International Society *for* Research *on* Emotion

Emotion Researcher

ISRE's Sourcebook for Research on Emotion and Affect

http://emotionresearcher.com

February 2022

Editors

Cain Todd

Lancaster University
c.todd@lancaster.ac.uk

Eric A. Walle

University of California, Merced
ewalle@ucmerced.edu

Published by the International Society for Research on Emotion

Table of Contents

Editors' Column

How does it feel?

Cain Todd & Eric A. Walle

After a relatively long hiatus, we are excited to publish this issue devoted to the connection between emotion and feeling. Although hardly a neglected topic, the nature and role of feeling in emotion remains rather opaque and controversial, as evidenced by all three contributions to this issue.

In the first invited article, Ralph Adolphs (California Institute of Technology) addresses the question of where the study of feelings ought to fit within current scientific investigations of the nature of emotion. Stressing the importance of studying emotional behavior, he suggests that it is curious how, unlike other states – such as memory or perception – researchers often want to put the conscious experience of emotion centre stage. Against theories, such as those of LeDoux and Barret, that give an essential place to feelings in their accounts of what emotions are, Adolphs contends that in order to study feelings we first need to know what emotions are, and this can only be achieved by understanding their functional role. He concludes that scientists should think of emotions as latent variables that provide causal explanations of behavior rather than of conscious.

In the second contribution, Jonathan Gratch (USC Institute for Creative Technologies) also emphasizes the distinction between emotional behavior and emotional experience that arises when confronting significant problems facing various forms of Affect Recognition technology. Gratch examines reasons to be skeptical of the idea that a person's emotional state can be accurately inferred by surface cues such as facial expressions and voice quality, or through physiological signals such as skin conductance or heart rate variability. This skepticism is justified, he argues, insofar as the components of emotion are loosely connected, and expressions are highly-dependent on the social context. He highlights several ways in which affect recognition technologies can yield misleading results, in particular concerning problems in recognizing affective expressions, and problems in understanding what can be concluded from these expressions. Nonetheless, if complex contextual information can be appropriately taken into account, Gratch posits that some algorithms can play significant predictive roles and thus calls to ban affect recognition are misguided.

In the final article, David Sander (CISA, University of Geneva) examines the complex issue of the content of feelings within multicomponential accounts of emotion. He suggests that understanding the relationship between the feeling component and other components of emotion may help to illuminate a long-standing debate in emotion research; namely, whether the bodily changes associated with an emotion are a cause or a consequence of the emotion. He observes that the feeling may be (at least partly) determined by a change in the body state, while the other components of emotion may not be caused this body state. Further, he suggests that just like one can have a physiological feeling (e.g., feeling of an increased heart rate), one could also have a feeling of appraisal outcomes (e.g., feeling of uncertainty). This opens up the interesting question, for further research, of whether appraisal outcomes or action tendencies can be felt as direct inputs, or whether they are felt only via bodily changes

ISRE Spotlight

We are pleased to highlight the innovative work by Alan Cowen in our Spotlight feature. Cowen provides a refreshing take on the important claim that emotion theory be guided by data, rather than the other way around. His premise is supported by the potential offered by big data and artificial intelligence. As a graduate student and now Chief Scientist of Hume AI, Cowen has developed powerful tools that leverage vast datasets with sophisticated computational approaches, resulting in a data-driven approach to studying emotional expressions and experiences. The results provide a clearer understanding of how data supports (or refutes) existing theories of emotion, as well as showcase the potential of this approach for catapulting the field of emotion science forward.

Announcements

In addition to the excellent contributions in this issue, there are also some important announcements and points of mention.

First, there is an exciting announcement regarding the upcoming Biennial ISRE Meeting to take place in Los Angeles, California, USA. Conference Organizers Jonathan Gratch and Stacy Marsella describe their preparations for the meeting in a special piece in the current issue.

Second, we want to highlight the continued work being done by the Early Career Researchers Section. This group has put together numerous initiatives, including guest speakers, panels, and awards. We commend them for their excellence and for making ISRE a more well-rounded group of researchers.

Finally, we would like to convey our appreciation for the patience shown by our readership. The past two years have presented numerous challenges for us as editors, ranging from soliciting articles and interviews from busy researchers to balancing our work-life obligations. As parents with small children, the uncertainties and unexpected schedule changes have made for slow progress in publishing new issues of Emotion Researcher. However, we maintain our optimism that the pandemic will recede and a sense of normalcy will return.

Wishing each of you a safe, healthy, and productive year,

Warmly,

Eric & Cain

Cain Todd is Senior Lecturer in Philosophy at Lancaster University (UK). His research covers a wide range of issues centring on emotions and evaluative experience, most recently the phenomenology and objectivity of emotional experience and the role of attention and imagination therein. His co-edited collection *Emotion and Value* (OUP) was published in 2010, and his new monograph *Aesthetics and Emotion* (Bloomsbury) is due to appear in 2022/23.

Eric Walle is Associate Professor of Psychological Sciences at the University of California, Merced. His theoretical writings emphasize the functions of emotions, particularly in interpersonal contexts. His empirical work examines emotional development, principally in infancy and early childhood, as well as how individuals perceive and respond to emotional communication. He is also a co-editor of the *Oxford Handbook of Emotional Development* (2022).

ISRE Matters

Ursula Hess

Professor of Social and Organizational Psychology, Department of Psychology
Humboldt University, Berlin
Ursul.Hess@hu-berlin.de

Ursula Hess, ISRE President

Dear ISRE Members,

As is customary I am using this space to wish you all a good year to come and to highlight some of the exciting content of Emotion Researcher. But first a piece of good news. Even though the last year did not deliver on the anticipated return to normalcy, we do need to continue to make plans in hopes of some form of normalcy.

In this vein, the bi-annual ISRE conference will take place from July 15 to July 18, 2022 at the University of Southern California. We look forward to exciting keynotes by Antonio Damasio, Barbara Fredrikson and Eran Halperin as well as to a stimulating program of talks, posters and symposia. I hope to see many of you there.

One potential stimulus for discussion are the featured articles in this Emotion Researcher. Ralph Adophs (con) and David Sander (pro) present arguments for and against the usefulness of the concept of feelings in emotion research. Two articles that I found thought provoking in the best way.

A complementary point is made by Jonathan Gratch who discusses the important difference between recognizing or classifying emotion expressions and understanding emotions. He points to the important role of context in this discussion and notes that emotions cannot really be understood without their context and that both context and the real-world knowledge of the perceiver contribute to understanding. Finally, Alan Cowan discussed “Semantic Spaces, Big Data, and AI in Emotion Science.”

Your President,
Ursula

ISRE Early Career Researchers Section

ISRE Early Career Researchers Section: Update on Initiatives

Tanja S. H. Wingenbach, Claire Ashley, Manuel F. Gonzalez, Soohyun (Ashley) Lee, Olivia S. Mendoza, Daeun Shin, Marwen Belkaid, & Yong-Qi Cong

The International Society for Research on Emotion - Early Career Researchers Section (ISRE ECRS) is a platform for emotion/affective science researchers from any field, discipline, method, or culture. The ISRE ECRS organizes professional and social meetings for early career emotion researchers, both during ISRE conferences and between meetings. Additionally, the ISRE ECRS strives to create and maintain member support through awards, career development opportunities, expert feedback, webinars, and more.

The ISRE ECRS continues to grow since its launch in 2013 and has implemented several initiatives for early career emotion researchers. In 2022, the ISRE ECRS will host the career development series, student poster award, and social event at the bi-annual ISRE conference.

Career Development Series:

The career development series is designed to enable ISRE's early-career researchers to explore and prepare for successful careers as emotion scientists. Now entering its second year, the career development series will include virtual workshops regarding research grant and paper development, virtual writing sessions, and an in-person professional development workshop at the 2022 ISRE conference. We believe these career development opportunities will help early-career emotion researchers gain insights from senior researchers and grow their professional networks through interacting with fellow ISRE members of all career stages.

Student Poster Award:

The student poster award aims to recognize outstanding ISRE early-career researchers and increase chances for competitive funding. The top three posters will be selected based on the research and presentation quality by on-site judges. The three finalists will receive a personalized award certificate at the 2022 ISRE conference's closing awards ceremony, and the best poster winner will receive a monetary prize. We believe this poster award initiative will be an important step for early-career researchers to boost their career trajectory and increase their research visibility.

Social Event:

The goal of the social event is to help build a sense of community among ISRE's early-career researchers and enable ECRs to build their networks at the 2022 conference site. We plan to organize an on-site social event on the first conference day, which will be a great platform for ISRE's early-career researchers to establish social networks and find future collaborators.

The aforementioned career development series, student poster award, and social event will be advertised soon. Please keep an eye out for further information on the ISRE website, Listserv, and social media outlets!

Our team is excited to implement initiatives that align with the interests of ISRE and support early career emotion researchers. We are grateful for ISRE's support in implementing these initiatives, the publishers that have supported our initiatives financially, the senior researchers who participate in our initiatives, and the early career researchers who have been part of our journey thus far.

We would like to thank Melina West for all her work within the ECRS over the last couple of years and wish her all the best for her future.

The ECRS welcomes its new team additions: Daeun, Marwen, and Yong Qi.

Would you like to volunteer within the ISRE ECRS?

If you are an ISRE Associate Member[1] and would like to get involved, please get in touch. We are excited for you to help us best support our emotion research community.

Please note that volunteer commitment should be at least 1 year and requires continuous involvement.

If you are interested in playing an active role in the ISRE ECRS, please email Tan (tanja.wingenbach@bath.edu). In your interest email, include a short bio, a statement of which initiative you prefer to get involved with, and why.

Join us on Facebook!

Are you an early career emotion scientist or faculty that support early career emotion scientists? Join our Facebook page: **https://www.facebook.com/groups/ISRE.JRS/?ref=br_rs**

For any other questions or comments, please email Claire Ashley (claire.ayako@gmail.com)

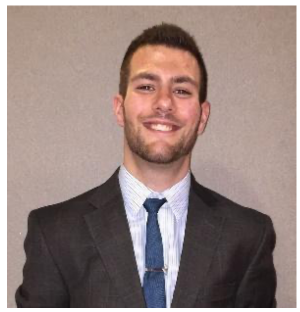
Current ISRE ECRS Board

Chair: Tanja S. H. Wingenbach, PhD (Postdoctoral Senior Research Fellow, University of Zurich/University Hospital Zurich, Switzerland)

Tan coordinates and initiates activities, liaises with the ISRE president/board, serves as a spokesperson of the ECRS, and represents the ECRS within the ISRE board.

Secretary: Claire A. Ashley, M.Sc, (Psychology Assistant, Park Terrace Care Center, USA)

Claire is responsible for internal and external communications and liaising with ISRE conference organizers.



Career Development Series – Events Coordinator: Manuel F. Gonzalez, PhD (Assistant Professor, Seton Hall University, USA)

Manuel oversees all aspects of the career development series, including developing and scheduling events, as well as recruiting panelists and speakers.

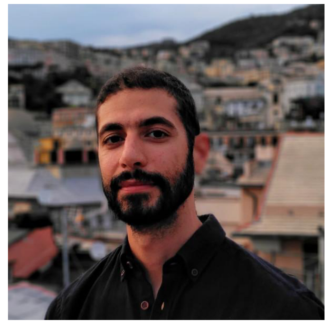
Poster Award Coordinator: Soohyun (Ashley) Lee, PhD Candidate (Baruch College & The Graduate Center, City University of New York, USA)

Ashley manages the poster awards at the ISRE conference (e.g., contacting ISRE board, communicating with the jury, soliciting submissions).

[1] ISRE Associate Membership is defined as: "less-established emotion researchers who have not yet obtained the terminal degree in their field or are engaged in postgraduate training. Associate Members are typically advanced graduate students or postdoctoral students."

Additional volunteers:


Olivia S. Mendoza, M.A. (University of the Philippines Baguio, Philippines)


Marwen Belkaid, PhD (Postdoc researcher, Istituto Italiano di Tecnologia, Italy)


Daeun Shin, PhD student (Arizona State University, USA)


Yong-Qi Cong, PhD Candidate (University of Amsterdam, The Netherlands)

ISRE Biennial Meeting Update

The ISRE 2022 Meeting

Jonathan Gratch[1] & Stacy Marsella[2]

Conference Organizers

[1]Computer Science, Psychology, and Media Arts and Practice
University of Southern California

[2]Computer Science and Psychology
Northeastern University

We are happy to announce that the bi-annual ISRE (International Society for Research on Emotion) conference will take place in-person on the 15-18th of July 2022 on the campus of the University of Southern California, Los Angeles USA.

The ISRE conference is an exciting opportunity to meet international colleagues, present your work, and to stay up-to-date with the latest developments in emotion research. ISRE members study emotions from a wide range of disciplines including psychology, neuroscience, philosophy, sociology, linguistics, affective

Ronald Tutor Campus Center, site of ISRE 2022.

computing, history, anthropology, art, and design.

The ISRE conference 2022 will include keynote addresses by Antonio Damasio, Barbara Fredrickson and Eran Halperin. Additionally, the meeting will feature 6 preconference sessions providing a more specialized environment for discussion of important topics relating to emotion.

We look forward to seeing you in sunny California in July!

Go to isre22.org to register or learn more details.

15-18 July
isre
International Society
for **Research** *on* **Emotion**
'22 LOS ANGELES

ISRE Spotlight

Semantic Spaces, Big Data, and AI in Emotion Science

Alan Cowen, PhD

Chief Scientist of Hume AI
alan.cowen@berkeley.edu

The great philosopher David Hume wrote that "Reason is, and ought only to be, the slave of the passions, and can never pretend to any other office than to serve and obey them" [1]: emotion is the force that propels thought and action.

To understand Hume's thesis, think about what happens when a kid asks you "why?" over and over again. Why you wake up in the morning, why go to work every day, why you need money, why you want to buy tickets to the Westminster Dog Show—your answer is eventually something about what brings you amusement, joy, or satisfaction, and what leaves you sad, angry, or regretful. And your next answer, if you're familiar with the state of emotion science, will be that "there's no scientific consensus about that yet."

If emotions are so fundamental, why is there so little scientific consensus about them? Emotion scientists can't seem to agree on answers to some of the most basic questions about emotion. Why do some people find dogs amusing? Why do we seek "amusement" out? Is "amusement" an evolved response, shared by all humans? Is it culture-specific?

In this Spotlight, I will discuss how I have sought out more definite answers to these questions. When I began on this journey around ten years ago, it was as a skeptical data scientist. I quickly saw that scientists with different theories of emotion were interpreting the same data in radically different ways. And the problem wasn't with the theories, I concluded—it was with the data. The existing data were often too small, too constrained, to provide unequivocal support to any of the complex theories that had become popular in the field.

I began to look for answers using richer, broader, more open-ended approaches, and found that new tools—the internet, large-scale statistics, machine learning—offered new possibilities for understanding human emotion. This marked a turning point on a journey that would take me from improv clubs to museums to record emotional expressions in audiences and ancient sculptures, to the offices of big tech companies, where I advised teams seeking to build technology that accounts for human emotion.

In this Profile, I summarize that journey, discuss why the future of emotion science hinges on big data and artificial intelligence, and provide resources that I hope will make it easier for scientists without a computational background to begin applying state-of-the-art AI to their data.

The Need for "Steel Manning" in Emotion Science

Before delving into big data and AI, let us return to the question of why emotion science has not reached much consensus.

In philosophy, there's a practice called "steel manning"—before critiquing an argument, one restates it in a manner with which its proponents would agree. It's the opposite of "straw manning." I believe this is something that emotion scientists need to start doing more often; if they did, I think the need for bigger data and new methods would be more apparent.

Proponents of one popular theory of emotion often begin papers by pointing out the misconception that emotions are completely variable across cultures. Their evidence to the contrary shows consistency in emotional behavior—across individuals, demographics, cultures, contexts, and so forth.

Proponents of the contrasting theory often point out that there are scientists who believe a fixed number of emotions are universal. Their evidence shows variability in emotional behavior—again, across individuals, demographics, cultures, contexts, etc.

Papers from each approach begin with the premise that some people think there is only consistency, or only variability, in a set of emotional behaviors. But to "steel man" these theories, you would have to start by acknowledging that everyone already agrees that there is both consistency and variability in emotional behavior. The real question is, what exactly *is* consistent in emotional behavior? What *is* it that varies across individuals, demographics, cultures, and contexts?

If the goal is to explain human behavior, then this is where the meat of emotion theory lies. It is not in broadly writ notions of universality or variability in human behavior across cultures. It is in the more specific, auxiliary claims that you often find closer to the results section of a paper, such as the following:

- People in different cultures perceive similar levels of "valence"—the level of pleasure or displeasure—in a facial expression. They also perceive similar levels of "arousal"—the level of calmness or excitement. Other dimensions, like "amusement," are variable across cultures, except insofar as they are correlated with valence and arousal [2–4].
- People in different cultures perceive similar levels of emotions like "amusement" in facial expressions, above and beyond similarities in valence and arousal [5].
- People in different cultures perceive things like unexpectedness, abruptness, and goal-congruence in facial expressions, and this is what explains the perception of emotions like "amusement" [6]

These are "steel man" versions of claims that have been made in prominent papers in emotion science. They are testable claims about what it is precisely that is consistent across people and what it is that is variable. They are mutually exclusive. If we are to take these claims seriously, we must ask next: What does it really take to put them to the test?

A Computational Approach to Emotion

Scientists have come up with dozens of theoretical constructs to explain emotional behaviors: from the "basic six" emotions—anger, disgust, fear, happiness, sadness, and surprise – to valence and arousal, to dimensions like unexpectedness, abruptness, and goal-congruence. In everyday life, people commonly use dozens of different concepts to describe their emotions. Meanwhile, emotional behaviors have dozens, if not hundreds, of parameters—40 or so facial muscles, lots of ways we can move our bodies, lots of ways we can manipulate our voice,

Alan Cowen, PhD

and so forth. And we experience emotion in an incredibly wide variety of situations.

What this means is that to really put any of the claims I laid out above to the test, you need to examine a plethora of theoretical constructs, everyday concepts, emotional behaviors, or experiences, and map out their relationships. To do this systematically is to derive what I have termed a "semantic space" of emotion.

Semantic spaces of emotion are defined by three properties (Figure 1A). The first is their *dimensionality*: How many different kinds of emotion are there? The second is the *distribution* of states within the space: Are there discrete boundaries between emotions, or is there overlap [7,8]? The third is the *conceptualization* of emotion: What concepts most precisely capture the variation in the emotional experiences and emotional expressions that people consider to be distinct [9,10]? Do experiences and expressions correspond to specific emotions, like "interest," "sadness," and amusement," or broader evaluations like "valence" and "arousal" [2,7,11] or "certainty" [12]?

To capture semantic spaces of emotion, we needed new kinds of data and new statistical approaches. It turns out that the small number of emotions and prototypical stimuli [3,13] that are by far the most studied capture only a fraction, about 30%, of the information conveyed by emotion concepts and emotional expressions

[14]. Accurately characterizing the meaning of what we say we feel and what we express turns out to require measuring participants' responses to vast arrays of evocative stimuli and expressions [15]. It requires moving beyond traditional statistical approaches like recognition accuracy [13] and factor analysis [2,12], approaches that presuppose either one-to-one mappings between emotional behaviors and concepts (e.g., "anger") or that these relationships reduce to a small number of dimensions.

Semantic spaces of emotion satisfy a broader goal: to separate signal—what the data at hand is capable of explaining—from noise. To carry signal, or meaning about emotion, all instances of a particular behavior (e.g., a smile) do not need to map to the same emotional state. Indeed, our studies have shown that the same facial expressions used in everyday life are sometimes used in multiple, very distinct contexts, such as sentimental expressions of musical performers that resemble expressions of pain [16].

Grounded in these principles, we have used large-scale data and new statistical tools to derive semantic spaces of emotion in facial-bodily expression [18], nonverbal vocalization [19], speech prosody [21], and the feelings evoked by music [20] and video [8], within and across cultures [20,21]. In different studies, we had thousands of people evaluate music samples provided by hundreds of other participants in the U.S. and China; speech samples recorded by hundreds of actors in five countries; vocalizations from improv actors; and much more (Figure 1B-E). We even toured museums to study facial expressions in ancient American sculptures (Figure 1F). Our results were both surprising and consistent. Over 25 emotions are associated with distinct profiles of behavior, many more than are typically account for in studies of emotion. Specific emotions like "amusement," more than valence and arousal, organize experience, expression, and neural processing. Emotions are not discrete, but systematically blended.

When we move beyond traditional models to study these broader semantic spaces, we uncover much more depth and nuance in human emotion than emotion scientists are used to accounting for. Many of these findings harken back to the observations of Charles Darwin, who described similarities and differences in dozens of emotional behaviors across mammalian species and diverse cultures [23].

Advancing Emotion Science with Artificial Intelligence

The goal of emotion science is not just to characterize emotional behavior. It's to understand how these behaviors shape relationships from the first moments of life [24], guide judgment, decision-making, and memory [25,26], and contribute to our health [27] and well-being [28]. To better understand these processes, it is critical to examine how emotional behavior unfolds in everyday life around the world.

Evidence of this kind is surprisingly lacking in emotion science. It is extremely difficult to capture expressive behavior in real-life contexts that trigger strong emotions. The hand-coding of emotional expressions is time consuming [29]. Moreover, because emotional expressions and the contexts in which they occur are complex, estimating associations between them requires extensive data [30]. For scientists to study emotional behavior in real life at such a scale, we needed new methods.

A few years into grad school, I was approached by tech companies attempting to build AI with empathy who were drawn to my research on semantic spaces of emotion. People were interested in building defenses against bullying, harassment, and depression into social media apps; methods to detect frustration and tiredness into voice assistants; tools to diagnose mental health conditions; and more.

One of the companies I worked with was Google, where I helped establish research efforts focused on the recognition of emotional behaviors. Over a couple of years, I helped develop the most accurate and nuanced algorithm ever built for measuring facial expression. This was a deep neural network (DNN), an algorithm that applied multiple layers of transformation to videos to predict the emotions perceived in facial expressions. My involvement in this effort created an incredible opportunity to study, for the first time, how facial expressions systematically co-vary with specific social contexts around the world.

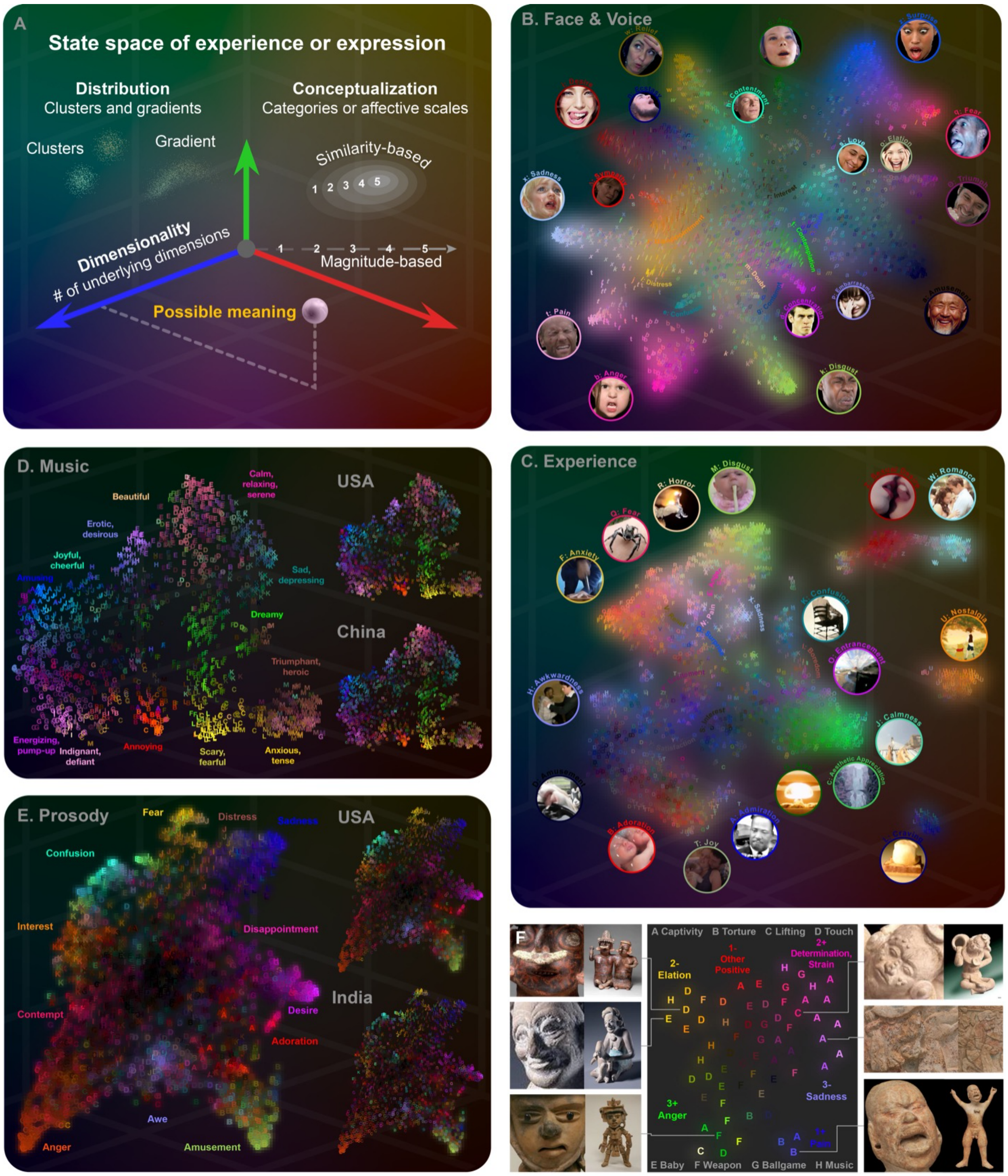

A
State space of experience or expression
Distribution
Clusters and gradients
Conceptualization
Categories or affective scales
Clusters
Gradient
Similarity-based
1 2 3 4 5
Dimensionality
# of underlying dimensions
1 2 3 4 5
Magnitude-based
Possible meaning
B. Face & Voice
C. Experience
D. Music
Calm, relaxing, serene
Beautiful
Erotic, desirous
Joyful, cheerful
Sad, depressing
Dreamy
Triumphant, heroic
Energizing, pump-up
Indignant, defiant
Annoying
Scary, fearful
Anxious, tense
USA
China
E. Prosody
Fear
Distress
Confusion
Interest
Disappointment
Contempt
Desire
Adoration
Awe
Anger
Amusement
USA
India
F
A Captivity
B Torture
C Lifting
D Touch
2+ Determination, Strain
1- Other Positive
2- Elation
3- Sadness
3+ Anger
E Baby
F Weapon
G Ballgame
H Music

Figure 1 (adapted from [17]). **Semantic spaces of experience and expression. A. The semantic space framework.** A semantic space of emotion is described by (1) its dimensionality, or the number of distinct meanings of emotion concepts or emotional expressions within the space; (2) the way that these meanings can be most accurately described in terms of mental states, intentions, or appraisals; and (3) the distribution of emotional experiences or expressions within the space, capturing clusters or blends of states. **B. Semantic space of facial-bodily and vocal expression.** 3,523 expressions are lettered, positioned, and colored according to 28 distinct emotions that people reliably attribute to them (28 in facial expression [18] and 24 in vocal expression [19]). Within the space are gradients in expression between emotions traditionally thought of as discrete, such as "fear" and "surprise." To explore these expressions, see the interactive maps (face: https://s3-us-west-1.amazonaws.com/face28/map.html, voice: https://s3-us-west-1.amazonaws.com/vocs/map.html). **C. Semantic space of emotion evoked by 2,185 brief videos.** At least 27 distinct affective states are reliably captured in reports of emotional experience evoked by video, best conceptualized in terms of emotion concepts such as "fear" [8]. Again, gradients bridge emotion concepts traditionally thought of as discrete, such as "fear" and "surprise". Interactive map: https://s3-us-west-1.amazonaws.com/emogifs/map.html. **D. Semantic space of emotional experience evoked by 1,841 music samples in multiple cultures** [20]. Music samples are positioned and colored according to 13 emotions with which they are reliably associated in both the US and China. Within the space, we find gradients among these states. The similarities in affective response across cultures were most reliably revealed in the use of specific emotion concepts (e.g., "desire", "fear"). Interactive map: https://s3.amazonaws.com/musicemo/map.html. **E. Semantic space of emotion conveyed by prosody in 2,519 lexically identical speech samples.** Across the US and India, at least 12 kinds of emotion are preserved in the recognition of mental states from speech prosody, most reliably revealed in the use of emotion concepts [21]. Interactive map: https://s3-us-west-1.amazonaws.com/venec/map.html. **F. Emotional expression in ancient American art.** From [22]. Ancient American sculpture was found to portray at least five distinct kinds of facial expression that accord, in terms of the emotions they communicate to Westerners, with Western expectations for the emotions that might unfold in the 8 contexts portrayed. Colors of individual faces (letters) are weighted averages of colors assigned to each kind of perceived facial expression. Eight example sculptures are shown. (To explore all 63 sculptures, see online map: https://s3.amazonaws.com/precolumbian/map.html.) Credit, from top left down: (i) Metropolitan Museum of Art 2005.91.12, gift of the Andrall and Joanne Pearson Collection, 2005; (ii) Princeton University Art Museum 2003-26, gift of G. G. Griffin; (iii) Metropolitan Museum of Art 1979.206.578, Michael C. Rockefeller Memorial Collection, Bequest of Nelson A. Rockefeller, 1979; (iv) Kerr Portfolio 342, Jaina Figure, photo by J. Kerr; (v) Kimbell Art Museum, Fort Worth, Texas, AP 1971.07, Presentation of Captives to a Maya Ruler (detail); and (vi) Photograph: Museum of Fine Arts, Boston 1983.288, gift of L. T. Clay.

In our first study using our DNN, we examined emotional behavior at a scale previously unheard of in emotion science: 16 types of facial expression in thousands of contexts found in 6 million videos from 144 countries [16]. We found that each kind of facial expression had distinct associations with a set of contexts that were 70% preserved across 12 world regions. Consistent with these associations, regions varied in how frequently different facial expressions were produced as a function of which contexts were most salient. These results revealed fine-grained patterns in human facial expressions that were well-preserved across the modern world.

In a second study conducted primarily at the University of California, Berkeley, we recorded 45,231 reactions to 2,185 evocative videos (Figure 2), largely in North America, Europe, and Japan, collecting participants' self-reported experiences in English or Japanese and DNN annotations of facial movement [31]. Facial expressions predicted at least 12 dimensions of experience, despite individual variability. We also identify culture-specific display tendencies—many facial movements differed in intensity in Japan compared to the U.S. and Europe, but represented similar experiences. With newfound precision, these results revealed how people experience and express emotion around the world: in high-dimensional, categorical, and complex fashion.

The Future of Big Data and AI in Emotion Science

Unfortunately, there were still important limitations in AI for emotion recognition. The algorithms that had been trained to date, including our DNN, were not well suited to recognize emotional behaviors found more rarely

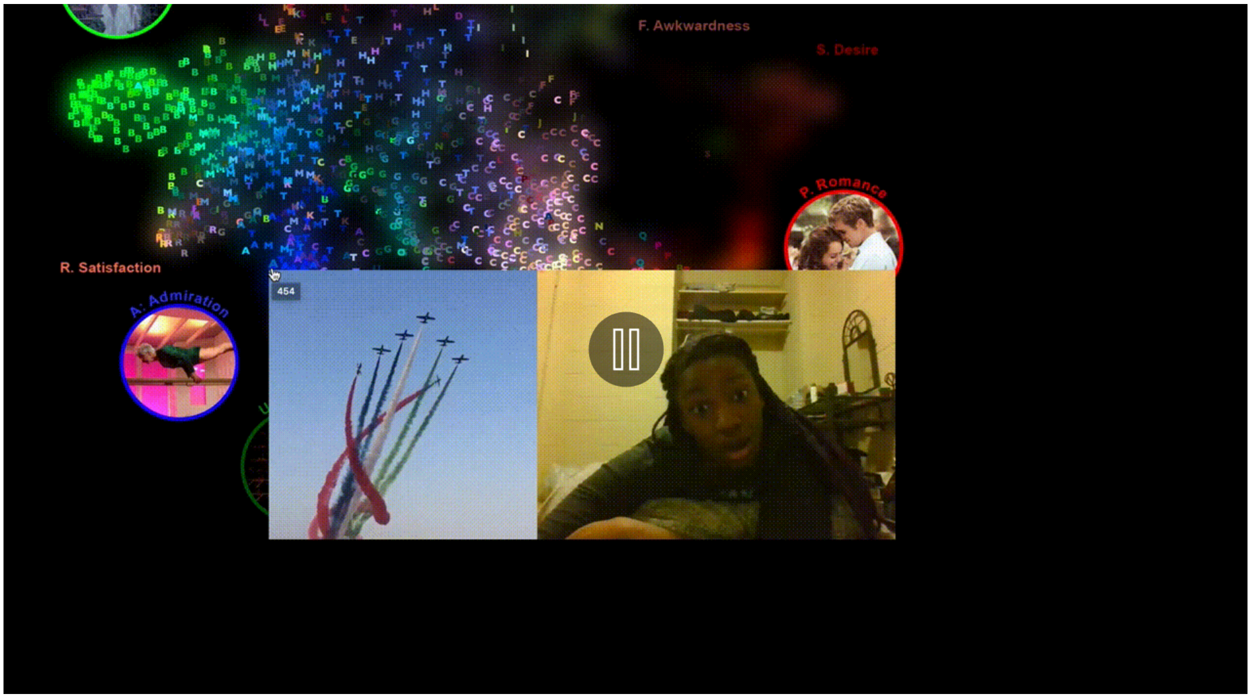


Figure 2. [still frame] Participant reactions from the Berkeley Reactions to Affective Video Elicitors (BRAVE) Dataset. See https://tinyurl.com/yywa7kjf for the full interactive map and [31] for details.

in publicly available data, such as expressions of disgust or fear. They were also confounded by certain perceptual biases—for instance, our DNN labeled anyone wearing sunglasses as expressing pride (as a result, we threw away many of its outputs). Finally, and perhaps most importantly, the DNN trained at Google wasn't publicly available for use by other researchers.

To train more accurate algorithms, we would need large-scale, globally diverse data with a multitude of emotional expressions and contexts. To advance the field of emotion science, we would also need to clear the way to share these AI tools with other researchers.

These are the goals of the private lab I started seven months ago, Hume AI (hume.ai). Hume has been gathering a new kind of rich, globally diverse, psychologically valid emotion data at scale. We have now gathered 3 million self-report and perceptual judgments of 1.5 million human emotional behaviors (Figure 3). With this data, we have trained algorithms that can infer human emotional behavior in the face (https://hume.ai/solutions/facial-expression-model) and voice (https://hume.ai/solutions/vocal-expression-model) with more accuracy and nuance than ever before.

(We are now making our algorithms available free of charge to research groups with compelling data to analyze. If you are interested in using our AI algorithms in your research, please feel free to reach out at hello@hume.ai.)

The Way Forward

With new data and methods, I hope that researchers take the opportunity to look beyond entrenched theories and debates. Rather than look to defend broadly writ notions of universality or variability in human behavior, we can ask, with newfound precision, what exactly *is* consistent in emotional behavior? What *is* it that varies across individuals, demographics, cultures, and contexts? What *do* emotional behaviors, whether universal or culture-specific, indicate about our present beliefs, feelings, and relationships?

I hope that the field of emotion science looks beyond our disagreements, real or imagined—particularly disagreements in terminology, emphasis, or in the questions we feel are most important to answer—toward advances in the substance of our understanding of human emotional behavior. Such advances should build

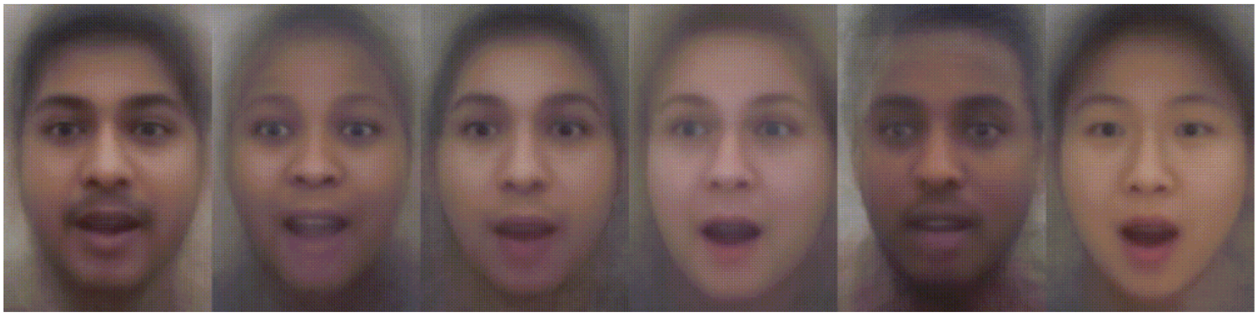


Figure 3. [still frame] Composite representations of 10 emotional expressions in six countries formed by averaging tens of thousands of expressions from real consenting participants based on their self-reported meaning. Hume AI has gathered over 470,000 facial expressions from six countries. From left to right: morphed average expressions in India, South Africa, Venezuela, U.S., Ethiopia, and China. These facial expressions were self-reported as expressing blends of anger, triumph, joy, amusement, love, awe, positive surprise, negative surprise, fear, and pain, shown in that temporal order.

on the fundamental consensus we all share, in and out of academia, that emotional behavior is not meaningless, nor is its meaning straightforward. That a human being is justified in reacting differently to a laugh coming from their living room than to a blood-curdling scream, even if they cannot be sure what it indicates. To draw quantitative insights that do justice to the nuances of human emotion, we will have to come to terms with the fact that what has impended progress in the field is not so much a deficiency in theory, but a deficiency in the tools scientists have had at their disposal to measure emotion behaviors with sufficient precision and at a sufficient scale to draw useful inferences from their occurrence in everyday life.

References

1 Hume, D. (2003) *A treatise of human nature*, Courier Corporation.
2 Russell, J.A. (2003) Core Affect and the Psychological Construction of Emotion. *Psychol. Rev.* 110, 145–172
3 Barrett, L.F. *et al.* (2019) Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements. *Psychol. Sci. Public Interes.* 20, 1–68
4 Gendron, M. *et al.* (2015) Cultural Variation in Emotion Perception Is Real: A Response to Sauter, Eisner, Ekman, and Scott (2015). *Psychol. Sci.* 26, 357–359
5 Sauter, D.A. *et al.* (2015) Emotional Vocalizations Are Recognized Across Cultures Regardless of the Valence of Distractors. *Psychol. Sci.* 26, 354–356
6 Gentsch, K. *et al.* (2015) Appraisals generate specific configurations of facial muscle movements in a gambling task: Evidence for the component process model of emotion. *PLoS One* 10,
7 Barrett, L.F. (2006) Are Emotions Natural Kinds? *Perspect. Psychol. Sci.* 1, 28–58
8 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
9 Shaver, P. *et al.* (1987) Emotion Knowledge: Further Exploration of a Prototype Approach. *J. Pers. Soc. Psychol.* 52, 1061–1086
10 Scherer, K.R. and Wallbott, H.G. (1994) Evidence for Universality and Cultural Variation of Differential Emotion Response Patterning. *J. Pers. Soc. Psychol.* 66, 310–328
11 Barrett, L.F. (2006) Valence is a basic building block of emotional life. *J. Res. Pers.* 40, 35–55
12 Smith, C.A. and Ellsworth, P.C. (1985) Patterns of Cognitive Appraisal in Emotion. *J. Pers. Soc. Psychol.* 48, 813–838
13 Elfenbein, H.A. and Ambady, N. (2002) On the universality and cultural specificity of emotion recognition: A meta-analysis. *Psychol. Bull.* 128, 203–235
14 Cowen, A.S. *et al.* (2019) Mapping the Passions: Moving from Impoverished Models to a High Dimensional Taxonomy of Emotion. *Psychol. Sci. Public Interes.* in press,

15 Cowen, A.S. *et al.* (2019) Mapping the Passions: Toward a High-Dimensional Taxonomy of Emotional Experience and Expression. *Psychol. Sci. Public Interes.* 20, 69–90

16 Cowen, A.S. *et al.* Sixteen facial expressions occur in similar contexts worldwide. *Nature*

17 Cowen, A.S. and Keltner, D. (2021) Semantic space theory: A computational approach to emotion. *Trends Cogn. Sci.*

18 Cowen, A.S. and Keltner, D. (2019) What the Face Displays: Mapping 28 Emotions Conveyed by Naturalistic Expression. *Am. Psychol.* DOI: 10.1037/amp0000488

19 Cowen, A.S. *et al.* (2018) Mapping 24 Emotions Conveyed by Brief Human Vocalization. *Am. Psychol.* 22, 274–276

20 Cowen, A.S. *et al.* (2020) What music makes us feel: Uncovering 13 kinds of emotion evoked by music across cultures. *Proc. Natl. Acad. Sci.*

21 Cowen, A.S. *et al.* (2019) The primacy of categories in the recognition of 12 emotions in speech prosody across two cultures. *Nat. Hum. Behav.* 3, 369–382

22 Cowen, A.S. and Keltner, D. (2020) Universal emotional expressions uncovered in art of the ancient Americas: A computational approach. *Sci. Adv.*

23 Darwin, C. and Darwin, F. (2009) *The expression of the emotions in man and animals*, Oxford University Press.

24 Niedenthal, P.M. and Ric, F. (2017) Expression of Emotion. In *Psychology of Emotion* (2nd edn) (L. F. Barrett, M. Lewis, & J. M. H. J., ed), pp. 98–123, Guilford Press

25 Keltner, D. *et al.* (2014) The Sociocultural Appraisals, Values, and Emotions (SAVE) Framework of Prosociality: Core Processes from Gene to Meme. *Annu. Rev. Psychol.* 65, 425–460

26 Holland, A.C. and Kensinger, E.A. Emotion and autobiographical memory. , *Physics of Life Reviews*, 7. (2010) , 88–131

27 Kok, B.E. *et al.* (2013) How Positive Emotions Build Physical Health: Perceived Positive Social Connections Account for the Upward Spiral Between Positive Emotions and Vagal Tone. *Psychol. Sci.* 24, 1123–1132

28 Diener, E. *et al.* The evolving concept of subjective well-being: the multifaceted nature of happiness. *Advances in Cell Aging and Gerontology*, 15. (2003) , 187–219

29 Cohn, J.F. *et al.* (2007) Observer-based measurement of facial expression with the Facial Action Coding System. In *The handbook of emotion elicitation and assessment* (Coan, J. A. and Allen, J. J. B., eds), pp. 203–221, Oxford University Press

30 Gatsonis, C. and Sampson, A.R. (1989) Multiple Correlation: Exact Power and Sample Size Calculations. *Psychol. Bull.* 106, 516–524

31 Cowen, A.S. *et al.* How emotion is experienced and expressed in multiple cultures: A large-scale experiment. *under Rev.* at <https://psyarxiv.com/gbqtc>

Feature article: Emotion & Feelings

A Science of Emotion without Feelings

Ralph Adolphs

Division of Humanities and Social Science
California Institute of Technology

No, I am not advocating that researchers working on emotion should become cold and unsympathetic people. I am also not arguing that feelings do not accompany emotions, or are irrelevant to emotions. Indeed, I think feelings are important and fascinating phenomena well worth study. I just don't think it is necessary to study them in a science of emotion. More than that, I think it's generally a bad idea to study them if you're studying emotions. They are not the place to begin.

Of course, this all depends on what it is that you are interested in explaining. Many people who study emotions, and in particular those who study emotions in animals, want to explain certain types of behaviors (e.g. so-called "facial expressions" of emotion; see the entries in *Emotion Researcher* on that debate; apparently even mice have them [1]), or certain effects on other cognitive processes (e.g., effects of emotion on memory). If this is indeed your primary interest, I don't see why you need to study feelings. Another way of motivating this conclusion is to imagine intelligent aliens, or perhaps AIs, who for the sake of the thought-experiment do not have conscious experiences or concepts for them, landing on earth and forging a science of the brains and behaviors of the many animals including humans that they find there. My intuition would be that they could do just fine without having feelings, or any other conscious experiences, in their inventory of psychological states. They would need to invent emotions in their psychological science, but their concept of emotions would not be grounded in experiences [2].

If, on the other hand, feelings are your primary interest, then I take it you are interested in explaining (perhaps one type of) conscious experience. Or perhaps the concepts and words we use to describe such conscious experiences. That seems like a different topic. If you are interested in explaining conscious experience, I would suggest you study visual perception perhaps, since it is much easier to link this to well quantified psychophysics. If you are particularly interested in conscious experiences that we might call feelings, I would suggest you study something like pain, which can also be better linked to external stimuli than can the kinds of conscious experiences that often accompany emotions in humans. If you are interested in studying the concepts and words we use to describe conscious experiences of emotions, you're going to be doing a lot of text analysis and NLP and you can't study this topic in animals at all. So, again: these are all interesting topics, but they are quite different topics, and they are secondary to emotion research per se because they assume some fact of the matter about what emotions are supposed to be in the first place (except the last case, where arguably you are not interested in emotions at all, but simply in whatever people say or think or write about emotions using language).

When I offer the above survey of the landscape to most colleagues, they often agree with me. But when I begin to say more about how to study emotion, they keep slipping back into what they had previously acknowledged could be bracketed, and start asking questions about conscious experiences of emotions. It is curious that we seem to have this problem, of continuously wanting to bring in conscious experiences, much more with emotions than with, say, memory, decision-making, or perception. With these latter three kinds of psychological processes, nobody seems to have a big problem studying them in animals or even in brain slices, without worrying about consciousness. To be sure, like emotions, they often are accompanied by conscious experiences. It is also the case that studies of these processes in humans are almost always done in conscious subjects, and even that the subjects are conscious of the stimuli and of the buttons they might press in response to them. So consciousness is around, at least in human participants. But many or most of the studies on perception, memory and decision-making are not primarily about consciousness. It's like saying that people's heart is beating and they are breathing while they are doing the experiment—yes, that helps. But you don't need to be studying cardiology in your memory experiments.

Perhaps there are such things as non-conscious emotions [3], but under most circumstances we are conscious of having emotions, at least if they are of any intensity, even when we may not be conscious of what triggered them. I take this to be an interesting and probably important fact that tells us something about emotions, and something about the respects in which they may differ from perception and memory. Emotions may require an integrated, coordinated response across many different effectors, and that broadcasting of the causal effects of an emotion state may be accompanied by (or indeed constitutive of) conscious experience (at least in the sense of "access consciousness" [4]). But, again, this is subsequent to a clear operationalization and study of emotion. Simply put: you need to know something about what emotions are before can study what conscious experiences of emotions might be.

Joe LeDoux's View

Let me briefly say something further about emotions vs. feelings by commenting on the positions others have advocated (and all misrepresentations will of course be my error). My colleague David Anderson and I recently co-authored a book, *The Neuroscience of Emotion* [5], that is in many ways a reaction to the recent views of Joe LeDoux. This is because LeDoux argues that emotions are feelings, and that work in animals that has purported to be about emotion is therefore not about emotion (but instead about what he calls survival circuits). It may be that the disagreement is just semantic [6], in which case there is little of interest to argue about; I'm assuming, for the sake of the argument here that it's more substantive.

There is a class of psychological processes, many of them historically entrenched, that are functionally defined without appeal to conscious experience and that figure prominently also in emotional behaviors: variables like motivation, reward, drive. Reward or value is perhaps worked out in the most detail in computational

Professor Ralph Adolphs

models of decision-making. Some of those same decision-making frameworks have also been used to operationalize emotions [7]. There is a lot to be worked out in the details, but this seems to me to be on the right track: emotions are functional states, and at least a part of their function can be formalized into models that describe the control of behavior through well-studied systems (Pavlovian, instrumental model-based and model-free, etc.). Some very specific proposals have been made here; an intriguing one is that emotions (or more particularly moods) should be thought of as "momentum" in the accumulation of a history of unexpected outcomes [8]. While powerful, these models also make it tricky to disentangle some variables from others [9] and in the case of emotion it is unclear what the scientific, let alone the metaphysical, status of emotion in a computational model is: one approach would be to say that it has simply been reduced to a set of other variables and we can get rid of the term "emotion". Perhaps this is LeDoux's view also — the only thing that can rescue emotion, as distinct from just a collection of other processes, might be feelings.

The most detailed functional and neurobiological models for understanding emotion come into play in the case of fear. Here again LeDoux's treatment is informative: he is at pains to distinguish non-conscious processes that control fear-related behaviors from those cognitive processes contributing to fear behaviors that are accompanied by (or perhaps constitutive of) conscious feelings of fear (which according to him depend on re-representations of the contents of working memory) [10]. By contrast, one of my colleagues here at Caltech, Dean Mobbs, has also written about such detailed functional proposals for understanding fear, but has no problem using the word "fear" [11]. Further viewpoints have been summarized in some recent fun debates [12]. These diverse views have real consequences for how we do science. For instance, Joe LeDoux and Danny Pine have used LeDoux's distinction between survival circuits and circuits for the conscious experience of fear to argue that animal models for anxiolytic drugs are invalid; presumably they would say the same for antidepressants. While I am with people like Michael Fanselow in disagreeing with this view [13], the intuition makes sense: if all that a drug accomplishes is to get you to approach anxiety-evoking situations, or to get out of bed in the morning, but when asked you say you feel just as anxious or depressed, the drug has clearly not achieved one of its main purposes (although one could argue that it has indeed achieved some useful purpose even under this scenario).

Lisa Barrett's View

For Lisa Barrett (and many others), emotion is all about conscious experience. Following work by people like Jim Russell, the basic idea is that there is an essential core to the conscious experience of emotion ("core affect" [14]), typically with something like two dimensions of valence and arousal. Various ingredients are then added to this core affect (consciousness of the eliciting circumstances, the context, the consequences, etc.). I think in her later writings Barrett would allow that not all the ingredients of an emotion episode need to figure in our conscious experiences of them (they are assembled from many different brain systems, not all of whose operation is necessarily tied to consciousness [15]), but I take it that all emotions require core affect as one necessary ingredient, so at least you are conscious of that when you have an emotion.

I have had a number of stimulating arguments with Barrett about emotion [16-19], and her views have evolved over time. Her current view is more focused on concepts as such, and she has developed a specific notion of that term that is very neurobiologically based. In a nutshell, I take her to propose that emotions are states of the brain that are assembled on the fly across many different brain systems as required by a particular, context-dependent circumstance — a "conceptual act" [15]. A key feature of this view is that emotions are not biological (or any other natural) kinds — they are entirely derivative to other processes in the brain and the only thing that qualifies them as emotions is our (socially shared) concept of what an emotion is (or, more specifically, what a particularly type of emotion, like fear, is).

So both LeDoux and Barrett think emotions require feelings. In LeDoux's case that leads him not to use the word "emotion," whereas in Barrett's case it leads her to include core affect as one necessary component of emotions. But the most interesting aspect of both views is that they seem to want to get rid of what I would take to be the most salient aspect of emotions: what it is they accomplish functionally (and hence presumably what it is that guided their evolution). Darwin had this in his concept of "serviceable associated habits" [20], Herbert Simon had it in the concept of an interrupt mechanism [21], and I think the layperson's concept of an emotion has it as well. Emotions are elicited by particular challenges in the environment, and their function is to help cope with those challenges by engaging a host of coordinated cognitive and behavioral responses. Indeed, this involves many systems (agreeing with Barrett), and indeed much of this processing is not necessarily accompanied by conscious experience (agreeing with LeDoux). The fact that their function engages many other components does not reduce emotions to these other components, because they cohere— they are packages at a particular level of behavioral control that resides intermediate between reflexes and deliberated behavior [22].

Situating Emotion Science as a Science of the Mind

There are three main ways people think about psychology, and about the mind. The layperson tends to think of consciousness. On this view, a psychologist studies aspects of conscious experience or entities defined by relation to conscious experience (the "subconscious"). Many psychologists and neuroscientists instead think about psychology, and about the mind, in terms of the attributions that we make of other people. Experiments asking people to judge emotions from facial expressions, "theory of mind" tasks, and so forth all study this human ability. Developmental psychology studies how it emerges in infancy and how it might be compromised in autism; comparative psychology has been working hard to describe how it might be present in great apes.

Finally, the third way of thinking about psychology and the mind is the one I am discussing here. No doubt, laypeople think of emotions in terms of feelings; and no doubt there is a rich story to tell about how we are able to infer emotions in other people from observing them. But I am asking the question, How should scientists think of emotions? I think they should think of emotions as latent variables, much like should be the case for all other psychological variables. They are variables that, taken together, provide causal explanations of behavior (not of conscious experience), realized in the brain.

While it is true that we have invented much about our own minds and those of others in our folk psychology, I don't see why this needs to discredit scientific psychology. Contrary to some views [23], I don't think the mind is flat: there is an architecture of cognition, and much of the task of psychology should be to figure out what to put into the boxes and where to draw the arrows. If one accepts that we are currently very far from having this all figured out, I think it gives a new stance on emotion science also. We don't really know what emotions are, the various theories about them notwithstanding. There is no quick "essence," like their phenomenal "feel", that will solve the problem. They will have to be accommodated into a mental architecture, and it will remain to be seen whether "emotion" as a generic category, let alone specific categories or dimensions of emotions, will eventually correspond to the ones we currently have available in our folk psychology. No doubt revision will be required. But the starting point should be phylogenetically continuous study of behaviors and abilities, whose explanation will require a mind composed of many types of latent variables. Emotion will be one of those variables, and I think we have rough idea of the criteria required for inferring it. Conscious experiences, or reports of them, may well be one of those criteria in humans under many circumstances.

References

1. Dolensek, N., et al., Facial expressions of emotion states and their neuronal correlates in mice. Science, 2020. **368**: p. 89-94.
2. Adolphs, R. and D. Andler, Investigating emotions as functional states distinct from feelings. Emotion Review, 2017. **in press**.
3. Winkielman, P. and K.C. Berridge, *Unconscious emotions.* Current Directions in Psychological Science, 2004. **13**: p. 120-123.
4. Block, N., On a confusion about a function of consciousness. Behavioral and Brain Sciences, 1995. **18**: p. 227-287.
5. Adolphs, R. and D.J. Anderson, *The Neurobiology of Emotion: A New Synthesis*. 2018, Princeton, NJ: Princeton University Press.
6. LeDoux, J., Semantics, surplus meaning, and the science of fear. Trends Cogn Sci, 2017. **21**: p. 303-306.
7. Bach, D.R. and P. Dayan, Algorithms for survival: a comparative perspective on emotions. Nature Reviews Neuroscience, 2017. **18**: p. 311-319.
8. Eldar, E., et al., *Mood as representation of momentum.* Trends in Cognitive Sciences, 2016. **20**: p. 15-24.
9. O'Doherty, J.P., *The problem with value.* Neuroscience and Biobehavioral Reviews, 2014. **43**: p. 259-268.
10. LeDoux, J. and N. Daw, Surviving threats: neural circuit and computational implications of a new taxonomy of defensive behavior. Nature Reviews Neuroscience, 2018. **19**: p. 269-282.
11. Mobbs, D., et al., *The ecology of human fear: survival optimization and the nervous system.* Frontiers in Neuroscience: Evolutionary Psychology and Neuroscience, 2015. **9**: p. 55.
12. Mobbs, D., et al., Viewpoint: approaches to defining and investigating fear. Nature Neuroscience, 2019. **22**: p. 1205-1216.
13. Fanselow, M. and R. Pennington, A return to the psychiatric dark ages with a two-system framework for fear. Behav Res Therapy, 2018. **100**: p. 24-29.
14. Russell, J.A., Core affect and the psychological construction of emotion.

Psychological Review, 2003. **110**: p. 145-172.
15. Barrett, L.F., How Emotions are Made: The Secret Life of the Brain. 2017, New York: Houghton Mifflin Harcourt.
16. Adolphs, R., Reply to Barrett: affective neuroscience needs objective criteria for emotions. Social Cognitive and Affective Neuroscience, 2017. **12**: p. 32-33.
17. Adolphs, R., How should neuroscience study emotions? By distinguishing emotion states, concepts, and experiences. Social Cognitive and Affective Neuroscience, 2017. **12**: p. 24-31.
18. Barrett, L.F., Functionalism cannot save the classical view of emotion. Social Cognitive and Affective Neuroscience, 2017. **12**: p. 34-36.
19. Barrett, L.F., The theory of constructed emotion: an active inference account of interoception and categorization. Social Cognitive and Affective Neuroscience, 2017. **12**: p. 17-23.
20. Darwin, C., *The Expression of the Emotions in Man and Animals*. 1872/1965, Chicago: University of Chicago Press. 372.
21. Simon, H.A., Motivational and emotional controls of cognition. Psychological Review, 1967. **74**: p. 29-39.
22. Adolphs, R., *Emotions are functional states that cause feelings and behavior*, in *The Nature of Emotion*, A. Fox, et al., Editors. 2019, Oxford University Press: New York. p. 6-10.
23. Chater, N., *The Mind is Flat*. 2018, New Haven: Yale University Press.

Feature article: Emotion & Feelings

Emotion Recognition ≠ Emotion Understanding: Challenges Confronting the Field of Affective Computing[1]

Jonathan Gratch

USC Institute for Creative Technologies
Departments of Computer Science, Psychology, and Media Arts and Practice
University of Southern California

Affect recognition: a useful or dangerous tool?

Many assume that a person's emotional state can be accurately inferred by surface cues such as facial expressions and voice quality, or through physiological signals such as skin conductance or heart rate variability. Indeed, this assumption is reflected in many commercial "affect recognition" tools. For example, Table 1 illustrates some of the findings of a recent survey performed by the Association for the Advancement of Affective Computing (AAAC) on how commercial affect recognition is marketed.

Affective science has long debated the linkages between affective states, emotional expressions and self-reported feelings. The quotes in Table 1 come from products that adopt what I have come to call "context-ignorant" emotion recognition. For example, Affectiva claims to tell you "how your customers and viewers feel" solely from video of a human face without regard to the physical, social, or cultural context in which the face was captured. Digging deeper into their documentation, the company clarifies that the algorithm mimics what third-party observers, also ignorant of the context, would say the face is showing. Early versions of Ekman's basic emotion theory argued such context-ignorant inferences are meaningful. In this view, emotional state, emotional expressions, and self-reported feelings are tightly linked and essentially act as a single circuit (Ekman, 1992). The implication is that the recognition of behavior in one component is highly diagnostic of other components, and in particular, that feelings can be readily predicted from surface cues such as facial expressions. However, the emerging consensus within affective science is that components of emotion are loosely connected, and expressions are highly-dependent on the context, shaped by social norms and regulation, and many expressions are completely disconnected from underlying feelings – e.g., arising from deliberate communicative acts or even the articulatory movements required to produce speech (e.g. see Barrett, Adolphs, Marsella, Martinez, & Pollak, 2019; Crivelli & Fridlund, 2018; Scarantino, 2017). Even contemporary proponents of Ekman emphasize the context-specificity of emotional expressions, even when arguing for emotion's universality (see Cowen et al., 2020). Thus, automatically recognizing emotional state or felt emotions from decontextualized signals is a difficult, if not quixotic enterprise.

This seeming disconnect between the claims of many affect recognition companies and the science of affective signals has raised alarm in some circles. One prominent AI watchdog identified affect recognition as their #1 societal concern, recommending that "regulators should ban the use of affect recognition in important decisions that impact people's lives and access to opportunities" (Crawford et al., 2019, pg. 6). In fact, I was quoted, misleadingly, as evidence for such a ban (pg. 51). But this reaction also lacks context. For example, given that "affect" encompasses moods and chronic states such as depression and PTSD, a literal interpretation would prevent doctors from using validated scales to identify patients at risk of mental illness (in that such scales constitute a primitive form of affect recognition "technology"). As I will highlight, affective signals can meaningfully inform decision-making as long as appropriate care is taken in their measurement and interpretation.

[1] This article is based on a recent webinar presented at ISRE. This article benefited from contributions from USC's Affective Computing Group but especially Su Lei and Kelsie Lam

Professor Jonathan Gratch

One of the key problems with these tools comes down to terminology. "Affect recognition" overstates the capabilities of these systems, as most focus on *expression* recognition. On the other hand, "affect recognition" understates the utility of these methods, as expression contain important information even if this information is disconnected from underlying feelings and emotional state. Unfortunately, such tools rarely come with the appropriate disclaimers or concrete advice on how to avoid their misuse. In this article, I will give a broad overview on how these methods work, the many ways they can yield misleading results, and the emerging engineering advances that address the most common points of failure. These failures broadly fall into two categories: problems in *recognizing* affective expressions, and problems in *understanding* what can be concluded from these expressions (even if they have nothing to do with emotion). In illustrating these challenges, I will focus on facial expressions but similar issues arise in other modalities such as affective speech or text analysis.

Table 1
Highlights from a survey on the marketing of Affect Recognition software

Company	Advertised Applications	Quotes from Company Website
ADOREBOARD	Market research, human resources	Adoreboard emotics tells you what your customers and employees actually feel
:) Affectiva	Market research, Media analytics, transportation	Understand how your customers and viewers feel when they can't or won't say so themselves
audEERING™	Customer service	A few seconds of speech are enough to determine the emotional state of the caller, even while they are still in the queue.
BEHAVIORAL SIGNALS	Customer service	BSP algorithms can determine what behaviors and emotions caused a reaction and when there were turning points.
Empath	Customer service, health care	We can identify your emotion whatever language you speak.
motive	Market research, Customer service, human resources	Understand the emotion and intent that's driving your employees and customers — across all channels of data in real-time — and what it means for your business
Noldus	Market research, UX, human factors	The [recognized] valence indicates whether the emotional state of the subject is positive or negative

Challenges facing expression recognition

Most of the algorithms advertised as “emotion recognition” or “affect recognition” are designed to recognize expressions, not “emotion” or “feeling”. These algorithms use machine learning to map from some input (e.g., an image or video of a face) to some label. Some labels are clearly about facial expressions. For example, OpenFace and AFFDEX map an image of a face to a vector of Facial Action Units. The algorithm takes a large database of images that were labeled by trained coders and mimics the skill of these experts. But the majority of these algorithms produce other labels, such as Ekman’s seven basic emotions or other affective or mental states such as frustration, confusion or fatigue. To understand the meaning of these labels, we have to look at the details (which are often, frustratingly, not provided). Many algorithms are trained to recognize prototypical expressions generated by an actor. For example, Emoreader uses a” training database is comprised of 72,800 faces from 3,092 actors.” Other algorithms rely on third-party judgments. For example, Affectiva used trained human coders to look at images and label them for the presence or absence of disgust (McDuff, Kaliouby, Cohn, & Picard, 2015). Either way, these labels are best seen as expressions as the underlying feelings or emotional state is unknown (in the case of third-party observations) or unrelated (in the case of acted expressions).

Recognizing expressions is not the same as recognizing feelings or emotional state, but classifying facial signals is still valuable if classified correctly. Unfortunately, the measurement context can systematically bias the output of these methods. Figure 1 illustrates some of common problems we have encountered in our own research based on superficial details of the recording environment or participant appearance. These details include:

- *Scene complexity:* One of the first steps in expression recognition is to find the face and Figure 1a highlights that face detection can become confused when the scene is complex or other faces lurk in the background. In this case, the face of Abraham Lincoln appears more interesting than the person in the foreground. Automatic expression recognition works best when a person sits alone with a neutral background, but many studies deviate from this ideal, either because of a lack of knowledge of this limitation or other pragmatic concerns. For example, video data collected “in the wild” typically involves high scene complexity. This can undermine the accuracy of results. *Head orientation:* Part of the issue with the image is 1a is that the person is looking down. Indeed, head orientation is a major concern

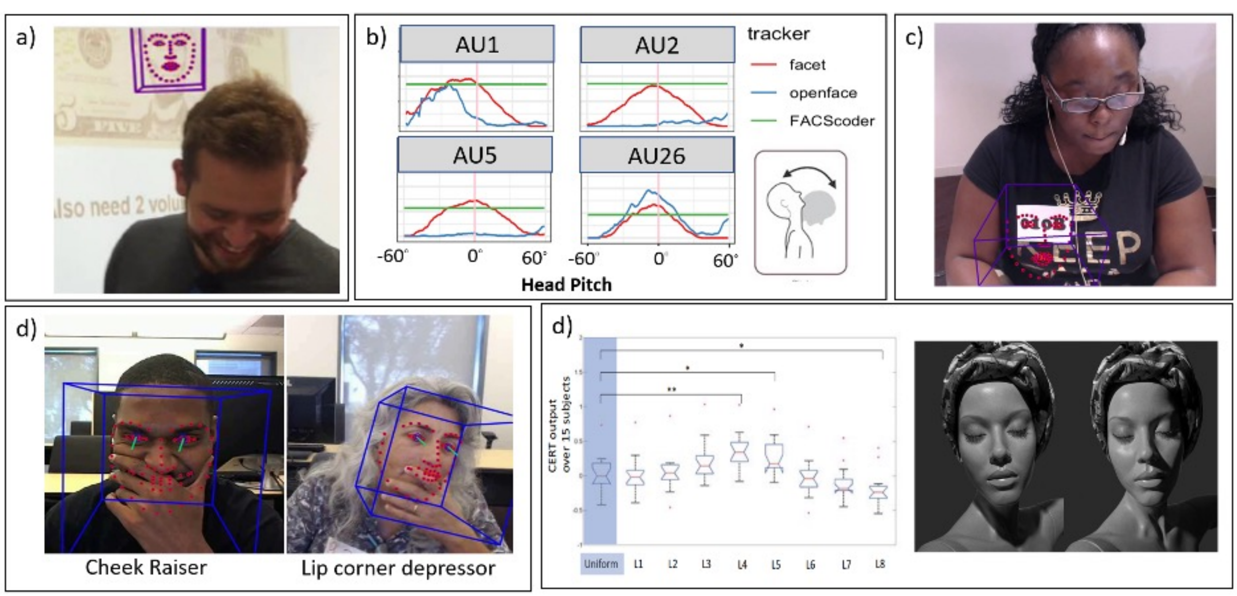


Figure 1. An illustration of several errors in expression recognition.

for expression recognition (see Kappas, Hess, Barr, & Kleck, 1994). Many face detectors loose track of the face when the head rotates more than ten or twenty degrees off axis. But perhaps more problematic, small but systematic changes in head orientation can bias the results of an experiment. Figure 1b illustrates how recognition of a prototypical surprise expression (i.e., AU1 + AU2 + AU26) changes as a function of head pitch (the graphs show expert FACS coder ratings contrasted against two expression recognizers). This highlights that the placement of a camera (e.g., above or below a monitor) or even the height of a participant can create confounds in a study (e.g., observed differences in expressions between men or women might simply reflect that men are taller on average).

- *Skin tone:* Several researchers have highlighted that face and expression recognition algorithms exhibit racial bias. Figure 1c illustrates an example where the face detector is fixated on an apparent white face in this image (the letters "0 1 0" on this black participant's id number). In our own work, we have found that these effects can introduce systematic biases into machine learning algorithms that use this input. For example, we developed a technique to classify workers as engaged or disengaged with a task based on their expression. Black participants were more frequently misclassified as disengaged workers than white participants.
- *Occlusions:* People frequently touch their face and this can undermine the accuracy of expression recognition. Figure 1d shows some examples of how these facial occlusions can change what is reported. We discovered these images by looking for cases in one of our corpora where specific action units showed high activation. For example, in the woman on the left, the space between her fingers are misrecognized a mouth. Such effects could systematically bias experimental results. For example, self-touching has been argued to be an indicator of stress (Ekman & Friesen, 1974). If true, automatic findings on the association of facial expressions and stress must control for facial touching.
- *Lighting:* Photographers and cinematographers have long understood that lighting changes perceptions of emotion. Even before photography, Japanese Noh actors used this effect to convey very different expression from a static mask (Kawai, Miyata, Nishimura, & Okanoya, 2013). Unfortunately, this phenomenon also shapes the output of expression recognition techniques. Figure 1d shows the reported activation of AU4 (brow lowerer) using the FACET commercial expression recognizer under a variety of lighting conditions (Stratou, Ghosh, Debevec, & Morency, 2012). This highlights that some experimental findings on expressions could be influenced by lighting artefacts. For example, studies on the effect of time-of-day on facial expressions might report spurious correlations if video is collected in a room with an open shade.
- *Non-expressive sources of facial motion:* The face moves for many reasons including the articulatory movements required for speech, chewing, swallowing and breathing. Third-party observers would tend to filter out these movements when judging facial expressions but many algorithms don't. Most automatic analysis of facial expressions are performed on individual images and an overall expression is calculated by aggregate measures such as mean, max or velocity over some time window. At the level of individual frames, the mouth shape from smiling and the mouth shape from saying the word "cheese" appear identical. As a consequence, an experimental finding that depressed individuals show flattened affect in a study might simply reflect that they spoke less, unless analysis was restricted to areas of non-speech or if the amount of speech was statistically controlled.

Affective computing researchers are well aware of these concerns and engineering solutions are in the works. Facial detection accuracy is quickly advancing as more and more research focuses on collection in the wild. Problems with head orientation are being addressed by training algorithms on 3D data to better account for how the appearance changes when a 3D face is projected on a 2D image (Jeni, Cohn, & Kanade, 2015). Problems with skin tone are often traced to the lack of data collected on minority groups and many of the more egregious problems have now been corrected (Raji & Buolamwini, 2019) though many challenges remain. Researchers are developing "occlusion-aware" facial expression methods that can minimize the impact of self-touching (Li, Zeng, Shan, & Chen, 2019) and algorithms can automatically infer how a scene is illuminated, making it possible to correct for the impact of lighting on facial appearance (Xie, Zheng, Lai, Yuen, & Suen, 2011). Algorithms are also being developed to filter out non-expressive sources of facial motion, such as speech (Kim & Mower Provost, 2014).

Until these methods are perfected, researchers can minimize the impact of these issues by taking care during expression measurement. This can include collecting data in windowless rooms with uncluttered backgrounds and standardized lighting and camera locations. Participants can be cautioned to avoid self-touching. To avoid the impact of head orientation, consider presenting stimuli on a computer screen and avoid secondary tasks that lead people to look away from the screen. For example, studying expressions during team tasks will introduce fewer expression artefacts if the team works on Zoom rather than in-person (as eye contact can be maintained without large head rotations). When these factors cannot be eliminated, confidence in findings can be enhanced by ensuring these factors do not vary systematically with experimental condition. This can be addressed statistically. For example, head orientation or lighting can be measured and controlled for in any analysis.

Challenges facing expression understanding

Assuming we can accurately recognize facial expressions, surely these can support useful inferences. For example, Taco Bell could show customers a video of their new "Loaded Nacho Taco" and predict if people will buy it based on if they smile or show disgust. As illustrated in Table 1, market research is a heavily promoted application of this technology and some research suggests it can be effective. A large study supported by Affectiva showed that expressive reactions to product advertisements can predict if people like an advertisement and, to a somewhat lesser extent, intentions to purchase the product (McDuff et al., 2015). While an important validation of the use of automatic expression recognition, the study also emphasizes the importance of context in shaping these predictions. Rather than using summary statistics (e.g., checking if people smiled more during an ad), expressions were fed, along with several contextual features, into machine learning algorithms. Contextual features included the type of product (e.g., pet care, food, or candy) and country of origin. Ads were also classified by the type of emotions they were designed to evoke (e.g., amusement, inspiration, sentimentality). Findings showed that context mattered. Ad liking was best predicted when focusing on ads designed to be amusing and when the product category and country of origin were used as inputs to the learn algorithm. More broadly, the findings are specific to a certain context: reaction to TV ads by participants from four affluent western countries (US, UK, Germany, France). It remains unclear if the models learned from this collection of videos would generalize to other contexts (e.g., other products, other countries, or videos drawn at a different period in time). It should also be noted that emotion reactions were extremely rare (only 17% of the frames showed an expression). Thus, it is not possible to infer if a specific individual would like the ad or product. Rather, predictions must be made from a panel of members watching the identical video. Using such panels is common in market research but many of the applications in Table 1 claim to provide information about individuals (e.g., identifying that a particular customer is angry), a claim that needs to be taken with some skepticism.

Moving from expression recognition to expression understanding must account for context. Affective science has documented an array of contextual factors that shape the association between emotion, expression and feeling, but these factors are rarely highlighted by the marketing materials of commercial affective recognition products. The social context around

Figure 2: Facial reactions during one round of an iterated prisoner's dilemma game.

an interaction exerts a particularly strong influence:

- *Alone vs. Social*: The presence of an audience clearly shapes expressions. When alone, the frequency and intensity of expressions tend to be significantly reduced, whereas self-reported feelings appear less impacted by the presence of others, at least for tasks that don't rely on a social component (e.g., Fridlund, 1991). Even the presence of an experimenter in the room can influence findings and data collection "in-the-wild" rarely event attempts to control this confounding factor.
- *Friends vs. strangers:* People tend to be more expressive in the presence of friends compared with strangers and some research suggests the match between expressions and feelings is stronger in the presence of friends (Gratch, Cheng, Marsella, & Boberg, 2013; Hess, Banse, & Kappas, 1995). Tracking the relationship between individuals should [?] expression understanding.
- *Impression management*: In social settings, people are often concerned about politeness or maintaining a good impression. In customer service jobs, employees are paid to convey a particular emotion (Hochschild, 2003). Although some research seeks to distinguish "authentic expressions" from impression management attempts (Ambadar, Cohn, & Reed, 2009), algorithms might benefit more from controlling for impression management demands.
- *Co-construction:* The above-mentioned effects occur even in the mere presence of others, but many socially situations involve the back-and-forth display of emotion between two or more people. Rather than reflecting an individual's feelings, expressions may reflect momentary adjustments to expressions displayed by their interaction partner (Parkinson, 2009). For example, Figure 2 illustrates a common pattern in data we've collected in competitive games (see Lei & Gratch, 2019). Here a woman conveys sadness after losing a round, which is immediately mimicked by her partner, leading both players to smile. Some expressions may be better understood as analogous to words in a conversation than conveying some underlying emotional state (see also Scarantino, 2017).
- *Social power*: The power dynamics between individuals shapes what people express. For example, low-power individuals tend to be more expressive and more willing to engage in mimicry than high-power individuals (e.g., Tiedens & Fragale, 2003). Representing or automatically recognizing power relationship (as in Hung, Huang, Friedland, & Gatica-Perez, 2011) could benefit expression understanding.
- *Social goals and motives*: Expressions often reflect what people are trying to accomplish in a situation (Crivelli & Fridlund, 2018) and these motives will differ across individuals. For example, in a negotiation, some individuals may be focused on material outcomes (winning) whereas others are focused on maintaining the relationship. Thus, the same expression (smile) may reflect quite different meanings across different individuals (de Melo, Carnevale, Read, & Gratch, 2014).

This is not to say that algorithms cannot infer meaningful information about a situation from patterns of facial expressions. As discussed above, algorithms can infer intentions to buy a product with some accuracy from facial reactions to an advertisement. Algorithms have also been shown to detect risk of depression or suicide form expressions produced in a clinical interview (Cummins et al., 2015). The reason these algorithms are successful is that they control for context: i.e., they are trained on data collected in a specific context and applied to similar situations. Algorithms that avoid this care should be viewed with suspicion.

Towards Knowledge-based affect understanding

I have argued that context-ignorant affect recognition is quixotic enterprise, likely doomed to fail. But that is not a problem with the technology. It is a problem with its use and marketing. People don't make sense of expressions in the absence of context. For example, in one of our recent studies on competitive games, peoples inferences about the other players emotions were better predicted by context than their opponent's expressions (Hoegen, Gratch, Parkinson, & Shore, 2019). Affect recognition methods are successful when they control for the context. But automated methods could be even more powerful if they explicitly reason about social situations. Although this article has focused on methods that infer emotion from shallow signals (e.g., facial expressions), another thread of affective computing research has focused on how to predict emotion from deep representations of situations. Much of this work is based on appraisal theory. Algorithms reason about how emotions arise from an appraisal of how an individual's goals are impacted by events (see Marsella, Gratch, & Petta, 2010 for a review). For example, if a person is known to have a goal of winning and they lose, and the situation does not afford opportunities to reverse the loss, we might reasonably conclude they are sad, even in the absence of obvious expressions. Combining these threads (knowledge-based reasoning with expression recognition) could yield even more robust and accurate inferences (see Yongsatianchot & Marsella, 2016 for one example).

In sum, calls to ban affect recognition are misguided and distract attention from the real issues. Like most human innovations, they can provide clear benefits when used appropriately but clear harms through ill-informed use. Companies and researchers have a responsibility to educate consumers on the constraints and limitations of the technology. This is perhaps even more important with affect recognition as everyone feels they are an expert on human emotion. As is clear from Table 1, we still have a long way to go.

References

Ambadar, Z., Cohn, J. F., & Reed, L. I. (2009). All smiles are not created equal: Morphology and timing of smiles perceived as amused, polite, and embarrassed/nervous. *Journal of nonverbal behavior, 33*(1), 17-34.

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional Expressions Reconsidered: Challenges to Inferring Emotion From Human Facial Movements. *Psychological Science in the Public Interest, 20*(1), 1-68. doi:10.1177/1529100619832930

Cowen, A. S., Keltner, D., Schroff, F., Jou, B., Adam, H., & Prasad, G. (2020). Sixteen facial expressions occur in similar contexts worldwide. *Nature*. doi:10.1038/s41586-020-3037-7

Crawford, K., Dobbe, R., Dryer, T., Fried, G., Green, B., Kaziunas, E., . . . Sánchez, A. N. (2019). AI now 2019 report. *New York, NY: AI Now Institute*.

Crivelli, C., & Fridlund, A. J. (2018). Facial Displays Are Tools for Social Influence. *Trends in Cognitive Sciences, 22*(5), 388-399. doi:https://doi.org/10.1016/j.tics.2018.02.006

Cummins, N., Scherer, S., Krajewski, J., Schnieder, S., Epps, J., & Quatieri, T. F. (2015). A review of depression and suicide risk assessment using speech analysis. *Speech Communication, 71*, 10-49. doi:https://doi.org/10.1016/j.specom.2015.03.004

de Melo, C., Carnevale, P. J., Read, S. J., & Gratch, J. (2014). Reading people's minds from emotion expressions in interdependent decision making. *Journal of Personality and Social Psychology, 106*(1), 73-88. doi:10.1037/a0034251

Ekman, P. (1992). An argument for basic emotions. *Cognition and Emotion, 6*(3-4), 169-200.

Ekman, P., & Friesen, W. V. (1974). Nonverbal behavior and psychopathology. *The psychology of depression: Contemporary theory and research*, 3-31.

Fridlund, A. J. (1991). Sociality of social smiling: Potentiation by an implicit audience. *Journal of Personality and Social Psychology, 60*, 229-240.

Gratch, J., Cheng, L., Marsella, S., & Boberg, J. (2013). *Felt emotion and social context determine the intensity of smiles in a competitive video game*. Paper presented at the 10th IEEE International Conference on Automatic Face and Gesture Recognition, Shanghai, China.

Hess, U., Banse, R., & Kappas, A. (1995). The intensity of facial expression is determined by underlying affective state and social situation. *Journal of Personality and Social Psychology, 69*(2), 280-288. doi:10.1037/0022-3514.69.2.280

Hochschild, A. R. (2003). *The managed heart: Commercialization of human feeling*: Univ of California Press.

Hoegen, R., Gratch, J., Parkinson, B., & Shore, D. (2019). *Signals of Emotion Regulation in a Social Dilemma: Detection from Face and Context*. Paper presented at the 8th International Conference on. Affective Computing & Intelligent Interaction Cambridge, UK.

Hung, H., Huang, Y., Friedland, G., & Gatica-Perez, D. (2011). Estimating Dominance in Multi-Party Meetings Using Speaker Diarization. *IEEE Transactions on Audio, Speech, and Language Processing, 19*(4), 847-860. doi:10.1109/TASL.2010.2066267

Jeni, L. A., Cohn, J. F., & Kanade, T. (2015, 4-8 May 2015). *Dense 3D face alignment from 2D videos in real-time.* Paper presented at the 2015 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG).

Kappas, A., Hess, U., Barr, C. L., & Kleck, R. E. (1994). Angle of regard: The effect of vertical viewing angle on the perception of facial expressions. *Journal of nonverbal behavior, 18*(4), 263-280. doi:10.1007/bf02172289

Kawai, N., Miyata, H., Nishimura, R., & Okanoya, K. (2013). Shadows alter facial expressions of Noh masks. *PLoS One, 8*(8), e71389.

Kim, Y., & Mower Provost, E. (2014). *Say cheese vs. smile: Reducing speech-related variability for facial emotion recognition.* Paper presented at the Proceedings of the 22nd ACM international conference on Multimedia.

Lei, S., & Gratch, J. (2019, 3-6 Sept. 2019). *Smiles Signal Surprise in a Social Dilemma.* Paper presented at the 2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII).

Li, Y., Zeng, J., Shan, S., & Chen, X. (2019). Occlusion Aware Facial Expression Recognition Using CNN With Attention Mechanism. *IEEE Transactions on Image Processing, 28*(5), 2439-2450. doi:10.1109/TIP.2018.2886767

Marsella, S., Gratch, J., & Petta, P. (2010). Computational Models of Emotion. In K. R. Scherer, T. Bänziger, & E. Roesch (Eds.), *A blueprint for affective computing: A sourcebook and manual* (pp. 21-46). New York: Oxford University Press.

McDuff, D., Kaliouby, R. E., Cohn, J. F., & Picard, R. W. (2015). Predicting Ad Liking and Purchase Intent: Large-Scale Analysis of Facial Responses to Ads. *IEEE Transactions on Affective Computing, 6*(3), 223-235. doi:10.1109/TAFFC.2014.2384198

Parkinson, B. (2009). What holds emotions together? Meaning and response coordination. *Cognitive Systems Research, 10*, 31-47.

Raji, I. D., & Buolamwini, J. (2019). *Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products*. Paper presented at the Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and

Society, Honolulu, HI, USA. https://doi.org/10.1145/3306618.3314244

Scarantino, A. (2017). How to Do Things with Emotional Expressions: The Theory of Affective Pragmatics. *Psychological Inquiry, 28*(2-3), 165-185. doi:10.1080/1047840X.2017.1328951

Stratou, G., Ghosh, A., Debevec, P., & Morency, L.-P. (2012). Exploring the effect of illumination on automatic expression recognition using the ICT-3DRFE database. *Image and Vision Computing, 30*(10), 728-737. doi:https://doi.org/10.1016/j.imavis.2012.02.001

Tiedens, L. Z., & Fragale, A. R. (2003). Power moves: Complementarity in dominant and submissive nonverbal behavior. *Journal of Personality and Social Psychology, 84*(3), 558-568.

Xie, X., Zheng, W., Lai, J., Yuen, P. C., & Suen, C. Y. (2011). Normalization of Face Illumination Based on Large-and Small-Scale Features. *IEEE Transactions on Image Processing, 20*(7), 1807-1821. doi:10.1109/TIP.2010.2097270

Yongsatianchot, N., & Marsella, S. (2016). Integrating Model-Based Prediction and Facial Expressions in the Perception of Emotion. In B. Steunebrink, P. Wang, & B. Goertzel (Eds.), *Artificial General Intelligence: 9th International Conference, AGI 2016, New York, NY, USA, July 16-19, 2016, Proceedings* (pp. 234-243). Cham: Springer International Publishing.

Feature article: Emotion & Feelings

Feelings, and the Multicomponential Approach to Emotions

David Sander

Swiss Center for Affective Sciences
University of Geneva

With the rise of affectivism (Dukes et al., 2021), affective scientists are increasingly investigating the mechanisms that underly emotions and their interactions with cognitive processes such as attention, memory, and decision-making. While intense debates exist in our field, it seems to me that there are also important agreements that have been reached. I argue below that whereas most current theories of emotion adopt a "multiaspect" perspective to emotion, with feeling being one aspect (or component) of an emotion, the content of feelings remains debated. Below, I also discuss how understanding the relationship between the feeling component and the other components may help solve a secular debate in emotion research concerning the so-called James-Lange theory of emotion, and may also help frame new research questions.

Since at least Aristotle, the idea that the complexity of emotions can be understood with respect to more simple dimensions has been discussed in many of the various fields that now constitute the affective sciences. For instance, Aristotle's idea that we experience pleasure and pain when we have emotions (e.g., Dow, 2015) is arguably present in most contemporary theories of emotions, with the term "valence" being typically used to refer to the displeasure /pleasure dimension(s). Other dimensions have been described as important when it comes to reducing the complexity of our emotional experiences; for instance, arousal, dominance, and unpredictability have all been suggested to underly emotions and other affective phenomena.

The explicit search for which elements can be considered as constituents of an emotion is not very recent (see e.g., McCosh, 1880; Irons, 1897), and in addition to the above-mentioned dimensions, emotion researchers have also described what we could refer to as "aspects" or "components" of emotions (see Sander, 2013). Although there are conceptual distinctions that could be proposed between the constructs of "dimensions" and "components", what I aim at highlighting here is that it has often been suggested that one can decompose an emotion into some constituent elements. For instance, McCosh (1880) considered four elements to be involved in emotion: "*First, there is the affection, or what I prefer calling the motive principle, or the appetence; (…) Secondly, there is an idea of something, of some object or occurrence, as fitted to gratify or disappoint a motive principle or appetence; (…) Thirdly, there is the conscious feeling; (…) Fourthly, there is an organic affection.*" (McCosh, 1880, p. 1-4).

As can be seen from McCosh's early componential approach to emotion, the conscious feeling was already suggested as one of several aspects of emotion almost 150 years ago. Focusing on definitions from the 20th century, Kleinginna and Kleinginna (1981) reminded us that, when considering the 11 categories of definitions of emotion that they established, the "multiaspect" category was the largest one. Indeed, this category included 32 definitions suggested by scholars of the 20th century who emphasized that "*emotion contains several important components*" (Kleinginna and Kleinginna, 1981, p. 352). While each major contemporary psychological theory keeps its specificities with respect to many principles that are proposed to govern an emotional episode, there seems to be a consensus among these theories about the general idea that emotions are multicomponential phenomena. Let us consider, for instance, Basic Emotion Theories, Core-Affect Theories, and Appraisal Theories.

For instance, Paul Ekman is a prominent representative of the Basic Emotion Theories, and according to Matsumoto and Ekman (2009, p. 69): "*A match (…) initiates a group of responses, including expressive behaviour, physiology, cognitions, and subjective experience. The group of responses is coordinated, integrated, and organized, and constitutes what is known as an emotion. (…). In our view, the term 'emotion' is* *a metaphor that refers to this group of coordinated responses."*

A very different perspective in many respects, but similar in terms of a multicomponential approach, is the Core-Affect Theory, which insists on the idea that one should understand the psychological construction of emotions by considering several components. According to Jim Russell, who is a prominent representative of the Core-Affect Theories, "*Psychological construction is not one process but an umbrella term for the various processes that produce: (a) a particular emotional episode's 'components' (such as facial movement, vocal tone, peripheral nervous system change, appraisal, attribution, behaviour, subjective experience, and emotion regulation); (b) associations among the components; and (c) the categorisation of the pattern of components as a specific emotion.*" (Russell, 2009, p. 1259). Arguably, when Matsumuto and Ekman (2009), and Russell (2009) use "subjective experience" as a component of emotion, they use it in the broad sense of "subjective feeling".

As a third family of theories of emotion, appraisal theories are different to the two previously mentioned theories in many ways but typically share the multicompoenential perspective. In this respect, the "Component Process Model" of emotion proposed by Scherer (1984; 2005) is certainly the appraisal theory that, since the 1980s, most explicitly emphasizes the role of various components of emotion. Indeed, "*in the framework of the component process model, emotion is defined as an episode of interrelated, synchronized changes in the states of all or most of the five organismic subsystems in response to the evaluation of an external or internal stimulus event as relevant to major concerns of the organism."* (Scherer, 2005, p. 697). In Scherer's approach, the 5 components are the cognitive component (appraisal processes), the physiological component (bodily symptoms), the motivational component (action tendencies), the motor expression component (facial and vocal expression), and the subjective feeling component (emotional experience). Not all components have the same status: one of them (the appraisal component) determines changes in the other four components (see also Moors, 2014), and the subjective feeling component integrates the activity of the other components (see Grandjean, Sander, and Scherer, 2008).

Professor David Sander

It is important to note that whether the feeling component is necessarily conscious can be a matter of conceptual discussion (see Grandjean, Sander, & Scherer, 2008). However, it seems to me that the terms "feeling," "emotional consciousness," or "emotional experience" are typically used interchangeably in the literature. Inspired by the multicomponential appaisal approach, Pool & Sander (2021) recently proposed Figure 1 as a general way to represent both the emotion components and their inter-relations: a particular event is first appraised by the individual according to their current concerns, values, and goals (motivational processes displayed in yellow). Then, this elicitation process can trigger an emotional response in multiple components: autonomic physiology, action tendency, expression, and feeling. The emotional processes (displayed in red) closely interact with several cognitive processes (displayed in blue) such as attention, memory, learning, and decision-making.

When measuring emotions, we typically use only a few measures (e.g., only self-reports, only psycho-physiological measures, only appraisal questionnaires, only action tendency questionnaires, only expression coding, or only approach-avoidance tendencies). The multicomponential approach suggests that any measure of emotion may only give a probabilistic

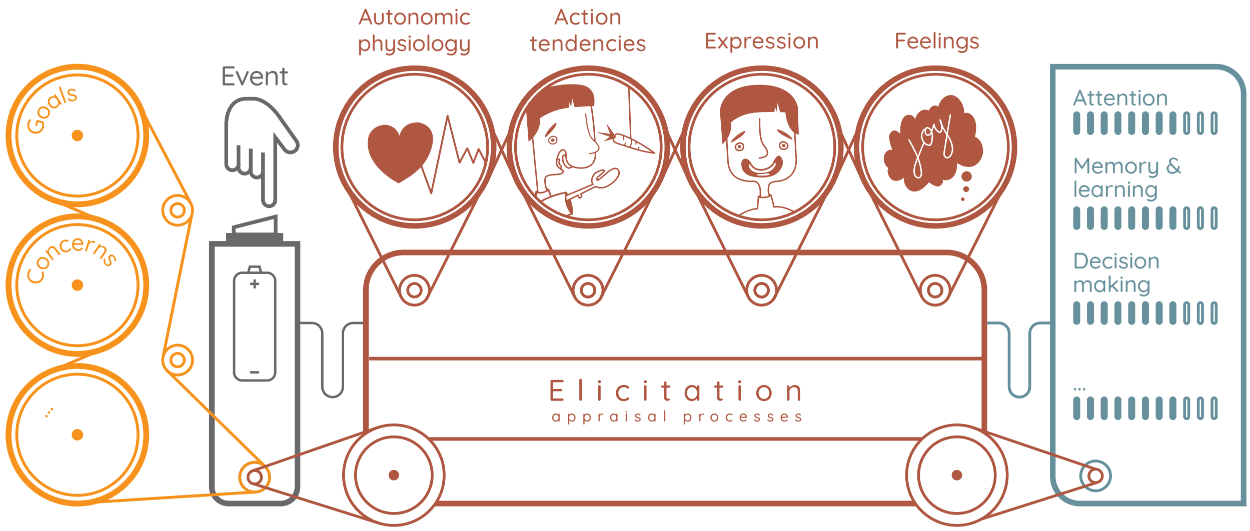


Figure 1. From Pool and Sander (2021).

account of emotion inference if one aims at measuring a specific emotion (e.g., fear, happiness, or awe), and that multiple measures would be needed to achieve converging evidence across the components of emotions (see Delpalnque & Sander, 2021). Indeed, single measures of emotions are not markers of emotions: there is a risk of making wrong reverse inferences when focusing only on single measures; this is why Delplanque & Sander (2021) argued that this risk is decreased when converging evidence is obtained from several components instead of only one. Other ways to measure emotions include measures of brain systems, with methods of high temporal (e.g., EEG) and spatial (e.g., fMRI) resolution.

Although further systematic investigation is needed, it seems to me that the existence of several components is consistent with the way the emotional brain is organized (see Sander, Grandjean, and Scherer, 2018). In fact, the study of the emotional brain (Adolphs & Anderson, 2018), and investigations of the neural networks involved in emotion is a lively development in the affective sciences (see Pessoa, 2018). In affective neuroscience, it has also been considered useful to separate the mechanisms involved in feeling from those involved in other emotional components (for discussion, see Damasio, 1998; Leitao et al., 2020; Sander, 2013), with the idea that some particular brain systems (e.g., the anterior cortical midline structures, see Heinzel et al., 2010) or some modes of interactions between several brain networks (e.g., Thagard and Aubie, 2008; Grandjean, Sander, and Scherer, 2008) allow the specific emergence of feelings.

Interestingly, adopting a multicomponential approach to emotion also appears to be useful beyond the study of emotion itself. For instance, another key domain in affective sciences is the study of emotion regulation mechanisms, and in this domain too, considering various components has proven useful, with different specific component-related strategies (e.g., reappraisal, expressive suppression, or physiological intervention) being studied and compared (see McRae & Gross, 2020). Another example of the usefulness of the multicomponential approach can be found in attempts to understand other affective phenomena than emotions; for instance, a distinction between “wanting” and “liking” components in reward processing (see Berridge & Kringelbach, 2015, Pool et al., in press) can be related to some components of emotion (see Sander, & Nummenmaa, 2021).

Considering the links between the feeling component and the other components leads to the complex question concerning the content of feelings. What are the inputs that are processed by the feeling component during an emotional episode?

In relation to this question, bridges between the Jamesian approaches to emotion and the multicomponential approaches to emotion may be particularly useful when it comes to solving a secular debate in emotion research: the peripheralist-centralist debate. James famously defined standard emotions as follows: "My thesis, on the contrary, is that the bodily changes follow directly the PERCEPTION of the exciting fact, and that our feeling of the same changes as they occur IS the emotion." (James, 1884, p.189-190). It can be noted here that James mentioned both the notions of "bodily changes" and "feeling" in his definition: bodily states would be a necessary condition to actually feel the emotion. Importantly, James explicitly restricted his definitions to those emotions that have a distinct bodily expression (James, 1884, p.189; for discussion, see Friedman, 2010). This is noteworthy because, contrary to what is sometimes assumed, James' definition was not about all emotions but only about this subset of emotions. This suggests that, even for James, some feelings could be elicited via another process.

The focus on the role of bodily changes in emotion and feeling is well known in theories of James, Lange and Sergi, but these ideas were also found before, and, obviously, developed later on (see Damasio, 1998; Damasio & Carvalho, 2013; Friedman, 2010). For instance, traces of the peripheralist view can be recognized in the writings of McCosh who, a few years before James, already wrote: "*If it be true that emotion produces a certain bodily state, it is also true that some bodily states tend to produce the corresponding feeling*" (McCosh, 1880, p. 105; see also Ruckmick, 1934). There is little doubt that James linked bodily changes to a particular aspect of emotion: the conscious aspect. Indeed, James's (1894) introductory sentence is clear in this respect: "*In the year 1884 Prof. Lange of Copenhagen and the present writer published, independently of each other, the same theory of emotional consciousness*" (p. 516).

Therefore, the multicompential approach may help solving the debate of whether the bodily changes are a cause or a consequence of the emotion by considering feeling as a component: the feeling may be (at least partly) determined by a change in the body state, while the other components of emotion may not be caused this body state. For instance, according to Frijda (2005), the emotional experience "*generally contains conscious reflections of the four major nonconscious components of the process of emotions: affect, appraisal, action readiness, and arousal. In addition, it may include the emotion's felt 'significance'*" (p. 494). Assuming that an emotional experience is to be considered as synonymous with an emotional feeling, then it would suggest that, in addition to the bodily changes, the outcomes of the other components should also be considered when it comes to understanding the content of an emotional feeling. This may also mean that, just like one can have a physiological feeling (e.g., feeling of an increased heart rate), one could also have a feeling of appraisal outcomes (e.g., feeling of uncertainty).

Interestingly, in addition to using interoception to feel physiological processes that are not emotional (e.g., hunger, thirst, well-being, cold, or fatigue, see Damasio and Carvalho, 2013; Pace-Schott et al., 2019), one may also feel outcomes of cognitive processes (e.g., feeling of remembering). Following James, there may be emotional experiences that cannot be explained in terms of felt distinct bodily expressions; these experiences could also be explained by a model considering that not only bodily changes can be felt, but that appraisals or action tendencies can also be felt. This would correspond to an extension of the view suggested by Damasio and Carvalho (2013) according to whom "*Feelings are mental experiences of body states*" (p. 143). They suggest that drives and emotions can elicit feelings, but that their definition also excludes the use of "feeling" in the sense of "thinking" or "intuiting". With this view, a research question is therefore whether appraisal outcomes or action tendencies can be felt as direct inputs, or whether they would be felt only via bodily changes.

To conclude, it seems to me that the multicomponential approach to emotion has the potential to bring affective scientists together showing that, despite the useful diversity of theories of emotions, there is a general framework that can be shared. As emotion researchers, we tend to focus on specific processes or components, and study for instance appraisal/elicitation processes, expressions, autonomic responses, action tendencies, or feelings in relative independence. Therefore, a better understanding of the relationships between the components of an emotional episode would also be a way to further the disciplinary and interdisciplinary collaborations between affective scientists who specialize in the study of different components.

In other words, bringing the components together would also bring researchers together. Among the numerous questions that such a framework can support, we can cite a few: How do the components interact together (e.g., in terms of psychological mechanisms, brain networks, and dynamic systems)? How does the inter-relation between components represent specific emotions (e.g., fear or pride)? Are there different weightings of the components as inputs to the feeling component for different emotions (e.g., with bodily states being more weighted for "basic" emotions than for other emotions)? Are some components typically unconscious (e.g., autonomic physiology) while others are typically conscious (e.g., feeling)? How do the components and their relations develop over time, both phylogenetically and ontogenetically? How do the components, and their sub-processes, specifically explain the differential effects of emotions on many cognitive mechanisms and on behaviours? It seems to me that such questions, and many others, while having their roots in historical theories of emotion, have the potential to use multicomponential approaches in order to bring new answers for a better general understanding of emotion.

References

Adolphs, R., & Anderson, D. J. (2018). *The Neuroscience of Emotion: A New Synthesis.* Princeton University.

Berridge, K. C. & Kringelbach, M. L. (2015). Pleasure systems in the brain. *Neuron 86,* 646–664 (2015).

Damasio, A. R. (1998). Emotion in the perspective of an integrated nervous system. *Brain Research Reviews, 26*(2–3), 83-86.

Damasio, A., & Carvalho, G.B., (2013). The nature of feelings: evolutionary and neurobiological origins. *Nature Reviews Neuroscience,14,* 143–152.

Delplanque, S., & Sander, D. (2021). A fascinating but risky case of reverse inference: from measures to emotions! *Food Quality and Preference, 92*, 104183.

Dukes, D., Abrams, K., Adolphs, R., Ahmed, M. E., Beatty, A., Berridge, K. C., Broomhall, S., Brosch, T., Campos, J. J., Clay, Z., Clément, F., Cunningham, W., Damasio, A., Damasio, H., D'Arms, J., Davidson, J., de Gelder, B., Deonna, J., de Sousa, R., Ekman, P., Ellsworth, P., Fehr, E., Fischer, A., Foolen, A., Frevert, U., Grandjean, D., Gratch, J., Greenberg, L., Greenspan, P., Gross, J., Halperin, E., Kappas, A., Keltner, D., Knutson, B., Konstan, D., Kret, M., LeDoux, J., Lerner, J., Levenson, R. W., Loewenstein, G., Manstead, A., Maroney, T., Moors, A., Niedenthal, P., Parkinson, B., Pavlidis, I., Pelachaud, C., Pollak, S., Pourtois, G., Röttger-Rössler, B., Russell, J. A., Sauter, D., Scarantino, A., Scherer, K., Stearns, P., Stets, J., Tappolet, C., Teroni, F., Tsai, J., Turner, J., Van Reekum, C., Vuilleumier, P., Wharton, T., & Sander, D. (2021). The rise of Affectivism. *Nature Human Behaviour, 5,* 816–820.

Dow, J. (2015). Aristotle's Theory of the Passions—Passions as Pleasures and Pains. In J. Dow (2015), *Passions and Persuasion in Aristotle's Rhetoric*, Oxford University Press.

Friedman, B.H. (2010). Feelings and the body: the Jamesian perspective on autonomic specificity of emotion. *Biological Psychology, 84*, 383–393.

Frijda, N.H. (2005). Emotion experience. *Cognition & Emotion, 19*, 473-498.

Grandjean, D., Sander, D., & Scherer, K. R. (2008). Conscious emotional experience emerges as a function of multilevel, appraisal–driven response synchronization. *Consciousness and Cognition, 17*(2), 484-495.

Irons, D. (1897). The primary emotions. *The Philosophical Review, 6*, 626–645.

James, W. (1884). What is an emotion? *Mind, 9*, 188-205.

James, W. (1894). The physical basis of emotion. *Psychological Review, 1,* 516- 529.

Kleinginna, P. R. Jr., & Kleinginna, A. M. (1981). A categorized list of emotion definitions with suggestions for a consensual definition. *Motivation and Emotion, 5,* 345-379.

Leitao, J., Meuleman, B., Van De Ville, D., & Vuilleumier, P. (2020). Computational imaging during video game playing shows dynamic synchronization of cortical and subcortical networks of emotions. *Plos Biology, 18*(11), e3000900.

Matsumoto, D. & Ekman, P. (2009). Basic emotions. In D. Sander and K.R. Scherer (Eds.). *The Oxford Companion to Emotion and the Affective Sciences* (pp. 69-72). New York and Oxford: Oxford University Press.

McCosh, J. (1880). *The Emotions,* New York: Charles Scribners'sons.

McRae, K., & Gross, J. J. (2020). Emotion regulation. *Emotion, 20*(1), 1-9.

Moors, A. (2014). Flavors of appraisal theories of emotion. *Emotion Review, 6*(4), 303–307.

Pace-Schott, E. F., Amole, M. C., Aue, T., Balconi, M., Bylsma, L. M., Critchley, H., Demaree, H. A., Friedman, B. H., Gooding, A. E. K., Gosseries, O., Jovanovic, T., Kirby, L. A. J., Kozlowska, K., Laureys, S., Lowe, L., Magee, K., Marin, M.-F., Merner, A. R., Robinson, J. L., . . . VanElzakker, M. B. (2019). Physiological feelings. *Neuroscience and Biobehavioral Reviews, 103,* 267–304.

Pessoa, L. (2018). Emotion and the Interactive Brain: Insights From Comparative Neuroanatomy and Complex Systems. *Emotion Review, 10*(3), 204–216.

Pool, E. R., Munoz Tord, D., Delplanque, S., Stussi, Y., Cereghetti, D., Vuilleumier, P., & Sander, D. (in press). Differential contributions of ventral striatum subregions in the motivational and hedonic components of affective processing of reward. *The Journal of Neuroscience.*

Pool, E. R., & Sander, D. (2021). Emotional learning: Measuring how affective values are acquired and updated. In H. L. Meiselman (Ed.), *Emotion measurement* (2nd ed., pp. 133-165). Woodhead Publishing.

Ruckmick, C. A. (1934). McCosh on the emotions. *The American Journal of Psychology, 46,* 506–508.

Russell, J. A. (2009). Emotion, core affect, and psychological construction. *Cognition & Emotion, 23*(7), 1259 -1283.

Sander, D. (2013). Models of emotion: The affective neuroscience approach. In J. L. Armony & P. Vuilleumier (Eds.), *The Cambridge handbook of human affective neuroscience* (pp. 5-53). Cambridge, UK: Cambridge University Press.

Sander, D., Grandjean, D., & Scherer, K. R. (2018). An appraisal-driven componential approach to the emotional brain. *Emotion Review, 10*(3), 219-231.

Sander, D., & Nummenmaa, L. (2021). Reward and emotion: An affective neuroscience approach. *Current Opinion in Behavioral Sciences, 39*, 161-167.

Scherer, K. R. (1984). On the nature and function of emotion: A component process approach. In K. R. Scherer, & P. Ekman (Eds.), *Approaches to emotion* (pp. 293–317). Hillsdale, NJ: Erlbaum.

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information, 44*(4), 695-729.