We suggest, moreover, that while the problem of sensing the surroundings and other road users appears to be yielding impressive progress, the challenge of traffic ‘negotiation’ has scarcely been addressed either by the specialist literature on transportation research or by the cognitive sciences [13]. Indeed, the rate of progress on this challenge within the cognitive sciences may prove a decisive limiting factor in the development of autonomous vehicles.

Moreover, the safety of more limited steps to autonomy, where control is handed back and forth to a human driver, may depend on progress on understanding and modelling ‘negotiation’. The situations illustrated in Figure 1 arise routinely and unpredictably in urban driving, so such situations ought not to be classified as ‘too difficult’ and handed back to human users. In part, this is because a human will not be able to attend to, and resolve how to act in, such an interaction when previously engaged in some other task; but also because identifying the ‘difficult’ cases that require human-level negotiating skills may not be accurate without the deployment of such skills (just as it is difficult to accurately identify ‘difficult’ chess positions without actually attempting, and struggling, to decide what to do in such positions). Table 1 outlines some of the cognitive science challenges and possible pathways for the development of autonomous vehicles (leaving aside important ethical issues, questions of acceptability of even a small number of accidents and problems of the opacity of computer algorithms, which have been discussed elsewhere [14,15]): it includes one scenario in which the challenge of negotiation is addressed and three ways in which it might be skirted.

We believe that the challenge of autonomous vehicles, which promises great gains in human welfare through improved mobility, safety and environmental impacts, brings to light fundamental challenges for cognitive science and artificial intelligence, not just in sensing and control (where machines may potentially exceed human performance – e.g., in response times), but also in mimicking or seamlessly meshing with human behaviour in driving interactions. The problem of understanding how we ‘negotiate’ the traffic also provides a microcosm of deep questions concerning human social interaction and communication more generally.

Acknowledgments

This work was supported by RCUK/Jaguar Land Rover Grant EPSRC EP/N012380/1. The views expressed are solely those of the authors, not the sponsoring bodies. We thank two anonymous reviewers for their valuable input.

1Richardson, Z. B. (2006) Joint action: bodies and minds moving together. Trends Cogn. Sci. 10, 70–76
2Klein, G. et al. (2003) Common ground and coordination in joint activity. In Organizational Simulation (House, W.B. and Boff, K.R., eds), pp. 139–184, John Wiley & Sons
3Porto, F. et al. (2014) Drivers’ communicative interactions: on-road observations and modelling for integration in future automation systems. Ergonomics 57, 1795–1805
4Swan, L.A. and Owens, M.B. (1988) The social psychology of driving behavior: communicative aspects of joint-action. Mot Am. Rev. Social. 13, 59–67
5Schelling, T.C. (1983) The Strategy of Conflict, Harvard University Press
6Brown, B. and Lauter, E. (2017) The trouble with autolots: assisted and autonomous driving on the social road. Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, pp. 416–429, ACM
7Miyak, J. et al. (2016) Instantaneous conventions: the emergence of flexible communicative signals. Psychol. Sci. 27, 1555–1561
8Miyak, J.B. et al. (2014) Unwritten rules: virtual bargaining underpins social interaction, culture, and society. Trends Cogn. Sci. 18, 512–519
9Bacharach, M. (2006). In Beyond Individual Choice: Teams and Frames in Game Theory. In Beyond Individual Choice: Teams and Frames in Game Theory (Gold, N. and Sugden, R., eds), Princeton University Press

References

10. Shafto, P. et al. (2014) A rational account of pedagogical reasoning: teaching by, and learning from, examples. Cogn. Psychol. 71, 55–89
11. Goodman, N.D. and Frank, M.C. (2016) Pragmatic language interpretation as probabilistic inference. Trends Cogn. Sci. 20, 818–829
12. Pesutlo, G. et al. (2013) Human sensorimotor communication: a theory of signaling in online social interactions. PLoS One 8, e79876
13. Shashua, A. (2016) Sensing and Beyond: Towards Full Autonomous Driving. http://ir.mobilitye.com/investor-relations/events-and-presentations/ CES-2016-Presentation
14. Bonnfon, J.F. et al. (2016) The social dilemma of autonomous vehicles. Science 352, 1573–1576
15. Shariff, A. et al. (2017) Psychological roadblocks to the adoption of self-driving vehicles. Nat. Hum. Behav. 1, 694

Spotlight

Forecasting Faces in the Cortex

Comment on ‘High-Level Prediction Signals in a Low-Level Area of the Macaque Face-Processing Hierarchy’, by Schwidersk and Freiwald, Neuron (2017)

Lucy S. Petro1 and Lars Muckli1,*

Although theories of predictive coding in the brain abound, we lack key pieces of neuronal data to support these theories. Recently, Schwidersk and Freiwald found neurophysiological evidence for predictive codes throughout the face-processing hierarchy in macaque cortex. We highlight how these data enhance our knowledge of cortical information processing, and the impact of this more broadly.

Despite having no direct access to the world, neuronal networks must ultimately represent useful information about the external environment. Fortunately, our prior experiences are stored as information in the brain. If the brain can use this stored information to generate a prediction that matches the current environment, then the neuronal microcircuits has
represented something useful about the
world. This hypothetical model of brain
function is conceptualized by predictive
coding [1].

Predictive coding theories state that neu-
rons generate predictions higher up in the
cortical hierarchy, and test these predic-
tions against incoming sensory informa-
tion in lower areas, producing an error
signal if there is a mismatch between the
prediction and current sensory informa-
tion [1]. This theory, whereby an inter-

nal testing mechanism flows down the
cortical hierarchy, could support a range
of brain functions including social cogni-
tion. Consider recognizing the face of a
friend from a distance: your brain might predict that the face of the accompanying
person is that of his or her partner. Such
context-dependent expectations are essential for successful social behavior.

However, there was no prior neuronal
evidence that prediction error signals
generated in a lower area tested the pre-
diction of a higher area – a hallmark stip-
ulated by predictive coding. In a recent
paper, Schwiedrzik and Freiwald [2]
revealed how prediction-generating and
prediction-testing neurons in the macaque face network exchange codes until predictions match the input. The
neuronal guessing-game starts with a
template generated from memory that
neuronal processes then test against sen-
sory evidence lower in the cortical hier-
archy. The macaque face network is an
attractive system for scrutinizing predic-
tion principles. Face-selective cortical
patches form an interconnected hier-
archical network but have different and
specific functional preferences for faces.
Therefore, we have a hypothesis about the
prediction that each patch generates and
to which patch it will communicate
this prediction.

During a learning phase, the authors pre-
sented sequential face image pairs to
monkeys, training them to expect a
particular ‘successor’ image after seeing
the first ‘predictor’ image. Images differed
on dimensions to which face patches are
tuned (head orientation or identity). The
authors later manipulated the order of
images, inducing an error response in
neurons in the ‘ML’ (middle lateral) patch
because the new face pairs violated the
monkey’s expectation along dimensions
of identity, viewpoint, or both. By
comparing ML responses to the succes-
sor image under predictable and unpre-
dictable conditions, the authors showed
that prediction errors in ML reflect identity
specificity and view invariance.

Importantly, the prediction error prop-
erties are not those that ML is tuned to – but
are instead the properties of higher areas.
This suggests that ML constructs the
error from the comparison with top-down
predictions, a mechanism by which a
lower area inherits the features of higher
areas. This contrasts with previous
accounts of top-down feedback that con-
ceptualize the recovery of high-resolution
features by probing lower-hierarchy visual
areas [3]. The authors also confirm the
behavioral relevance of this neuronal pro-
cess. A separate, near-identical experi-
ment revealed that humans detected
faces faster in predictable versus unpred-
dictable conditions.

This study provides evidence of predic-
tive signaling in a brain system where
information-processing determinants
are well established. The data encour-
ge the re-evaluation of cortical processing
strategies. The currency which process-
ing might trade in is the violation of a
prediction, a signal that the cortex broad-
casts upstream to higher areas until
those areas generate a better hypothesis
that matches the input. Thus, cortical
communication might be well conceptu-
alized as a recurrent negotiation loop
attempting to resolve prediction error.
This model fits well for perception in
which higher-order representations are
well constrained and predict precise sen-
sory inputs; here we can estimate what
an error-detecting neuron might be
tuned to. However, we do not know
whether or how prediction sustains pro-
cessing for a sensory input that is not
immediately available but will arrive in
the future. This long-term mental fore-
casting, for example planning your route
home, involves complex and abstract
internal representations that unfold over
longer durations than sensory signaling.

Another challenge is to understand the
precise neuronal machinery for predictive
processing. The current data suggest that
cortical feedback carries predictions. Dis-
tal dendrites in upper layers of mamma-
lian cortex are a target of feedback and
could offer a means by which individual
neurons test and compare feedforward
and feedback signals [4], potentially
resolving the error. It is a challenge, how-
ever, to decode the language of feed-
back, in contrast to feedforward
responses that are driven directly by the
stimulus. In as much as feedforward
processing generalizes across viewpoints
(e.g., when identifying a face), feedback
projections might inherit viewpoint invari-
ance through a quick succession of pre-
dictions of mirror-symmetric alternatives
of the same identity. We also find this
generalization of predicted features in
feedback projections to primary sensory
areas. A scene remains predicted by the
top-down projection even if the spatial
frequency band of the input changes
[5]. Complex top-down predictions may
not only allow contextualization but also
a rapid reduction of prediction errors. A
neuron therefore signals not only the
presence of a feature but also whether
this feature was unexpected, and when
this surprise is resolved. Interestingly, the
current data also do not necessarily imply
a complex translation of one type of rep-
resentational format into another, but that
predictions can be communicated on the
backbone of straightforward wiring properties.

Schwiedrzik and Freiwald’s data are a striking demonstration that visual neurons perform predictive processing. Evidence of neural prediction marks a paradigm shift that engages theorists and empirical scientists in psychology, neuroscience, and philosophy. Philosophical themes being redefined using neurocomputational rules of prediction include conscious experience and embodied cognition (e.g., [6]). Algorithmic developments in artificial intelligence, brain-inspired computing, and robotics are rooted in neuronal prediction (e.g., [7]). Deficient predictive processing might contribute to psychotic symptoms [8] and neurodevelopmental disorders such as autism [9]. Proponents of predictive processing argue we can explain such data most parsimoniously in the framework of prediction. Others argue we have insufficient empirical evidence to substantiate predictive processing. Ultimately, theoretical frameworks will need to adjust to empirical data before we can model precisely how neurons predict or how prediction supports the full range of brain functions. Similar to the pioneering finding of Hubel and Wiesel of feature detectors in visual cortex [10], Schwiedrzik and Freiwald found prediction error-detection signals in the face-processing network. Such data are essential if prediction is to transform from a conceptual framework into a measurable and general mechanism of brain function.

Acknowledgements
This work was supported by the European Research Council (ERC) under the European Union’s Seventh Framework Programme (FP7) under grant agreement 26311751 (Brain reading of contextual feedback and predictions) and a Human Brain Project grant from the European Commission Horizon 2020 Research and Innovation Programme under grant agreement 720270 (HBP SGA1) (Context-sensitive multisensory object recognition: a deep network model constrained by multi-level, multi-species data), both awarded to L.M.

1Centre for Cognitive Neuroimaging, Institute of Neuroscience and Psychology, College of Medical, Veterinary and Life Sciences, University of Glasgow, 58 Hillhead Street, Glasgow, G12 8QB, UK

*Correspondence: Lars.Muckli@glasgow.ac.uk (L. Muckli).
https://doi.org/10.1016/j.tics.2017.12.001

References
1. Friston, K. (2005) A theory of cortical responses. Philos. Trans. R. Soc. Lond. B Biol. Sci. 360, 815-636
2. Schwiedrzik, C.M. and Freiwald, W.A. (2017) High-level prediction signals in a low-level area of the macaque face-processing hierarchy. Neuron 96, 89–97
3. Anissar, M. and Hochstein, S. (2004) The reverse hierarchy theory of visual perceptual learning. Trends Cogn. Sci. 8, 457–464
4. Larkum, M. (2013) A cellular mechanism for cortical associations: an organizing principle for the cerebral cortex. Trends Neurosci. 36, 141–151
5. Revina, Y. et al. (2017) Cortical feedback signals generalise across different spatial frequencies of feedforward inputs. Neuroimage Published online September 22, 2017, http://dx.doi.org/10.1016/j.neuroimage.2017.09.047
6. Clark, A. (2015) Surfing Uncertainty: Prediction, Action, and the Embodied Mind, Oxford University Press
7. Hawkins, J. and Blakeslee, S. (2004) On Intelligence: How a New Understanding of the Brain Will Lead to the Creation of Truly Intelligent Machines, Macmillan
8. Horga, G. et al. (2014) Deficits in predictive coding underlie hallucinations in schizophrenia. J. Neurosci. 34, 8072–8092
9. Van de Cruys, S. et al. (2014) Precise minds in uncertain worlds: predictive coding in autism. Psychol. Rev. 121, 649
10. Hubel, D.H. and Wiesel, T.N. (1959) Receptive fields of single neurones in the cat’s striate cortex. J. Physiol. 148, 514–591

Spotlight
Nature of Emotion Categories: Comment on Cowen and Keltner
Lisa Feldman Barrett,1,2,3,* Zulqarnain Khan,4 Jennifer Dy,4 and Dana Brooks4

Cowen and Keltner (2017) published the latest installment in a longstanding debate about whether measures of emotion organize themselves into categories or array themselves more continuously along affective dimensions. We discuss several notable features of the study and suggest future studies should consider asking questions more directly about physical and psychological variation within emotion categories as well as similarities between categories.

One long-standing debate in the science of emotion concerns whether emotions organize themselves into distinct classes, such as categories for fear, sadness, happiness, and surprise, or array themselves more continuously along affective dimensions such as valence and arousal. A recent article by Cowen and Keltner [1] is the latest installment in this debate. Cowen and Keltner asked whether people’s emotional experiences cluster together into a smaller set of categories or whether those experiences are best described along continuous affective gradients. To address these questions, participants reported their experience of emotions and other sentiments after watching film clips designed to evoke a range of feelings (Box 1). Based on their analyses, Cowen and Keltner concluded that reports of emotional experience were structured as 27 emotion clusters (i.e., categories), with fuzzy rather than firm boundaries. Furthermore, they argued that each category has a single pattern of affective features (e.g., anger is a high-arousal, unpleasant experience) and the similarity and differences among categories can be described by their proximity along affective features such as valence, arousal, and effort.

Study Designed for Robust and Replicable Findings
In this era of persistent concerns about the replicability and robustness of scientific findings, it is notable that Cowen and Keltner’s findings about self-reported emotional experiences replicate earlier lexical studies that more directly assess the structure of people’s emotion