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 number StG 2012-311751 (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.

[1]Centre 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, 815–836
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. Ahissar, 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. (2007) *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, 8072–8082
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, 574–591

Spotlight

Nature of Emotion Categories: Comment on Cowen and Keltner

Lisa Feldman Barrett,[1,2,3,*] Zulqarnain Khan,[4] Jennifer Dy,[4] and Dana Brooks[4]

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

concepts [2]. Furthermore, Cowen and Keltner introduced a methodological innovation that will move the field towards a more robust, generalizable science of emotion. Virtually all experiments in the science of emotion treat emotion-inducing stimuli as a fixed effect (all participants received all stimuli). Consequently, the majority of existing studies have likely overstated the strength and generalizability of the statistical evidence they report. In Cowen and Keltner's study, however, each participant viewed a subsample of the available movie clips (30 of the 2185 available), allowing the possibility of treating clips as a random-effects variable. Cowen and Keltner did not explicitly estimate stimulus variability and adjust for uncertainty due to stimulus sampling, but future studies can (for an example of modeling stimuli as a random effect, see [3]). Also of note, the analyses were conducted on the average of the individual ratings associated with each film clip, rather than on the individual ratings themselves. Because people differ in the number and precision of the emotion categories that describe their experience [4], it is important to establish how well these results apply to individual people.

Box 1. Brief summary of Cowen and Keltner (2017)

Eight hundred and fifty-three English speaking US participants each viewed a subset of 30 film clips designed to evoke a range of sentiments. A total of 2185 clips viewed by between nine and 17 participants. One group of participants reported on their emotional experiences by making categorical (yes/no) judgments of 34 emotion terms, such as anger, awe, confusion, and horror. A second group of participants reported on the affective features of their experiences by rating 14 affective properties such as valence, arousal, and effort, along a Likert-type scale (1–9; e.g., unpleasant to pleasant, subdued to stimulated, and no effort to enormous effort). A third set of participants freely labeled their experiences by choosing words from a database of 600 terms. Using a series of dimensionality reduction techniques to examine the covariances in the ratings, Cowen and Keltner reported that ratings of emotional experience were structured as 27 emotion groupings (i.e., categories), with fuzzy rather than firm boundaries, distributed along continuous affective gradients (i.e., dimensions). Additional details for the various analyses can be found in Cowen and Keltner [1], but one main observation is worth noting: the split-half canonical correlation analysis (CCA) that was used to identify the emotion groupings is equivalent to a supervised, confirmatory data-analytic approach. Participants categorized their experience using 34 emotion adjectives by making dichotomous yes/no responses, so a split-half CCA reflects how well the rating covariances conform to prespecified emotion categories rather than inductively discovering whether or not emotion categories exist in the first place. A more inductive approach would begin with participants making all ratings, including those involving emotion adjectives, on a continuous Likert-type scale, along with the use of an unsupervised dimensionality reduction approach using conventional statistical guidelines for extracting components (e.g., eigenvalues >1).

But What Kind of Categories?

The categories versus dimensions debate that motivated Cowen and Keltner is more properly a debate about the nature of emotion categories. Are emotion categories classical (or natural kind) categories with necessary and sufficient features, as proposed by the basic emotion approach (as described in Cowen and Keltner), or are they conceptual categories (on par with other biological categories as discussed by Darwin in On the Origin of Species [5], as discussed in [4])?

All instances of emotion have affective features – emotional experiences feel pleasant or unpleasant, they involve feelings of arousal or quiescence, they involve some degree of effort, and so on. So, scientists can ask: are instances of the same emotion category highly similar in their affective features (so that they strongly cluster) or do they substantially vary from one another in their affective features? For example, instances of anger, sadness, fear, and happiness can be either unpleasant or pleasant [6]. We can also ask: are emotion categories qualitatively different from one another in their affective features (with firm boundaries between categories) or are there instances of different emotion categories that are more similar in affective features (so that the categories have fuzzy boundaries or even overlap in their features)? For example, are there instances of anger, fear, and excitement that are experienced as similarly pleasant and high in arousal?

Cowen and Keltner found evidence of variation, but the pressing question is how much variation? Using their analysis strategy, they found that emotional experience (as reported by participants) configure as prototype categories with fuzzy boundaries (replicating [2]), meaning that the instances of an emotion category supposedly share a single pattern of affective features, but vary primarily in the degree to which those features are present. These findings are inconsistent with research in other measurement domains of emotion that clearly shows how instances of emotion belonging to the same category vary considerably in their features. Furthermore, instances of different emotion categories have considerable similarity in their features. This is true when measuring affective properties in emotional experiences (e.g., [6]), when measuring facial actions that serve as emotional expressions [7], when measuring autonomic nervous system (ANS) features of emotion [8], and when measuring the neural patterns associated with emotion categories [9].

A category is a group of instances that are similar for some purpose. A conceptual category has instances that are similar in function but that do not necessarily share perceptual features [10]. If the instances of the same emotion category have variable features across contexts (i.e., if the affective feelings differ, the facial movements differ, and the ANS features differ), then an emotion label does not uniquely describe any instance of that category; nor does the category prototype (which is a statistical abstraction over variable instances and not a biological essence

[4]). Knowing that a person is angry does not reveal whether she/he is feeling pleasant or unpleasant; it does not tell you whether she/he is preparing for attack or feels defeated. The features themselves (including affective features) provide additional information that the emotion label alone cannot. Such findings are inconsistent with classical views of emotion (such as the basic emotion approach of Cowen and Keltner) and are more consistent with constructionist views of emotion (e.g., [2,4]). It would be interesting to reanalyze Cowen and Keltner's data using a more inductive, discovery-based approach rather than the more confirmatory approach they chose (Box 1). Future studies that adopt a discovery-based approach will make it possible to empirically compare different hypotheses about the nature of emotion categories.

Acknowledgments

Preparation of this article was supported by a US National Cancer Institute grant (U01 CA193632) and a US National Institute of Mental Health grant (R01 MH109464) to LFB, a US Army Research Institute for the Behavioral and Social Sciences grant (W911NF-16-1-0191) to LFB and JD, and a US National Institute of Mental Health grant (R01 MH113234) to LFB, DB, and JD. No authors have any financial interests to disclose. The views, opinions, and/or findings contained in this paper are those of the authors and shall not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documents. No authors have any financial interests to disclose.

[1]Department of Psychology, Northeastern University, Boston, MA 02115, USA
[2]Psychiatric Neuroimaging Division, Department of Psychiatry, Massachusetts General Hospital and Harvard Medical School, 149 13th Street, Charlestown, MA 02129, USA
[3]Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital and Harvard Medical School, 149 13th Street, Charlestown, MA 02129, USA
[4]Department of Electrical and Computer Engineering, Northeastern University, Boston, MA 02115, USA

*Correspondence:
l.barrett@northeastern.edu (L.F. Barrett).

https://doi.org/10.1016/j.tics.2017.12.004

References

1. Cowen, A.S. and Keltner, D. (2017) Self-report captures 27 distinct categories of emotion bridged by continuous gradients. *Proc. Natl. Acad. Sci. U. S. A.* 114, E7900–E7909
2. Russell, J.A. (2003) Core affect and the psychological construction of emotion. *Psychol. Rev.* 110, 145–172
3. Jack, R.E. and Schyns, P.G. (2017) Towards a social psychophysics of face communication. *Annu. Rev. Psychol.* 68, 269–297
4. Barrett, L.F. (2017) *How Emotions Are Made: the Secret Life of the Brain,* Houghton–Mifflin–Harcourt
5. Darwin, C. (1859/2001) *On the Origin of Species* (Facsimile ed.), Harvard University Press.
6. Wilson-Mendenhall, C.D. *et al.* (2015) Variety in emotional life: within-category typicality of emotional experiences is associated with neural activity in large-scale brain networks. *Soc. Cogn. Affect. Neurosci.* 10, 62–71
7. Duran, J.I. *et al.* (2017) Coherence between emotions and facial expressions: a research synthesis. In *The Science of Facial Expression* (Fernandez-Dols, J.-M. and Russell, J. A., eds), pp. 107–129, Oxford University Press
8. Siegel, E.H. *et al.* Emotion fingerprints or emotion populations? A meta-analytic investigation of autonomic nervous system features of emotion categories. *Psychol. Bull.* (in press).
9. Clark-Polner, E. *et al.* (2017) Multivoxel pattern analysis does not provide evidence to support the existence of basic emotions. *Cereb. Cortex* 27, 1944–1948
10. Barsalou, L.W. (2012) The human conceptual system. In *The Cambridge Handbook of Linguistics* (Spivey, M.J., ed.), pp. 239–258, Cambridge University Press