

## Letter

Curious Inferences:  
Reply to Sun and  
Firestone on the Dark  
Room ProblemAnil K. Seth,<sup>1,2,3,\*</sup> Beren Millidge,<sup>4</sup>  
Christopher L. Buckley,<sup>1</sup> and  
Alexander Tschantz<sup>1,2</sup>

Sun and Firestone [1] presented a challenge to predictive processing (PP) accounts of brain function by reviving the Dark Room problem – the idea that if agents are mandated to minimise prediction error, the best thing for them to do is to seek out highly predictable environments where nothing changes, and stay there. They argued that standard responses to this challenge have the potential to render the PP account untestable and explanatorily empty. We disagree.

One standard response is that Dark Room type environments are intrinsically surprising, given the homeostatic imperatives of living organisms. One might worry that this response solves nothing, since it merely redefines what counts as ‘surprising’ for an agent. The reply by Van de Cruys and colleagues relieves us of this worry by highlighting the principled role of ‘optimistic predictions’ in driving actions [2].

A second, and related, response is that increasing prediction error in the short term – for example, by leaving a Dark Room to engage in curious exploration – may help to reduce prediction error in the long run. Sun and Firestone argued that this response is also inadequate because ‘not all motivations that drive us from Dark Rooms reduce to instrumentally valuable exploration, even over the long-term’ [1]. To support their point, they noted that some distinctive – though rare – human behaviours, such as riding rollercoasters

and reading poetry, do not seem to deliver instrumental (goal-oriented) benefit, even over the long-term. Fortunately, this objection loses its force when the role of action is properly taken into consideration, as exemplified by active inference formulations of PP [3,4].

Action holds a special position in active inference. Unlike inferences about sensory states and internal states, (proprioceptive) inferences leading to actions can directly change the environment. From the perspective of the agent, sequences of actions thus change the future. Given that an agent is compelled to minimise long-term prediction error, it is therefore also mandated to reduce its uncertainty about the world, so that it can better minimise the discrepancy between expected and actual sensory data across temporally extended sequences of actions. Such an agent will therefore engage in epistemic, information-seeking actions – such as leaving a Dark Room – even though such actions may transiently increase short-term prediction error. In short, in order to minimise surprise in the future, an agent needs to be a curious, sensation-seeking agent in the present.

Why does this response not fall foul of Sun and Firestone’s critique that it simply redefines what is surprising? The reason is that minimisation of long-term prediction error can be formalised in a way that makes

specific, testable predictions. The formalism that makes this possible is the free energy principle (FEP.). The FEP generalises PP to propose that organisms minimise the free-energy, a tractable (i.e., measurable from the perspective of the agent) upper-bound on the long-term average of sensory entropy, which generalises the notion of prediction error.

When it comes to Dark Rooms, agents must minimise free energy over long temporal horizons – over the long term. Mathematically, this means that agents must minimize not free energy *per se*, but the expected free energy – a quantity that can be formalized in various ways (Box 1). Minimising expected free energy entails minimising a (negative) expected information gain term, which rewards sampling those novel environmental states that (are predicted to) induce a large divergence between prior and posterior beliefs. This is why long-term free-energy-minimising agents are intrinsically drawn towards novel experiences (and thus out of Dark Rooms) that reduce uncertainty about the world, even at the expense of temporarily higher prediction error. Importantly, the free-energy functional (a function of a function) intrinsically balances this trade-off between immediate and long-term free-energy minimisation.

Formalising the situation in this way leads to testable hypotheses. For example, by

## Box 1. Extending Free Energy into the Future

We have highlighted that when agents minimise free energy across long temporal horizons, this naturally induces information-seeking (Dark Room escaping) behaviour. What does minimising free energy into the future mean? The idea is to conceptualise inference as operating over sequences of observations, states, and actions extending into the future. This leads to an imperative to minimise expected free energy, which quantifies the total free energy of a sequence of observations and actions.

Mathematically, the expected free energy can be expressed in several ways (through different free-energy functionals) [9]; each of which can be decomposed into instrumental terms, which promote the immediate fulfilment of prior beliefs, and – crucially – epistemic terms, which promote exploration of novel environmental contingencies. Importantly these epistemic terms arise naturally out of the mathematical formalism, instead of being bolted on *ad hoc*, and they only arise when performing inference over temporally extended sequences. This is because the epistemic terms function as Bayes-optimal mediators of the trade-off between short and long-term free-energy minimisation. Moreover, the properties of different free-energy functionals can be distinguished and tested empirically, thus leading to testable predictions about the specific functionals that best describe agents’ behaviour when they escape from Dark Rooms.

performing variational inference in computational models that have parameters corresponding to beliefs about actions, one can make specific predictions about epistemic actions such as eye movements [5]. By incorporating learning, one can also make predictions about the biases that may accrue to an agent's beliefs about the world, as it attempts to minimise expected free energy [6]. By reconstruing goals and rewards as prior expectations, these models can also make fine-grained predictions about the dynamics of reinforcement learning [7].

Will this approach extend to explain rollercoaster riding and poetry reading? In the details, perhaps not. But this is not a failure, nor is it – as Sun and Firestone suggested – a concern over the explanatory reach of PP and the FEP. Here, it is

important to recognise that the FEP is a framework, not a testable hypothesis in and of itself. The lasting value of the FEP lies in the fecundity with which it generates virtuous circles of testable hypotheses – such as those deriving from models of expected free energy – and not with any specific attempt to prove or falsify it [8].

<sup>1</sup>Department of Informatics, University of Sussex, Brighton BN1 9QJ, UK

<sup>2</sup>Sackler Centre for Consciousness Science, University of Sussex, Brighton, BN1 9QJ, UK

<sup>3</sup>Azrieli Program on Brain, Mind, and Consciousness, Canadian Institute for Advanced Research (CIFAR), Toronto, Ontario M5G 1M1, Canada

<sup>4</sup>Department of Informatics, University of Edinburgh, Edinburgh EH8 9AB, UK

\*Correspondence:

[a.k.seth@sussex.ac.uk](mailto:a.k.seth@sussex.ac.uk) (A.K. Seth).

<https://doi.org/10.1016/j.tics.2020.05.011>

© 2020 Elsevier Ltd. All rights reserved.

#### References

1. Sun, Z. and Firestone, C. (2020) The Dark Room Problem. *Trends Cogn. Sci.* 24, 346–348
2. Van der Cruys, S. et al. (2020) Controlled optimism: Reply to Sun and Firestone on the Dark Room Problem. *Trends Cogn. Sci.* (in press)
3. Friston, K.J. et al. (2010) Action and behavior: a free-energy formulation. *Biol. Cybern.* 102, 227–260
4. Buckley, C.L. et al. (2017) The free energy principle for action and perception: A mathematical review. *J. Math. Psychol.* 81, 55–79
5. Mirza, M.B. et al. (2018) Human visual exploration reduces uncertainty about the sensed world. *PLoS One* 13, e0190429
6. Tschantz, A. et al. (2020) Learning action-oriented models through active inference. *PLoS Comput. Biol.* 16, e1007805
7. Tschantz, A. et al. (2020) Reinforcement learning through active inference. *arXiv*, 2002.12636. Published online February 28, 2020
8. Lakatos, I. (1978) *The Methodology of Scientific Research Programmes: Philosophical Papers*, Cambridge University Press
9. Milidge, B. et al. (2020) Whence the expected free energy? *arXiv*, 2004.08128. Published online April 17, 2020