task dynamics define the contextual emergence of human corralling behaviors. *PloS one*, *16*(11), e0260046.

Palacios, P. (2022). From a free energy principle to a free energy model?: Comment on “The Markov blanket trick: On the scope of the free energy principle and active inference” by Vicente Raja et al. *Physics of Life Reviews*, 43, 96-97.

Parr, T (2022). Patterns and particles: Comment on “The Markov blanket trick: On the scope of the free energy principle and active inference” by Raja et al. *Physics of Life Reviews*, 43, 17-19.

Raja, V. (2022). Embodiment and cognitive neuroscience: the forgotten tales. *Phenomenology and the Cognitive Sciences*, 21, 603-623. <https://doi.org/10.1007/s11097-020-09711-0>

Raja, V., Valluri, D., Baggs, E., Chemero, A., & Anderson, M. L. (2021). The Markov blanket trick: On the scope of the free energy principle and active inference. *Physics of Life Reviews*, 39, 49-72.

Ramstead, M. J. D. (2022). One person's modus ponens…: Comment on “The Markov blanket trick: On the scope of the free energy principle and active inference” by Raja and colleagues (2021). *Physics of Life Reviews*, 43, 305-307.

Ramstead, M. J., Sakthivadivel, D. A. R., Heins, C., Koudahl, M., Millidge, B., Da Costa, L., ... & Friston, K. J. (2022). On Bayesian mechanics: a physics of and by beliefs. *arXiv*, 2205.11543.

Sakthivadivel, D. A. R. (2022). Weak Markov blankets in high-dimensional, sparsely-coupled random dynamical systems. *arXiv*, 2207.07620.

Tani, J. (2016). *Exploring Robotic Minds: Actions, Symbols, and Consciousness as Self-Organizing Dynamic Phenomena*.New York: Oxford University Press.

van Es, T., & Hipólito, I. (2022). Co-constructing Markov blankets: Tricky solutions: Comment on “The Markov blanket trick: On the scope of the free energy principle and active inference” by Vicente Raja et al. *Physics of Life Reviews*, 43, 29-31.

Walsh, K. S., McGovern, D. P., Clark, A., & O’Connell, R. G. (2020). Evaluating the neurophysiological evidence for predictive processing as a model of perception. *Annals of the New York Academy of Science*. <https://doi.org/10.1111/nyas.14321>

Wirkuttis, N., & Tani, J. (2021). Leading or Following? Dyadic Robot Imitative Interaction Using the Active Inference Framework. *IEEE Robotics and Automation Letters,* 6(3) 6024-6031.