Title: Building a New Field of Computational Psychiatry

Article Type: Invited Commentary

Corresponding Author: Dr. Michael Moutoussis,

Corresponding Author's Institution: WTCN, University College London

First Author: Michael Moutoussis

Order of Authors: Michael Moutoussis; Eran Eldar, PhD; Raymond J Dolan, FRS

Building a New Field of Computational Psychiatry

Michael Moutoussis[1*], Eran Eldar[1] & Raymond J. Dolan[12]

[1] Wellcome Trust Centre for Neuroimaging, University College London, London WC1N 3BG, United Kingdom

[2] Max Planck – UCL Centre for Computational Psychiatry and Ageing Research, London WC1B 5EH, United Kingdom

[*] Correspondence: m.moutoussis@ucl.ac.uk

Computational psychiatry is a flourishing, yet relatively novel, branch of psychiatric research. It represents an integrative approach to psychopathology that draws on advances in systems neuroscience, decision-making based on probabilistic principles, as well as on advances in biological and computational neural network research and novel epidemiological approaches. At this stage in its evolution it has yielded few clinically useful insights. We argue that progress in this domain can be improved by more ecologically valid studies, including the use of more affectively relevant tasks as well as the acquisition of real-world (e.g. experience sampling) data. Furthermore, progress can be enhanced by specifically targeting relevant individual variability. The field might also benefit from a renewed focus on therapeutic recovery research, a domain that has been neglected relative to the attention accorded to etiology and diagnosis. Finally, we advocate a stronger engagement with patients, therapy professionals, prescribers and psychopharmacologists.

Brain function is realized in efficient performance of computations that best serve human needs. Computational psychiatry is a discipline which exploits the notion that understanding deviations from this ideal will help understand psychiatric disorders, with important implications for treatment. The field progresses in two directions simultaneously: top-down, (formally computational, i.e. starting from normative ideas), and bottom-up, starting from neural constraints (implementational). The field has enjoyed success largely because of its exploitation of new tools that connect these levels. An important

new bridge between top-down and bottom-up perspectives has been explicit modeling of the brain as a decision-maker operating in an uncertain world (probabilistic inference). This has helped elucidate existing findings, especially in biological reinforcement learning and allowed testing of hypotheses about information processing at the level of systems neuroscience, especially in the context of analysis of brain imaging data.

New vistas on psychopathology are now been opened up. There is a much-needed emphasis in rethinking psychiatric nosology in terms of computational mechanisms using constraints from all levels of analysis. A closer integration between animal and human research has led to an important focus on conserved computations and their implementation. A wide spectrum of probabilistic cognition, from sensory perception to social decision-making is being analyzed. At the level of whole human brain, there is a rich exploitation of technologies such as high-field fMRI and MEG, often employing sophisticated methods informed by a Bayesian framework (1). Complementary approaches are provided by biological neural network modeling, where computational psychiatry has enthusiastically embraced developments in machine learning. Neuroscientific insights stemming from progress in deep learning neural networks are also emerging (2) and here neuropsychiatric applications are awaited with excitement. At the epidemiological level, computational analyses are increasingly applied to populations, especially longitudinal informed data (3). Population studies are starting to use machine learning data-driven approaches to improve diagnosis.

Early computational psychiatry studies, often produced modest or difficult to interpret results, a situation that can be improved by addressing important limitations. First, the ecological validity of computational tasks in addressing psychopathology needs addressing. Tasks are needed that accurately assess psychiatrically relevant individual differences, aiming to elucidate features beyond neurotypical mechanisms, for instance by means of emotionally engaging experimental manipulations (4). Furthermore, computational psychiatry is likely to benefit by looking beyond behaviors that only extend to a timescale of minutes (typically during interactions with a computer in a laboratory). Here it is important to appreciate that many aspects of psychiatric disorders emerge from dynamical processes that evolve over timescales of weeks or months (5) and are tightly linked to the specific context of a person's life. The advent of mobile technologies (6) can provide an economically realistic means for studying behavior over such long time scales that necessarily embraces a participant's personal context.

Inspired by progress in other branches of medicine (e.g. stroke research), we expect that the implementation of psychiatrically relevant computations in the brain is heterogeneous. Individual brain differences most probably form the substrates for vulnerability, but also resilience and recovery (7). This highlights a need to fully characterize heterogeneity in subclinical participants, seeking out 'differences that make a difference' in generating symptoms. The importance of subclinical participants, particularly those experiencing substantial symptoms but who do not fulfill diagnostic criteria for a disorder, is that they are relatively free of confounds that arise from treatments and secondary deficits as seen in clinically diagnosed patients. Subclinical participants in effect can form an optimal testing ground for the development of tasks *sensitive to symptoms*. How (and in whom) mechanisms lead to *clinical disorders* can then be addressed. In short, we propose that the field move from a largely case-control approach to a stepped individual-variability approach.

Computational psychiatry may be like a gifted, confident, teenager who hasn't had the chance to come across real disappointments and discover the value of complex interdependence. Whether or not psychiatric ill health turns out to have 'natural joints' to 'carve', computational psychiatry can also flourish through cultivating its inter-dependencies with the broader mental health economy. The first inter-dependency is with patients and front-line staff, specifically addressing the role of learning as part of therapy. Learning-based treatments primarily concern rehabilitation and wellbeing, not necessarily symptom elimination (8). Psycho-education is essential, and here the twin lenses of the brain-as-updater-of-beliefs (the 'Bayesian brain') and the brain-as-seeker-of-well-being (the 'neuroeconomic brain') are notions easily accessible to patients and staff (Figure 1). More generally, as well as asking 'how can we treat the patient's symptoms, e.g. delusions?' a recovery approach asks 'how can the patient *best pursue their life goals* given residual symptoms?' This therapeutic target can be naturally pursued by computational psychiatry, as it deals with optimal decision-making. In turn this suggests opening up a new domain of research to *support adaptive* mechanisms, rather than *treat etiologic* mechanisms, by means of behavioral interventions and pharmacological agents that that have a potential to improve learning and decision-making.

[Figure 1 here]

The second inter-dependency is with prescribers and psychopharmacologists. Prescribers need optimal

treatment algorithms to use with existing drugs. They are less interested in how a relative safe drug, say Citalopram, functions than in whom it will *not* work (and hence likely to benefit from alternatives such as Venlafaxine, Quetiapine, or Lithium). We should therefore direct research at these neglected groups, namely those who are *not likely to improve* using current clinical algorithms.

Finally, the most lucrative and ethically most demanding inter-dependency is with drug developers. Here computational psychiatry could help enhance the power of animal research. As an example, β-CaMKII activity in the lateral habenula *enhances* spike output of lateral habenula neurons and induces depression-like behaviours, akin to anhedonia and behavioral despair (9). Insights can emerge by elucidating the computations through which neural structures process aversive information. This would help reconcile animal findings with seemingly conflicting data from patients, which has suggested depression is associated with *attenuated* habenula activity (10). Thus, computational psychiatry can help triangulate animal and human findings, ensuring that comparable - neurotypical or atypical - neurocomputational processes are considered to enable development of valid assays for novel therapeutic drugs.

In conclusion, computational psychiatry is beginning to flourish as a discipline but has made few clinically useful discoveries. The field can benefit from establishing a richer and more comprehensive industry of systematic task development, sensitive to relevant individual differences. We also suggest an extension of its focus in order to study a resilient-to-diseased spectrum across individuals ('nature'), exposures ('nurture') and development (longitudinally). Finally, computational psychiatry needs to engage with the priorities of key mental health stakeholders so as to develop treatment paradigms that do not solely focus on disease aetiology but embrace disorder management.

supported by the Biomedical Research Council. None of the authors report biomedical financial interests, nor potential conflicts of interest.

References

1. Stephan KE, Binder EB, Breakspear M, Dayan P, Johnstone EC, Meyer-Lindenberg A, et al. Charting the landscape of priority problems in psychiatry, part 2: pathogenesis and aetiology. The Lancet Psychiatry. Elsevier; 2016;3(1):84–90.

2. Mante V, Sussillo D, Shenoy KV, Newsome WT. Context-dependent computation by recurrent dynamics in prefrontal cortex. Nature. Nature Publishing Group; 2013;503(7474):78–84.

3. Castellanos-Ryan N, Struve M, Whelan R, Banaschewski T, Barker GJ, Bokde ALW, et al. Neural and cognitive correlates of the common and specific variance across externalizing problems in young adolescence. American journal of psychiatry. Am Psychiatric Assoc; 2014;

4. Eldar E, Niv Y. Interaction between emotional state and learning underlies mood instability. Nature communications. Nature Publishing Group; 2015;6.

5. Eldar E, Rutledge RB, Dolan RJ, Niv Y. Mood as representation of momentum. Trends in cognitive sciences. Elsevier; 2016;20(1):15–24.

6. Brown HR, Zeidman P, Smittenaar P, Adams RA, McNab F, Rutledge RB, et al. Crowdsourcing for cognitive science–the utility of smartphones. PloS one. Public Library of Science; 2014;9(7):e100662.

7. Seghier ML, Neufeld NH, Zeidman P, Leff AP, Mechelli A, Nagendran A, et al. Reading without the left ventral occipito-temporal cortex. Neuropsychologia. Elsevier; 2012;50(14):3621–3635.

8. Davidson L. Recovery, self management and the expert patient–changing the culture of mental health from a UK perspective. Journal of Mental Health. Taylor & Francis; 2005;14(1):25–35.

9. Li K, Zhou T, Liao L, Yang Z, Wong C, Henn F, et al. βCaMKII in lateral habenula mediates core symptoms of depression. Science. American Association for the Advancement of Science; 2013;341(6149):1016–1020.

10. Lawson RP, Nord CL, Seymour B, Thomas DL, Dayan P, Pilling S, et al. Disrupted habenula function in major depression. Molecular psychiatry. Nature Publishing Group; 2016;

Figure Legend

Figure 1. Slide from a psycho-educational presentation we administer to patients with medically unexplained symptoms, e.g. non-epileptic seizures (NES). It illustrates what the brain does in terms of Bayesian inference. This can help resolve 'psychological-neurological' conflicts. NESs, for example, can be seen as understandable but maladaptive decisions that a biologically sensitized brain has learnt to make, similar to other biological 'emergency reactions' found in animals (e.g. freezing, defensive paralysis) and humans (dissociative analgesia; cf. neurosymptoms.org/#/attack-treatment/4533198184). This framework deconstructs the common stigma 'it's all in your mind'. We then examine the context and stages involved in this maladaptive inference and help patients make more helpful inferences.

The brain tells us what it *thinks* is the *most reasonable* thing to say

The brain does not see the world as it is, nor does it decide what to do based on the outside world as it is. Instead, it combines

- physical sensations with ...
- powerful *expectations*

... based on what sense it's made of the world up to now.
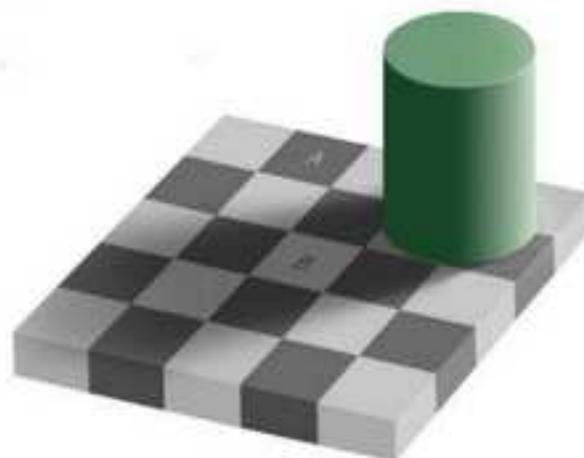
We can help the brain make the right decisions (is square A the same shade as B?) by guiding it along the right path.

http://www.123opticalillusions.com /pages/opticalillusions40.php