

Opportunities for emotion and mental health research in the resource-rationality framework

Evan M. Russek^{1,2,*}, Rani Moran^{1,2}, Daniel McNamee^{1,2}, Andrea Reiter^{1,2}, Yunzhe Liu^{1,2}, Raymond J. Dolan^{1,2}, Quentin J.M. Huys^{1,3,4}

1. Max Planck UCL Centre for Computational Psychiatry and Ageing Research, 2. The Wellcome Trust Centre for Neuroimaging, 3. Division of Psychiatry, University College London and 4. Camden and Islington NHS Foundation Trust

*For correspondence: e.russek@ucl.ac.uk

Abstract

We discuss opportunities in applying the resource-rationality framework towards answering questions in emotion and mental health research. These opportunities rely on characterization of individual differences in cognitive strategies; an endeavor that may be at odds with the normative approach outlined in the target article. We consider ways individual differences might enter the framework and the translational opportunities offered by each.

Main Text

The resource-rationality framework presented by Lieder and Griffiths has the potential to open up new computational approaches to emotion, including in the setting of mental illness. However, pitfalls can arise from a strong emphasis on normative modelling when this is at the expense of approaches that allow for measurement of individual differences. We consider the latter important in many translation efforts.

Minds face extremely complex decision-making tasks that far outstrip available computational resources. An individual must decide how to optimally apportion cognitive resources, such as attention, what memories to prioritize, what actions to evaluate and in what order, what future situations to simulate, whether to expend or save energy resources, and so on. The computational burdens imposed by resource-rationality, to choose optimal internal actions, is huge and often mandates the use of fast, automatic and approximate solutions.

Emotions and moods have long been characterized as states involving coordinated biases in all these domains (e.g. Scherer, 2009), suggesting that they could function as psychological and physiological mechanisms by which humans implement approximate resource-rational computations (Huys & Renz, 2017). For example, different emotions prioritize distinct action sets for mental evaluation (Frijda, Kuipers, & ter Schure, 1989); an effect which can be interpreted as implementing a solution to the meta-cognitive problem of the order in which a large space of actions should be evaluated. Moods involve longer-lasting biases towards particular emotions, and hence result in more persistent sets of cognitive choices to such internal demands. For example, low mood states are characterized by preferentially choosing to attend to negative stimuli, a prioritization of negative memories, a preferential evaluation of avoidance actions and not expending energy.

If it is true that moods are coherent, systematically covarying sets of internal cognitive choices that approximately address resource allocation problems, the question arises why a particular set of cognitive choices that define a mood state tend to co-occur. Furthermore, how these

internal policies change and adapt becomes fundamental to understanding them. The framework of resource-rationality seems ideally placed in answering these questions. By defining what state features lead a resource rational agent to make the cognitive selections defined by a given mood, we can understand why particular cognitive actions become associated with one another in distinct meta-action sets, as well as what pieces of environmental information trigger mood shifts, and hence why the range of internal actions becomes partitioned into a particular set of emotions.

Cognitive action selections are likely aberrant in many mental illnesses. Most obviously, if emotions are indeed cognitive action policies, this suggests that mood disorders can arise from maladaptive solutions to internal resource allocation problems. Indeed, patients with depression are known to suffer from characteristic biases in their allocation of attention, memory recall, action and energy expenditure (Elliott, Zahn, Deakin, & Anderson, 2011; Mathews & MacLeod, 2005; Whitton, Treadway, & Pizzagalli, 2015). Similarly, a tendency to catastrophize might relate to resource-rational arguments for biasing mental simulation toward extreme events (Lieder, Griffiths, & Hsu, 2018). However, this counter-intuitively links normative functions to maladaptive psychopathology, and begs the question about individual variability: not everybody should or does normatively suffer from mental illness.

Accommodating individual differences into the resource rationality framework, as it currently stands, is hindered by an emphasis on a modeling approach in which algorithmic hypothesis are developed by purely resource-rational considerations, and experiments that test group-level predictions of those hypothesis. This approach is at odds with characterizing individual differences. Indeed, Lieder and Griffiths explicitly argue against the alternative approach, common in computational cognitive neuroscience, of fitting components of models to human behaviour. This latter approach could provide parametric accounts of individual differences in terms of component processes.

Nevertheless, even the current framework provides for some inter-individual variability. First, though viewed as a constant in the target article, individuals certainly differ in their capacity for, and hence cost of, cognitive operations. For one, some costs will be sensitive to representation and hence depend on individual experience. Individual cost functions might be measured through behavioral tasks while assuming resource-rational optimality with respect to that individual cost function. Such cost estimates could possibly assist in the development of tailored cognitive interventions. As a concrete example, multiple lines of evidence link the representation of time-based costs to an operation of specific catecholaminergic neuromodulators (Constantino et al., 2017; Hauser, Moutoussis, Purg, Dayan, & Dolan, 2018). This implies that dysregulated cognitive processing caused by irregular time-based costing could be amenable to pharmacological modulation.

Second, the tradeoff between cost and utility must be optimized with respect to an individual's environment. Similar to how different moods might reflect adaptive cognitive actions for different environmental states, inter-individual differences in cognitive heuristics may reflect adaptation to different environments. Mental illness may reflect cognitive strategies that are adaptive in certain environments. Though difficult, it is tantalizing to consider the possibility of using the resource rationality framework to characterize the environments that an individual's cognitive strategy might be optimized for.

Such considerations can also motivate research into the meta-learning processes by which individuals arrive at resource-optimal cognitive strategies. Aspects of cognitive behavioral therapy aimed at "cognitive restructuring" attempt to change an individual's cognitive heuristics

so as to make them more adaptive (Beck, 1979). For example, in cognitive therapy patients are asked to identify a recent emotion triggering situation (e.g. “i wasn’t invited to a party”), record the beliefs they had in that situation (e.g. “everyone hates me”) and then critically evaluate this belief (e.g. “how would someone else think about this?”). Through the lens of resource-rational analysis, such approaches can be conceptualized as individuals learning to perform cognitive actions that lead to more adaptive inference regarding their current situations. More recent metacognitive treatments (Wells, 2011) could even be viewed as altering the very cost function by which internal actions are evaluated.

In conclusion, the notion of resource rationality raises tantalizing possibilities in the realm of mental illness, and we encourage an expanded methodological approach that embraces individual differences.

References

- Beck, A. T. (1979). *Cognitive therapy of depression*. Guilford Press.
- Constantino, S. M., Dalrymple, J., Gilbert, R. W., Varanese, S., Rocco, A. Di, & Daw, N. D. (2017). A Neural Mechanism for the Opportunity Cost of Time. *bioRxiv*, 173443. <http://doi.org/10.1101/173443>
- Elliott, R., Zahn, R., Deakin, J. F. W., & Anderson, I. M. (2011). Affective cognition and its disruption in mood disorders. *Neuropsychopharmacology*, 36(1), 153–82. <http://doi.org/10.1038/npp.2010.77>
- Frijda, N. H., Kuipers, P., & ter Schure, E. (1989). Relations among emotion, appraisal, and emotional action readiness. *Journal of Personality and Social Psychology*, 57(2), 212–228. <http://doi.org/10.1037/0022-3514.57.2.212>
- Hauser, T. U., Moutoussis, M., Purg, N., Dayan, P., & Dolan, R. J. (2018). Beta-Blocker Propranolol Modulates Decision Urgency During Sequential Information Gathering. *Journal of Neuroscience*, 38(32), 7170–7178. <http://doi.org/10.1523/JNEUROSCI.0192-18.2018>
- Huys, Q. J. M., & Renz, D. (2017). A Formal Valuation Framework for Emotions and Their Control. <http://doi.org/10.1016/j.biopsycho.2017.07.003>
- Lieder, F., Griffiths, T. L., & Hsu, M. (2018). Overrepresentation of extreme events in decision making reflects rational use of cognitive resources. *Psychological Review*, 125(1), 1–32. <http://doi.org/10.1037/rev0000074>
- Mathews, A., & MacLeod, C. (2005). Cognitive Vulnerability to Emotional Disorders. *Annual Review of Clinical Psychology*, 1(1), 167–195. <http://doi.org/10.1146/annurev.clinpsy.1.102803.143916>
- Scherer, K. R. (2009). The dynamic architecture of emotion: Evidence for the component process model. *Cognition & Emotion*, 23(7), 1307–1351. <http://doi.org/10.1080/02699930902928969>
- Wells, A. (2011). *Metacognitive therapy for anxiety and depression*. Guilford Press.
- Whitton, A. E., Treadway, M. T., & Pizzagalli, D. A. (2015). Reward processing dysfunction in

major depression, bipolar disorder and schizophrenia. *Current Opinion in Psychiatry*, 28(1), 7–12. <http://doi.org/10.1097/YCO.000000000000122>