Can a predictive processing framework improve the specification of negative bias in depression?

Harmer CJ, Browning M

Corresponding Author : Catherine Harmer

University Department of Psychiatry, Oxford Health NHS Foundation Trust,

Warneford Hospital

Oxford

Ox3 7JX

UK

E-mail: Catherine.harmer@psych.ox.ac.uk

Tel: +44 (0)1865 618326
Fax +44 (0)1865 251076

Our understanding of the psychological mechanisms underpinning depression was transformed by Aaron Beck's 1979 cognitive model (1). Beck conceptualised the role of negative self-beliefs or schemata in depression, driving negative biases in information processing, and thereby maintaining the diverse symptoms seen in the disorder. These ideas have remained influential to this day. This model formed the basis for the development of psychological treatments such as cognitive behavioural therapy as well as shaping subsequent research and work in this area. The negative distortions in information processing have also been identified as a target for the effects of antidepressant drug treatment in depression (2), providing an explanation of how these drugs work and why their effects on clinical state are delayed.

In this issue of Biological Psychiatry, Kube et al., (3) develop the ideas within Beck's original model to integrate recent advances in cognitive neuroscience using a predictive processing computational framework. The focus of the paper is on the observation that depressed patients generate more pessimistic predictions about the future which are relatively resistant to disconfirming positive information. In terms of the predictive processing framework they argue that this cognitive phenomenon arises due to an increased estimated precision of negative prior beliefs. An example of this process would be a depressed patient discounting positive feedback they receive from a work colleague because they are so sure they are performing poorly that the colleague must simply be mistaken (e.g. I think my work performance is 10%, they are saying it is 90%, so they must be mistaken). Building on this framework, the authors generate a set of specific predictions about behavioural, cognitive and neuroimaging measures in depression and speculate about potential novel treatment targets.

The essence of the predictive-processing framework (which has a marked overlap with accounts based on the "free-energy principle") reviewed by Kube et al (2020) is that our brain functions as a complex prediction machine based on an hierarchy of different levels of prediction (4). Throughout the hierarchy upper levels project "predictions" (what the higher levels expect the input to the lower levels to be) to the lower levels and the lower levels send back "prediction errors" (the difference between this prediction and what their input actually was). Sensory information is fed into the bottom of the hierarchy with the predictions produced by successively higher levels being increasingly general (see Figure 1). This kind of approach has been influential in other areas of neuroscience, particularly investigations of sensory and motor systems where it has provided a compelling account of the detailed hierarchical organisation of the cortex (5).

The hypotheses put forward by Kube et al., (3), arise from a specific attribute of this process—the predictions and prediction errors are not represented as simple values (e.g. my work performance is 10%) rather they include estimates of the uncertainty of these estimates (e.g. my work performance is somewhere between 5-15%). The degree to which a prediction error, generated at the lower level, is able to influence the upper level prediction

is governed by the relative uncertainty of the two representations, in other words feedback that your work performance is 90% will have a greater influence on someone whose beliefs have a higher uncertainty (e.g. they think their performance is between 5-75%) than someone whose beliefs have lower uncertainty (e.g. their performance is between 5-15%). In essence, they argue that depressed patients maintain negative beliefs about themselves with low uncertainty and positive beliefs with high uncertainty. This is consistent with evidence that the estimated uncertainty of positive and negative outcomes are maintained separately and therefore that positive and negative beliefs have the potential to be differentially affected by evidence (6). The perspective offered by Kube et al., (3) thereby provides a useful framework which integrates evidence on the role of negative biases in attention, interpretation and beliefs in depression and further specifies how this can be conceptualised and tested using a computational framework.

The idea that beliefs are updated by negative outcomes to a greater extent than positive outcomes in depression is compelling and has the potential to explain many of the symptoms and clinical observations in this disorder. However, to date the experimental evidence that depression is associated with higher learning rates for negative and/or lower learning rates for positive outcomes is inconsistent, with arguably more evidence that depressed patients differ in the effective value they attribute to positive relative to negative events rather than how they learn from these events (7). As the authors report there is also little consistency in research exploring neural responses during learning in depression. Lastly, there is pertinent data which suggests that learning from negative outcomes may also be impaired, rather than enhanced, in depression. For example, a recent study reported reduced responses to the threat of shock in depressed patients in the habenula (often conceptualised as the anti-reward centre of the brain) (8). Indeed, it is suggested that such a failure to learn from negative outcomes could reduce the ability to avoid negative situations in depression, also affording clinical face validity. This view would instead suggest an inflexibility of updating learnt information irrespective of valence. Consistent with this, we also found that a common treatment for depression, the SSRI citalopram, increased prediction errors for both positive and negative outcomes i.e. increasing the ability to learn from emotional information in general (9). As Kube et al point out, there may be variation between patients with depression in these core learning processes and a more precisely defined framework, as outlined here, may help us elucidate these potential differences between patients and presumably the search for more targeted treatments.

Lastly, the detailed model put forward by Kube et al. (3), represents one of a number of predictive processing accounts of psychiatric illness. While these models provide a fertile framework for reconceptualising illness experience, cognition and neurophysiology there are a number of outstanding areas where the research evidence supporting them needs to be further developed. For example, predictive processing has been successful in the visual cortex where there is a well defined hierarchy of regions. It is not at all clear where or how

the brain represents the higher order beliefs which are relevant to depression or how the uncertainty of these beliefs is encoded. If predictive processing models of psychiatric illness are to live up to their promise it is essential that these more general issues are addressed.

In summary, Kube et al. (3) have provided a fascinating suggestion about how we should think of the cognitive processes that underlie depression. Their work highlight the potential conceptual and therapeutic applications of this framework as well as the research which needs to be done to test and apply it to depression.

Financial Disclosures

CJH is supported by the Oxford Health NIHR Biomedical Centre. CJH has received consultancy fees from P1vital Ltd, Jansen Research, Sage Pharmaceuticals, Zogenics, Pfizer and Lundbeck. MB is supported by a Clinician Scientist Fellowship from the MRC (MR/N008103/1) and by the NIHR Oxford Health Biomedical Research Centre. He has received travel expenses from Lundbeck for attending conferences and acted as a consultant for Jansen Research and CHDR

References

1. Beck AT, Rush A, Shaw B, Emery G (1979): Cognitive Therapy of depression. New York: Guildford Press
2. Harmer CJ, Duman RS, Cowen PJ (2017) How do antidepressants work? New perspectives for refining future treatment approaches. Lancet Psychiatry 4:409-418.
3. Kube T, Schwarting R, Rozenkrantz L, Glombiewski JA, Rief W (2020) Distorted cognitive processes in major depression: a predictive Processing Perspective. Biological Psychiatry, this issue
4. Friston KJ, Kilner J, Harrison, L (2006) A free energy principle for the brain. J Physiol 100: 70–87.
5. Rao RP, Ballard DH (1999) Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. Nat Neurosci *2*:79–87.
6. Pulcu E, Browning M (2017) Affective bias as a rational response to the statistics of rewards and punishments. Elife 6: e27879.
7. Robinson OJ, Chase HW (2017) Learning and Choice in Mood Disorders: Searching for the Computational Parameters of Anhedonia. Computational Psychiatry 1:208–233.
8. Lawson RP, Nord CL, Seymour B, Thomas DL, Dayan P, Pilling S, Roiser JP (2017) Disrupted habenula function in major depression. Mol Psychiatry 22:202-208.
9. Scholl J, Kolling N, Nelissen N, Browning M, Rushworth MF, Harmer CJ (2017) Beyond negative valence: 2-week administration of a serotonergic antidepressant enhances both reward and effort learning signals. PLoS Biol 15:e2000756.

Figure 1: Schematic representation of the cortical hierarchy in predictive coding. Information enters from the sensory organs and is compared to predictions in prediction error units (red circles). Prediction errors are passed up the hierarchy (red curved arrows) to prediction units (blue boxes) which adjust their predictions and feed these back down the hierarchy (blue curved arrows). Sequential levels of the hierarchy are able to represent more complex and abstract predictions. Kube et al. (3) suggest that high level negatively valanced predictions dominate bottom up evidence in depression so that patients maintain negative beliefs and discount disconfirmatory evidence. Questions remain about how and where the complex beliefs, and their associated uncertainty, that are relevant to depression are represented in the brain.