PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B

Phil. Trans. R. Soc. B (2009) **364**, 1181–1182
doi:10.1098/rstb.2008.0321

Introduction

Predictions: a universal principle in the operation of the human brain

Just as physicists can explain complex systems with a small set of elegant equations (e.g. Maxwell's), it might be possible for the multidisciplinary study of the brain to produce a list of well-defined universal principles that can explain the majority of its operation. Given exciting developments in theory, empirical findings and computational studies, it seems that the generation of predictions might be one strong candidate for such a universal principle. *Predictions in the brain* is the focus of the collection of papers in this special theme issue. The word 'focus' might be misleading, however, because these papers range from addressing cellular underpinnings to computational principles to cognition, emotion and happiness, and they cover predictions that range from the next turn for a rat navigating a maze to predictions required in social interactions.

The questions that you might expect to learn and be stimulated about by reading this issue are accordingly diverse, for example:

— the links between future-related mental processes and memory, including their underlying cortical mechanisms. A related intriguing idea is that recollection relies on reconstruction, a mechanism that also provides a valuable tool for generating future-related thoughts, as well as for evaluations and future 'memories'. What are the neural mechanisms mediating such constructive processes?
— the connections between mental reconstruction, pattern completion, imagery, simulation, action plans, spatial navigation and more;
— the cellular mechanisms that balance the need to store stable memories and the need to be able to update them with novel experiences;
— counterfactual reasoning for past and future: 'what if' questions and mental scenarios in the context of preparation; 'what if' questions that seem to serve fantasizing; 'what if' questions that pertain to the past versus the future; and 'what if' questions that relate to self versus others. Are they all different? Perhaps not as much as it would seem, with all promoting simulations and imagery, possibly creating 'memories' that can be applied in future situations;
— is prediction-related processing primarily cortical or subcortical?
— what are the computational elements involved in the ability to predict?
— how do neural oscillations help in information organization for predictions?
— how do we learn what we need for subsequent generation of accurate predictions, and how does error in predictions promote such learning?
— what is the role of language in learning and sharing future-related information?
— does the ability for future simulations exist exclusively in humans, or are we simply unique in the magnitude and level of complexity of our foresight?
— in social settings, what is the relationship between simulations and our understanding of other people's thoughts and state of mind? and
— what is the role of predictions in emotion? How does the association of past experiences with affective values influence how we predict and perceive related information? Are we as good as we think in predicting emotional consequences? If there is a discrepancy between our ability to foresee affective states and rewards and our ability to predict non-affective outcome, what is the source of this discrepancy? How does the ability to plan and simulate futures affect our happiness, and what might be the implications to well-being and mental disorder?

We learn, encode, recollect, attend, recognize, evaluate, feel and act. The papers presented in this issue put forth neural models describing the possible interactions between these rich processes and the predictions-related mechanisms that connect them.

Some of the papers may appear almost synonymously describing similar ideas, which is an encouraging sign of potential validity. For example, the idea that predictions rely on memory. Indeed, it is hard to think of what other source if not memory can mediate predictions, but those proposals are in agreement with many of the details, including the cortical structures involved (e.g. medial temporal lobe and medial prefrontal cortex), the complexity of the underlying memory structures, the way they are encoded and the reconstructive way by which they are recalled and used in predictions.

On the other hand, some papers might seem to contradict each other. For example, the issue of whether foresight is exclusive to humans. There seems to be evidence either way, and the question is whether the fact that rats, for example, show future planning-like operations such as transitive inference

One contribution of 18 to a Theme Issue 'Predictions in the brain: using our past to prepare for the future'.

and prediction of upcoming positions implies that they are able to predict in the same way as humans, or similarly but to a lower capacity, or not at all, and these demonstrations in non-humans can instead be interpreted as something different from foresight. Other interesting debates include whether the underlying computational principles are Bayesian in nature or not. Such orthogonal proposals are particularly interesting because they show how the same solutions can be reached via different mechanisms and pathways.

As much as this collection answers important questions, it raises and emphasizes outstanding ones. How are experiences coded optimally to afford using them for predictions? What is the mechanism underlying reconstruction, and how do we construct a new simulation from separate memories? How specific in detail are future-related mental simulations, and when do they rely on imagery, concepts or language? What is the difference in mechanism and cortical underpinnings of predictions that stem from sequence memory (i.e. replaying existing memories) and predictions that stem from construction? What is the role of hierarchies in representations and predictions? It is hoped that the questions that emerge from this issue will inspire and steer future research on future-related mental processes.

Let me conclude this introduction with an intriguing analogy. The fighter plane F-16 is the first aeroplane intentionally designed to have an aerodynamically unstable platform. This design was chosen to enhance the aircraft's manoeuvrability. Most aeroplanes are designed to be stable such that they strive to return to their original attitude following an interruption. While such stability is a desired property for a passenger aeroplane, for example, it opposes a pilot's effort to change headings rapidly and thus can degrade manoeuvring performance required for a fighter jet. This behaviour has led to a saying among pilots that 'you do not fly an F-16, it flies you' (http://en.wikipedia.org/wiki/F-16_Fighting_Falcon). As is evident from the collection of articles presented in this issue, the brain might be similarly flexible and 'restless' by default. This restlessness does not reflect random activity that is there merely for the sake of remaining active, but, instead, it reflects the ongoing generation of predictions, which relies on memory and enhances our interaction with and adjustment to the demanding environment.

I am deeply grateful to Jasmine Boshyan for her relentless and remarkably meticulous assistance with preparing this entire issue for publication.

Moshe Bar*
Martinos Center, Harvard Medical School, Charlestown, MA 02129, USA
*bar@nmr.mgh.harvard.edu