

There is no “inference within a model”

Commentary on:

Bruineberg, J., Dolega, K, Dewhurst, J. & Baltieri, M.(2021) The Emperor's new Markov blankets. Behavioral and Brain Sciences, <https://doi.org/10.1017/S0140525X21002351>

Abstract:

I argue that there is no viable development of the instrumentalist *Inference within a model* research program. I further argue that both Friston and Pearl blankets are not the right sort of tool to settle debates on philosophical internalism and externalism. For these reasons, the *Inference within a model* program is far less promising than the target article suggests.

Commentary Main Text

In this commentary, I want to focus on the *Inference within a model* research program, and briefly argue for two claims. If correct, these claims suggest that Bruineberg and co-authors present the inference within a model research program in a too bright light.

First claim: Bruineberg and co-authors mischaracterize the instrumentalist development of that research program. They say that it is viable, but uninteresting. I think the opposite is true: it is not a viable development, but, if viable, it would have been interesting.

Recall that, as the target article makes amply clear, (a) *Pearl* blankets only capture the patterns of (in)dependencies between variables in a model, which *need not* correspond to real boundaries in the world, and (b) *Pearl* blankets are model-dependent. Now, the instrumentalist inference within a model program uses *Pearl* blankets as a guide to find real boundaries in the world. Given (a) and (b), the success of such a program seems predicated on having *at least* a reliable rule of thumb to identify patterns of conditional independencies corresponding to real, worldly, systemic boundaries, as well as a reliable rule of thumb to identify the models that accurately capture the real structure of the modeled phenomena (as opposed to merely providing a parsimonious account of the data observed).

These rules of thumbs would be interesting epistemic tools in their own right, and their usage would allow us to learn a great deal about the world. Moreover, while possessing these rule of thumbs would not allow us to vindicate the most ambitious claims concerning *Friston* blankets in the literature on the free-energy principle (e.g. the claim that physical systems “possess” or “instantiate” *Friston* Blankets, see Friston 2013), possessing these rules of thumb *would* be sufficient to allow *Pearl* blankets to play the boundary-defining role *Friston* blankets are currently supposed to play in the philosophical literature on the free-energy principle (Hohwy 2016, 2017; Kirchhoff and Kiverstein 2021; Ramstead *et al.* 2019). This could (in principle) allow us to solve hotly discussed philosophical problems only by “doing the math” as many supporters of the Free-energy principle claim, which would be an interesting development.

Yet, the instrumentalist inference within a model program is not really viable, because it is subtly circular. Recall that a *Pearl* blanket is defined as the union of three sets: the sets of parents, co-parent and children *of a target variable*. This means that, in order to identify the

Pearl blanket of a variable (or set thereof), we must have already identified the target, “blanketed” variable. If this is correct, the identification of the target variable *logically precedes* the identification of its Pearl blanket. This means that, in order to identify the Pearl blanket of a real-world system, we must have already identified the variable(s) mapping over that system. Hence we *cannot* identify systems, and the various variables describing their behavior, *by identifying their Pearl blankets*, on the pain of circularity. So, although the usage of Pearl blanket to identify the boundaries of a system suggested by the instrumentalist inference within a model *would* have interesting consequences is viable, it does not really seem viable.

Second claim: Neither Friston nor Pearl blankets can be used to *satisfactorily* solve the disputes surrounding various forms of philosophical internalism and externalism.

Consider that Friston and/or Pearl blankets have been used to “identify in a principled manner” all the relevant factors constituting some phenomena of interest (see, for example, Hohwy 2016; Clark 2017, Kirchhoff and Kiverstein 2021). Consider further that in the relevant literature the presence of a Friston and/or Pearl blanket *defines* what counts as internal or external in the relevant sense. What is “surrounded” by the blanket counts as internal in the relevant sense, the rest counts as external (see Hohwy 2017: 6-7).

It is easy to see that the conjunction of these two ideas entails that all the factors constituting a phenomenon of interest count, by definition of “internal”, as internal. But this means that the conjunction of these two claims makes internalism true *by definition*. If this is a solution to the philosophical internalism/externalism debate, it is not a *satisfactory* solution.

For one thing, the truth value of some forms of internalism and externalism seems to depend on contingent matters of fact. For example, the debate concerning externalism/internalism about the vehicles of cognition and consciousness would surely be solved by the existence of non-biological props able to mimic cerebral processes sufficiently well (see Adams and Aizawa 2010, p. 78; Vold 2015). It is hard to see how such a debate, concerning at least in part matters of fact, could be solved by definition.

The same verdict holds for other debates concerning externalism and internalism. The truth value of externalism and internalism about knowledge and mental content, for example, depends on what content and knowledge are; that is, on their nature (cf. Egan 2009; Segal 2009; Bonjour 2005). But the nature of content and knowledge is not something that can be satisfactorily settled *by definition*. Surely no internalist or externalist should be allowed to win the day just by *defining* certain factors as internal or external!

Summarizing: in this commentary, I have tried to argue that the target article is wrong on the instrumentalist development of the inference within a model research program: it would be an exciting research program, if it were viable. But it is not viable. Furthermore, I have tried to argue that Friston and Pearl blankets are the wrong kind of tools when it comes to providing a satisfactory solution to the philosophical debates concerning externalism and internalism. If the arguments I have provided here are correct, the target article mischaracterizes the inference within a model research program: its chances of success are far slimmer than Bruineberg and co-authors suggest.

Conflict of interest statement Conflict of interest: none

Funding statement This research received no specific grant from any funding agency,

commercial or not-for-profit sector.

reference list

- Adams, F.; & Aizawa, K. (2010). Why the mind is still in the head. In P. Robbins, M. Aydede (Eds.), *The Cambridge Handbook of Situated Cognition* (pp. 78-96). New York: Cambridge University Press.
- Bonjour, L. (2005). Internalism and externalism. In P. K. Moser (Ed.), *The Oxford Handbook of Epistemology* (pp. 234-263). New York: Oxford University Press.
- Clark, A. (2017). How to knit your own Markov Blanket. In T. Metzinger and W. Wiese (Eds.) *Philosophy and Predictive Processing: 3*. Frankfurt am Main, The MIND Group. <https://doi.org/10.15502/9783958573031>
- Egan, F. (2009). Wide content. In A. Beckermann, B.P. McLaughlin, S. Walter (Eds.), *The Oxford Handbook of Philosophy of Mind* (pp. 406-425). New York: Oxford University Press.
- Friston, K. (2013). Life as we know it. *Journal of the Royal Society Interface*, 10(86), 20130475.
- Hohwy, J. (2016). The self-evidencing brain. *Noûs*, 50(2), 259-285.
- Hohwy, J. (2017). How to entrain your evil demon. In T. Metzinger and W. Wiese (Eds.) *Philosophy and Predictive Processing: 2*. Frankfurt am Main, The MIND Group. <https://doi.org/10.15502/9783958573048>
- Kirchoff, D. M., & Kiverstein, J. (2021). How to determine the boundaries of the mind: a Markov blanket proposal. *Synthese*, 198(5), 4791-4810.
- Ramstead, M. J. *et al.* (2019). Multiscale integration: beyond externalism and internalism. *Synthese*, 198(1), 41-70.
- Segal, G. (2009). Narrow content. In A. Beckermann, B.P. McLaughlin, S. Walter (Eds.), *The Oxford Handbook of Philosophy of Mind* (pp. 425-440). New York: Oxford University Press.
- Vold, K. (2015). The parity argument for extended consciousness. *Journal of Consciousness Studies*, 22(3-4), 16-33.