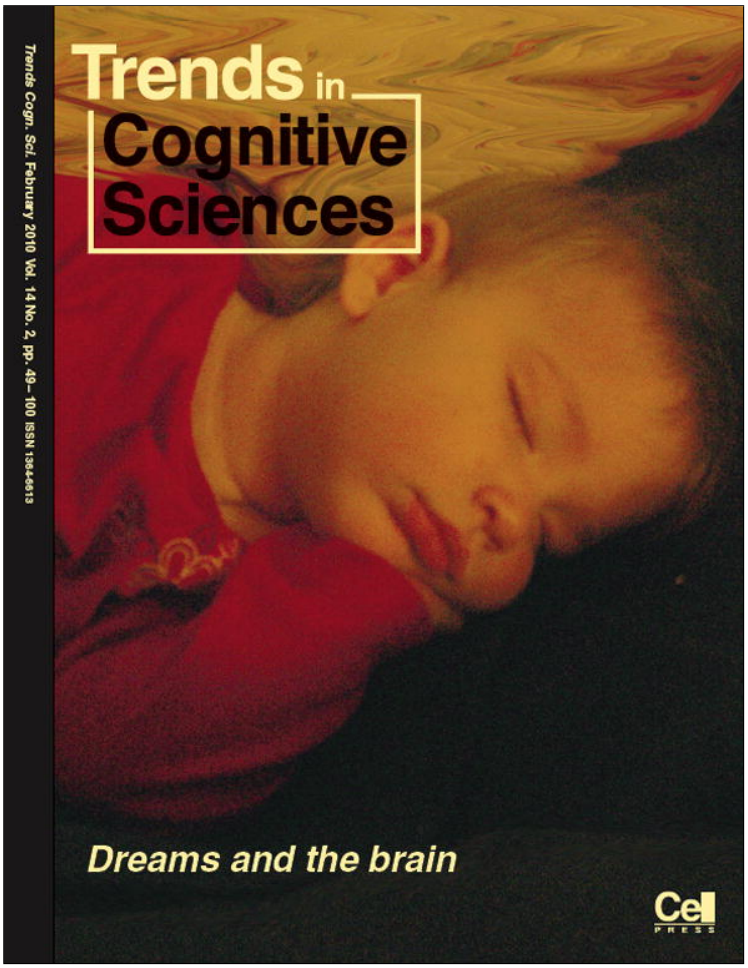
Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright

Author's personal copy

poses difficulties with regard to the specification of ϑ ranges, resolution of sensory ambiguity and inference of causal structure.

References

1 Friston, K. (2009) The free-energy principle: a rough guide to the brain? *Trends Cogn. Sci.* 13, 279–328
2 Goldstein, E. (1996) *Sensation and Perception*, (4th edn), Brooks/Cole Publishing Company
3 Friston, K. and Stephan, K. (2007) Free-energy and the brain. *Synthese* 159, 417–458
4 Friston, K. *et al.* (2009) Reinforcement learning or active inference. *PLoS One* 4, 1–13
5 Gatlin, L. (1972) *Information Theory and the Living System*, Columbia University Press
6 Campbell, J. (1984) *Grammatical Man*, Penguin

1364-6613/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tics.2009.11.010 Available online 23 December 2009

Letters Response

Some free-energy puzzles resolved: response to Thornton

Karl Friston

The Wellcome Trust Centre for Neuroimaging, University College London, Queen Square, London, WC1N 3BG, UK

Chris Thornton [1] poses some simple but key questions about the free-energy principle reviewed in [2]. These puzzles have simple and clear answers:

Puzzle: "A generative model of causal structure in the environment is [then] obtained, on which basis the agent is able to infer the 'causes of sensory samples' [ibid. p. 294]. What is unclear is how this mechanism would function where sensory samples are ambiguous" [1].

Answer: One of the main motivations for the free-energy principle is its appeal to [approximate] Bayesian inference where ambiguities are resolved by priors [3]. Priors are mandated by the (ill-posed) problems created by ambiguity and empirical priors are an integral part of hierarchical inference [2,Box 3]. This is not theoretical hand waving; in biophysics, the free-energy formulation is used routinely to solve difficult ill-posed inverse problems (e.g. [4]).

Puzzle: "On the face of it, no particular stand is taken on emergence of the structures that mediate minimization. But looking at the definition of free-energy, we find a significant role being played by the variable ϑ. It is values of this variable that encapsulate the brain's representation of 'environmental causes'" [1].

Answer: The representations are not environmental causes ϑ but the sufficient statistics μ of the brain's recognition density $q(\vartheta;\mu)$; these include synaptic activity and

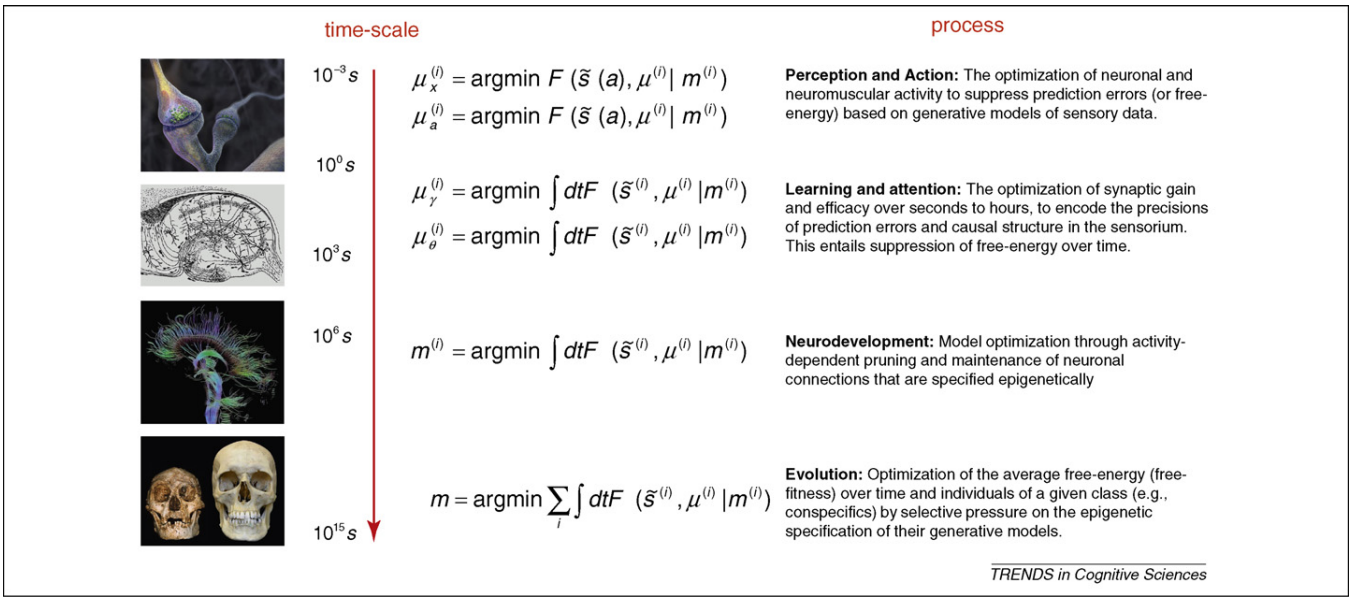


Figure 1. This schematic summarises the various timescales over which minimization of free-energy can be considered as optimizing the state (perception), configuration (action), connectivity (learning and attention), anatomy (neurodevelopment) and phenotype (evolution) of an agent. Here, $F(\tilde{s}, \mu^{(i)}|m^{(i)})$ is the free-energy of the sensory data (and its temporal derivatives - $\tilde{s}(a)$) and states of an agent $m^{(i)} \in m$ that belongs to class m. The states $\mu \supset \mu_x, \mu_\gamma, \mu_\theta$ correspond to synaptic activity, gain and strength, respectively, whereas a action determines the sampling of sensory data.

Corresponding author: Friston, K. (k.friston@fil.ion.ucl.ac.uk).

Author's personal copy

efficacy [2]. The implicit optimization of neuronal connections (i.e. perceptual learning) leads to hierarchical brain structures (models) that recapitulate causal structure in the sensorium. This optimization process can 'prune' the form or structure of the model (cf., synaptic pruning [5]) and is used routinely in model optimization (e.g. automatic relevance determination [6]). Furthermore, one could regard natural selection as optimizing the structural form of models at an evolutionary scale, through minimizing free-energy (where it is called free-fitness [7]). In a statistical setting, free-energy bounds on model evidence are used routinely in Bayesian model selection (where the log model evidence is negative surprise, e.g. [8];) (Figure 1).

Puzzle: "With the framework providing no principle for deciding the range of ϑ, the brain's representation of the conditional density is inevitably a 'slightly mysterious construct'" [1].

Answer: The range of ϑ (the values it can take) is specified by the form of the (generative) model and the priors it entails. For example, the equation in Box 2 [2] specifies the range of hidden states in the world $x^{(i)} \subset \vartheta$ with the range of a function, for example a neuronal activation function. The 'slightly mysterious' aspect of the recognition density is not its form (nor the implicit range of causes that are represented) but the fact that it is induced by the brain's physical states (which encode the recognition density).

Puzzle: "It is unclear how introduction of the 'free-energy' concept, specifically, adds explanatory content. . .*it is minimization of surprise that is explanatorily salient*" [1].

Answer: The explanatory advance furnished by free-energy is fundamental: it provides a means to minimize surprise. This is because surprise cannot be quantified by an agent, whereas free-energy can. Again, this is not abstract hand waving; the free-energy bound on surprise (or log-evidence for a model) plays an essential role in physics [9], machine learning [10] and statistics [11] for this reason.

References

1 Thornton, C. (2010) Some puzzles relating to the free energy principle: comment on Friston. *Trends Cogn. Sci.* 14, 53–54
2 Friston, K. (2009) The free-energy principle: a rough guide to the brain? *Trends Cogn. Sci.* 13, 293–301
3 Stone, J.V. *et al.* (2009) Where is the light? Bayesian perceptual priors for lighting direction. *Proc. Biol. Sci.* 276, 1797–1804
4 Phillips, C. *et al.* (2005) An empirical Bayesian solution to the source reconstruction problem in EEG. *Neuroimage* 24, 997–1011
5 Tessier, C.R. and Broadie, K. (2009) Activity-dependent modulation of neural circuit synaptic connectivity. *Front. Mol. Neurosci.* 2:8. doi:10.3389/neuro.02.008.2009
6 Tipping, M.E. (2001) Sparse Bayesian learning and the Relevance Vector Machine. *J. Mach. Learning Res.* 1, 211–244
7 Sella, G. and Hirsh, A.E. (2005) The application of statistical physics to evolutionary biology. *Proc. Natl. Acad. Sci. U S A* 102, 9541–9546
8 Daunizeau, J. *et al.* (2009) Variational Bayesian identification and prediction of stochastic nonlinear dynamic causal models. *Physica D.* 238, 2089–2118
9 Weissbach, F. *et al.* (2002) High-order variational perturbation theory for the free energy. *Phys. Rev. E*,66, 036129 DOI: 10.1103/PhysRevE.66.036129
10 Frey, B.J. and Jojic, N. (2005) A comparison of algorithms for inference and learning in probabilistic graphical models. *IEEE Trans. Pattern Anal. Mach. Intell.* 27, 1392–1416
11 Friston, K. *et al.* (2007) Variational free energy and the Laplace approximation. *NeuroImage* 34, 220–234

1364-6613/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tics.2009.11.008 Available online 22 December 2009

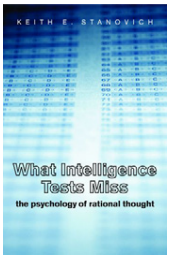
Book Review

Dysrationalia: intelligence without rationality

What Intelligence Tests Miss: The Psychology of Rational Thought by Keith Stanovich, Yale University Press, 2009. £16.00 (308 pp.) ISBN 978-0-300-12385-2.

David Over

Psychology Department, Durham University, Durham City DH1 3LE, UK



Despite a long tradition of research in both fields, the psychological study of intelligence and its tests has not been well integrated with the psychological study of rationality. Keith Stanovich's well-written and accessible book does integrate these studies and should, for that reason alone, have a highly beneficial impact on both.

Stanovich argues that standard intelligence tests miss the trait than can be of even greater value than relatively high intelligence: rationality. Although these tests measure something of value, it is not rationality: rationality is usually at best modestly correlated with scores on intelligence tests.

Corresponding author: Over, D. (david.over@durham.ac.uk).

Stanovich relies on the standard view in cognitive science that rationality should be defined in two related ways. It can refer, most fundamentally, to instrumental rationality, which is the ability to achieve one's goals or (more technically) maximize expected utility. It can also refer to epistemic rationality: the capacity to acquire the well-justified beliefs that are usually necessary for goal achievement. Stanovich establishes, both informally and on the basis of an extensive empirical literature, that there is a clear distinction between rationality so defined and intelligence. Intelligence can help to solve some problems about rational belief or action, but is of little help in other cases. An example is *myside bias*, the tendency to evaluate evidence from an egocentric point of view. Stanovich and his collaborator Richard West found no correlation between the magnitude of this bias and intelligence.