

Approximately half of the human genome, known as the repeatome, consists of repetitive DNA sequences. The repeatome includes more than one million tandem repeats – sections of DNA in which a sequence is replicated many times in tandem – whose biology remains largely unexplored. More than 50 diseases are known to be caused by expansion of a tandem-repeat sequence in a single gene; among them are Huntington's disease and fragile X syndrome<sup>1</sup>. But less-well understood is the role of tandem repeats in polygenic diseases, which have more-complex genetic underpinnings. On page 246, Mitra *et al.*<sup>2</sup> use a newly developed bioinformatics approach to

identify tandem repeats associated with one such condition, autism spectrum disorder.

Autism spectrum disorder (ASD) is highly prevalent, affecting approximately 1–2% of children in the United States (see [go.nature.com/38sqvhd](https://www.nature.com/38sqvhd)), although this varies internationally. It is characterized by atypical neurodevelopment, communication deficits, atypical social functioning, restricted interests and repetitive behaviours. Although progress is being made<sup>3</sup> in discovering the genetic basis of ASD, it remains poorly understood. Changes in the number of copies of large segments of DNA, along with other genetic variants, have been previously implicated<sup>3</sup>, but the capacity

people move their faces to express emotion and how other people infer emotional meaning in those facial movements<sup>5,8,9</sup>.

For some scientists, these observations seem consistent with alternative evolutionary hypotheses of emotion<sup>5</sup>, but for others, the findings introduce another question about universality. Might situated emotional expressions – how people express emotion in certain situations – be universal across cultures? This is the question that Cowen *et al.* set out to answer, by investigating whether certain facial configurations are found in similar contexts worldwide.

Previous studies have almost exclusively used artificial methods to observe how people express emotion with facial movements and infer emotional meaning in facial configurations. One common task is to arm people with a small, preselected set of emotion words (such as 'anger' or 'sadness') and to ask them to label posed, disembodied, contextless faces (such as a person smiling) with the word that they think best describes the emotion on each face. This method, when compared with others, has been consistently shown to inflate support for the universality hypothesis<sup>4–6,10,11</sup>. A related method offers people a single, impoverished scenario for an emotion category ('You have been insulted, and you are very angry about it', for instance). They are then asked to pose the facial configuration they believe they would make to express instances of that emotion category. Both approaches seem to encourage people to rely on stereotypes about emotional expressions that do not typically reflect the varied ways in which people express emotion in everyday life<sup>5</sup>.

One major strength of Cowen and colleagues' effort is that they analysed facial configurations in more-natural settings. The authors curated YouTube videos of people in natural social contexts, such as at weddings, sports events or playing with toys (Fig. 1). Another strength is the scope of their sampling: the paper sampled more than 6 million videos from 144 countries in 12 regions around the world. Perhaps just one other study<sup>12</sup> has so far matched this broad scale of sampling.

Cowen and colleagues used a powerful machine-learning method involving what are called deep neural networks (DNNs) to assess the extent to which specific facial configurations (called facial expressions by the authors) could be reliably observed in the videos across cultures. They trained one DNN to classify the facial configurations in the videos as emotional expressions and a second DNN to classify the videos' contextual elements. Then they estimated the associations between the two classifications to determine how frequently each class of facial expression occurred in videos that were classified as containing similar contextual elements, and compared

### Psychology

# Debate about universal facial expressions goes big

Lisa Feldman Barrett

An analysis of more than 6 million YouTube videos finds that people around the world make similar facial expressions in similar social contexts. The study brings data science to the debate about the universality of emotion categories. **See p.251**

When you are angry, do you scowl, cry or even laugh? To what extent do your facial movements depend on the situation you are in – whether you are in a formal meeting, say, or at home with your family? And do other people around the world express anger in such situations in the same way? These questions are at the centre of a contentious scientific debate about the nature of emotion that has raged for more than a century. On page 251, Cowen *et al.*<sup>1</sup> enter the fray. The authors used a type of machine learning to investigate whether people adopt the same facial expressions in similar contexts across cultures. Their results are sure to be the subject of lively discussion.

The universality debate is central to an understanding of the nature, causes and functions of emotions. The universality hypothesis proposes that people around the world consistently use certain configurations of facial-muscle movements to specifically express instances of a certain category of emotion. For example, people are said to frown in sadness consistently enough (and frown infrequently at other times) for frowning to be recognized as the expression of sadness the world over. The same goes for scowling in anger, smiling in happiness and so on. Such expressions are thought to have evolved to signal emotional information in contexts that posed fitness challenges for our hunter-gatherer ancestors<sup>2,3</sup>. Sometime in our

evolutionary past, the hypothesis goes, our ancestors capitalized on specific, universal emotional expressions as a means of communicating those challenges to one another, helping them to survive and reproduce.

There is considerable debate about whether the published empirical evidence provides support for the universality hypothesis. A large body of literature seems to back it up, but these studies have been consistently criticized for methodological problems<sup>4,5</sup>. Furthermore, a growing number of experiments

**“Perhaps just one other study has so far matched this broad scale of sampling.”**

(both in modern, urban populations<sup>5</sup> and in small-scale foraging communities<sup>6</sup>, such as the Hadza hunter-gatherer group in Tanzania<sup>7</sup>) using diverse methods seem to call the universality hypothesis into doubt. This research suggests that expressions of emotion might be context-specific, and therefore variable across instances – the facial movements in each instance being tailored to the immediate context, as is the case for all other motor movements. In fact, study after study has revealed the potent influence of context on both how

these association patterns across the 12 world regions.

In each region, certain facial configurations were observed relatively more frequently in certain contexts. The associations were subtle (that is, the magnitude of associations between facial expression and context tended to be weak), but, remarkably, the pattern of expression–context association observed in the videos from one world region were similar to those in other world regions. For example, in the various regions sampled, people in the videos made facial-muscle movements labelled as ‘awe’ more frequently in contexts that involved fireworks, a father, toys, a pet and dancing than in contexts that did not include these elements, such as those involving music, art, the police and team sports. In fact, facial expression–context patterns were 70% preserved across the 12 regions, suggesting a degree of universality in how people across the world express emotions in various situations.

To properly assess Cowen and colleagues’ findings and interpretation, let’s peek under the paper’s hood at a few of the many details worth considering. First, let’s consider the DNN used by Cowen *et al.* to detect facial configurations. The DNN learnt from human evaluators (‘raters’), who annotated the facial movements contained in each video clip by choosing from a set of English words describing emotion that were provided by the authors. According to previous research<sup>4,5,10,11</sup>, these annotations might have been strongly influenced by the word set given to raters, who might have offered different annotations if allowed to freely label the videos with words of their own choosing, as the authors acknowledge. In addition, by using emotion labels to annotate the faces (‘anger’, ‘fear’, ‘sadness’ and so on), rather than descriptive terms (such as ‘scowl’, ‘wide-eyed gasp’ and ‘frown’), the raters were, in effect, offering inferences about the emotional meaning of the facial movements. The authors’ goal might have been to train the DNN to identify similar sets of facial configurations, rather than to recognize instances of emotion per se, but, ultimately, the DNN was trained on raters’ inferences about the emotions. Neither the raters nor Cowen *et al.* can say with certainty which, if any, instances of emotion were actually being experienced by the people in the videos. Conflating descriptions of facial movements with the interpretation of their emotional meaning is potentially problematic because when more-sensitive methods are used, the expressive meaning of facial movements varies considerably between people and cultures<sup>5–8</sup>.

Second, the raters viewed the faces in context, not in isolation, so it is difficult to claim that the raters’ annotations were driven solely by the emotional meaning of facial movements. Instead, the annotations might



**Figure 1 | Children playing in South Africa (left) and the United Kingdom.** Cowen *et al.*<sup>1</sup> used machine-learning techniques to investigate whether people make similar facial expressions in similar contexts around the world.

refer to the meanings of the faces in context. A growing body of evidence suggests that this is the case<sup>5,9–11</sup>. If the raters applied emotion labels to the faces in a way that was indeed influenced by the context surrounding the faces, then it would mean that facial expressions labelled in one way (as happiness, for instance) might represent the raters’ inferences about the emotional meaning of the face in a given context. This would imply a lack of independence between the DNN trained to recognize facial expressions and the DNN being trained to classify contexts. (However, an experiment performed as a control in the current study suggests that contextual information did not influence how the first DNN labelled facial configurations.)

Third, and perhaps most central to the question of universality, the annotations were made by English speakers from a single country, India. Ultimately, then, the DNN was trained to assign facial configurations to emotion categories on the basis of the inferences of these raters alone, rather than being based on the perceptions of diverse people from countries around the globe. Evidence suggests that people from different countries would make different inferences<sup>5,6</sup>. In effect, the DNN enshrined a set of culture-specific beliefs about emotional expressions in its code.

Together, these details suggest that we cannot be sure that the findings reported by Cowen and colleagues reflect evidence of universal expressions in context. A sceptical reader might interpret these findings as showing how emotional miscommunication occurs across cultures, because one group of human raters used their own culture-specific beliefs to interpret the emotional meaning of facial

movements in context in other humans from across the modern world. The ultimate value of Cowen and colleagues’ study might lie not in the answers it provides, but in the opportunity for discovery that it opens up. Their work underscores the pressing need for a vigorous scientific effort to observe, describe and understand the myriad ways in which people move their faces to express emotion in real-world contexts, without depending solely on English speakers’ beliefs and stereotypes about emotional expressions.

**Lisa Feldman Barrett** is in the Department of Psychology, Northeastern University College of Science, Boston, Massachusetts 02115, USA. e-mail: l.barrett@northeastern.edu

1. Cowen, A. S. *et al.* *Nature* **589**, 251–257 (2021).
2. Shariff, A. F. & Tracy, J. L. *Curr. Direct. Psychol. Sci.* **20**, 395–399 (2011).
3. Cosmides, L. & Tooby, J. in *Handbook of Emotions* 2nd edn (eds Lewis, M. & Haviland-Jones, J. M.) 91–115 (Guilford, 2000).
4. Russell, J. A. *Psychol. Bull.* **115**, 102–141 (1994).
5. Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M. & Pollak, S. D. *Psychol. Sci. Public Interest* **20**, 1–68 (2019).
6. Gendron, M., Crivelli, C. & Barrett, L. F. *Curr. Direct. Psychol. Sci.* **27**, 211–219 (2018).
7. Gendron, M. *et al.* *Sci. Rep.* **10**, 3867 (2020).
8. Rychłowska, M. *et al.* *Proc. Natl Acad. Sci. USA* **112**, E2429–E2436 (2015).
9. Aviezer, H. *et al.* *Psychol. Sci.* **19**, 724–732 (2008).
10. Hoemann, K. *et al.* *Emotion* **19**, 1292–1313 (2019).
11. Gendron, M. & Barrett, L. F. in *The Science of Facial Expression* (eds Fernández-Dols, J.-M. & Russell, J. A.) 15–36 (Oxford Univ. Press, 2017).
12. Srinivasan, R. & Martinez, A. M. *IEEE Trans. Affective Comput. Comput.* <https://doi.org/10.1109/TAFFC.2018.2887267> (IEEE, 2018).

This article was published online on 16 December 2020.

### **Correction**

The original version of this News & Views described a control experiment as assessing the raters; instead it assessed the DNN.