Emotional Expressions Reconsidered:

Challenges to Inferring Emotion From Human Facial Movements

Lisa Feldman Barrett[1,2,3]*, Ralph Adolphs[4], Stacy Marsella[1,5,6], Aleix Martinez[7]

and Seth D. Pollak[8]

[1] Northeastern University, Department of Psychology, Boston, MA

[2] Massachusetts General Hospital, Department of Psychiatry and the Athinoula A. Martinos Center for Biomedical Imaging, Charlestown, MA

[3] Harvard Medical School, Department of Psychiatry, Boston MA

[4]California Institute of Technology, Departments of Psychology, Neuroscience, and Biology, Pasadena, CA

[5] Northeastern University, College of Computer and Information Science, Boston, MA

[6] University of Glasgow, Glasgow, Scotland

[7] The Ohio State University, Department of Electrical and Computer Engineering, and Center for Cognitive and Brain Sciences, Columbus, OH

[8] University of Wisconsin - Madison, Department of Psychology, Madison, WI

*Correspondence to: Lisa F. Barrett (l.barrett@neu.edu; 125 Nightingale Hall, Northeastern University, Boston MA 02115, USA).

In press, *Psychological Science in the Public Interest*

Table of Contents

Executive Summary

It is commonly assumed that a person's face gives evidence of emotions because there is a reliable one-to-one mapping between a certain configuration of facial movements, called a "facial expression," and the specific emotional state that is supposedly signals. This common view of facial expressions remains entrenched in consumers of emotion research, as well as in some scientists, despite an emerging consensus among affective scientists that emotional expressions are considerably more context-dependent and variable. Nonetheless, this common view continues to fuel commercial applications in industry and government (e.g., automated detection of emotions from faces), guide how children are taught (e.g., with posters and books showing stereotyped facial expressions), and impact clinical and legal applications (e.g., diagnoses of psychiatric illnesses and courtroom decisions). In this paper, we evaluate the common view of facial expressions against a review of the evidence and conclude that its rest on a number of flawed assumptions and incorrect interpretations of research findings. Our review is the most comprehensive and systematic to date, encompassing studies of healthy adults across cultures, newborns and young children, as well as people who are congenitally blind, and confirms that specific emotion categories -- anger, disgust, fear, happiness, sadness, and surprise – are each expressed with a particular configuration of facial movements, more reliably than would be expected by mere chance, but contrary to the common view, instances of these emotion categories are NOT expressed with facial movements that are sufficiently reliable and specific across contexts, individuals, and cultures to be considered diagnostic displays of any emotional state. Nor do human perceivers, in fact, infer emotions from particular configurations of muscle movements in a sufficiently reliable and specific way that similarly generalizes. Studies of expression production and perception both demonstrate multiple sources of variability that contradict the common view that smiles, scowls, frowns, and the like, are reliable and specific "expressions of emotion." We conclude the paper with specific recommendations for both scientists and consumers of science.

Abstract

It is commonly assumed that a person's emotional state can be readily inferred from the person's facial movements, typically called "emotional expressions" or "facial expressions." This assumption influences legal judgments, policy decisions, national security protocols, and educational practices, guides the diagnosis and treatment of psychiatric illness, as well as the development of commercial applications, and pervades everyday social interactions as well as research in other scientific fields such as artificial intelligence, neuroscience, and computer vision. In this paper, we survey examples of this widespread assumption, which we refer to as the "common view", and then examine the scientific evidence for this view with a focus on the six most popular emotion categories used by consumers of emotion research: anger, disgust, fear, happiness, sadness and surprise. The available scientific evidence suggests that people do sometimes smile when happy, frown when sad, scowl when angry, and so on, more than what would be expected by chance. Yet there is substantial variation in how people communicate anger, disgust, fear, happiness, sadness and surprise, across cultures, situations, and even within a single situation. Furthermore, similar configurations of facial movements variably express instances of more than one emotion category. In fact, a given configuration of facial movements, such as a scowl, often communicates something other than an emotional state. Scientists agree that facial movements convey a range of social information and are important for social communication, emotional or otherwise. But our review suggests there is an urgent need for research that examines how people *actually* move their faces to express emotions and other social information in the variety of contexts that make up everyday life, as well as careful study of the mechanisms by which people perceive instances of emotion in one another. We make specific research recommendations that will yield a more valid picture of how people move their

faces to express emotions, and how they infer emotional meaning from facial movements, as situations of everyday life. This research is crucial to provide consumers of emotion research with the translational information they require.

Introduction

Faces are a ubiquitous part of everyday life for humans. We greet each other with smiles or nods. We have face-to-face conversations on a daily basis, whether in person or via computers. We capture faces with smartphones and tablets, exchanging photos of ourselves and of each other on Instagram, Snapchat, and other social media platforms. The ability to perceive faces is one of the first capacities to emerge after birth: an infant begins to perceive faces within the first few days of life, equipped with a preference for face-like arrangements that allows the brain to wire itself, with experience, to become expert at perceiving faces (Arcaro et al., 2017; Cassia et al., 2004; Grossmann, 2015; Ghandi et al., 2017; Smith et al., 2018; Turati, 2004; but see Young & Burton (2018) for a more qualified claim). [1] Faces offer a rich, salient source of information for navigating the social world: they play a role in deciding who to love, who to trust, who to help, and who is found guilty of a crime (Todorov, 2017; Zebrowitz, 1997, 2017; Zhang, Chen & Yang, 2018). Dating back to the ancient Greeks (Aristotle, in 4th century BCE) and Romans (Cicero), various cultures have viewed the human face as a window on the mind. But to what extent can a raised eyebrow, a curled lip, or a narrowed eye reveal what someone is thinking or feeling, allowing a perceiver's brain to guess what that someone will do next?[2] The answers to these questions have major consequences for human outcomes as they unfold in the living room, the classroom, the courtroom and even on the battlefield. They also powerfully shape the direction of research in a broad array of scientific fields, from basic neuroscience to psychiatry research.

Understanding what facial movements might reveal about a person's emotions is made more urgent by the fact that many people believe we already know. Specific **configurations of facial muscle movements** appear as if they summarily broadcast or **display** a person's emotions, which is why they are routinely referred to as "**emotional expressions**" and "facial expressions."[3] A simple Google search using the phrase "emotional facial expressions" [see Box 1, in supplementary on-line materials (SOM)] reveals the ubiquity with which, at least in certain parts of the world, people believe that certain emotion categories are reliably signaled or revealed by certain facial muscle movement configurations – a set of beliefs were refer to as the **common view** (also called the classical view; Barrett, 2017a). Similarly, many cultural products testify to the common view. Here are several examples:

- Technology companies are investing tremendous resources to figure out how to objectively "read" emotions in people by detecting their presumed facial expressions, such as scowling faces , frowning faces and smiling faces in an automated fashion. Several companies claim to have already done it (e.g., https://www.affectiva.com/what/products/ ; https://azure.microsoft.com/en-us/services/cognitive-services/emotion/). For example, Microsoft's Emotion API promises to take video images of a person's face to detect what that individual is feeling. The application states: "The emotions detected are anger, contempt, disgust, fear, happiness, neutral, sadness, and surprise. These emotions are understood to be cross-culturally and universally communicated with particular facial expressions"(https://azure.microsoft.com/en-us/services/cognitive-services/emotion/).

- Countless electronic messages are annotated with emojis or emoticons that are schematized versions of the proposed facial expressions for various emotion categories (https://www.apple.com/newsroom/2018/07/apple-celebrates-world-emoji-day/).

- Putative emotional expressions are taught to preschool children by displaying scowling faces, frowning faces, smiling faces and so on, in posters (e.g., use "feeling chart for children" in a Google image search), games (https://www.amazon.com/s/ref=nb_sb_noss_2?url=search-alias%3Daps&field-keywords=miniland+emotion)and books (e.g., Cain, 2000; Parr, 2005), and on episodes of *Sesame Street* (among many examples, see https://www.youtube.com/watch?v=ZxfJicfyCdg , https://vimeo.com/108524970 , or https://www.youtube.com/watch?v=y28GH2GoIyc).[4]

- Television shows (e.g., *Lie to Me*), movies (e.g., *Inside Out*) and documentaries (e.g., *The Human Face*, produced by the British Broadcasting Company) customarily depict certain facial configurations as universal expressions of emotions.

- Magazine and newspaper articles routinely feature stories in kind: facial configurations depicting a scowl are referred to as "expressions of anger," facial configurations depicting a smile are referred to as "expressions of happiness," facial configurations depicting a frown are referred to as "expressions of sadness," and so on.

- Agents of the U.S. Federal Bureau of Investigations (FBI) and the Transportation Security Administration (TSA) were trained to detect emotions and other intentions using these facial configurations, with the goal of identifying and thwarting terrorists (Rhonda Heilig, special agent with the FBI, personal communication, December 15, 2014, 11:20 am; https://how-emotions-are-made.com/notes/Screening_of_Passengers_by_Observation_Techniques).[5]

- The facial configurations that supposedly diagnose emotional states also figure prominently in the diagnosis and treatment of psychiatric disorders. One of the most widely used task in autism research, the "Reading the Mind in the Eyes Test", asks patients to match photos of the upper (eye) region of a posed facial configuration with specific mental state words, including emotion words (Baron-Cohen t al., 2001). Treatment plans for people living with autism and other brain disorders often include learning to recognize these facial configurations as emotional expressions (Baron-Cohen et al., 2004; Kouo & Egel, 2016). This training does not generalize well to real-world skills, however (Bergren et al., 2018; Kouo & Egel, 2016).

- "Reading" the emotions of a defendant (in the words of Supreme Court Justice Anthony Kennedy -- to "know the heart and mind of the offender") is one pillar of a fair trial in the U.S. legal system and in many legal systems in the Western world (see Riggins v. Nevada, 1992). Legal actors like jurors and judges routinely rely on facial movements to determine the guilt and remorse of a defendant (e.g., Bandes, 2014; Zebrowitz, 1997). For example, defendants who are perceived as untrustworthy receive harsher sentences than they otherwise would (Wilson & Rule, 2015, 2016), and such perceptions are more likely when a person appears to be angry (i.e., facial structure is similar to the hypothesized facial expression of anger, which is a scowl (Todorov, 2017). An incorrect inference about a defendant's emotional state can cost someone her children, her freedom, or even her life (for recent examples, see Barrett, 2017, beginning on page 183).

But can a person's emotional state be reasonably inferred from that person's facial movements? In this paper, we offer a systematic review of the evidence, testing the common view that instances of emotion are signaled with a distinctive configuration of facial movements with enough consistently that it can serve as a diagnostic marker of those instances. We focus our review on evidence pertaining to six emotion categories that have received the lion's share of attention in the scientific literature -- anger, disgust, fear, happiness, sadness and surprise – and that, correspondingly, are the focus of common view (as evidenced by our Google search, summarized in Box 1, SOM), but our conclusions apply to all emotion categories that have thus far been scientifically studied. We open the paper with a brief discussion of its scope, approach, and intended audience. We then summarize evidence on how people *actually* move their faces during episodes of emotion, referred to as studies of **expression production** studies, following which we examine evidence for which emotions are actually inferred from looking at facial movements, referred to as studies of **emotion perception.** We identify three key shortcomings in the scientific research that have contributed to a general misunderstanding about how emotions are expressed and perceived in facial movements, and that limit the translation of this scientific evidence for other uses:

(1) *limited reliability* (instances of the same emotion category are neither reliably expressed with or perceived from a common set of facial movements);

(2) *lack of specificity* (there is no one-to-one mapping between a single configuration of facial movements and instances of the same emotion category); and,

(3) *limited generalizability* (the effects of context and culture have not been sufficiently documented and accounted for).

We then discuss our conclusions, followed by proposals for consumers on how they might use the existing scientific literature. We also provide recommendations for future research with consumers of emotion research in mind. We have included additional detail on some topics of import or interest in the supplementary on-line materials (SOM).

Scope, Approach and Intended Audience of Paper

The Common View: Reading an Inner Emotional State of Mind From A Set of Unique Facial Movements

In common English parlance, people refer to "emotions" or "an emotion" as if anger, happiness, or any emotion word refers to an object that is highly similar on every occurrence. But an emotion word refers not to a unitary entity, but to a **category** of instances that vary from one another in their **physical features**, such as facial expressions and bodily changes, and **mental features**. Few scientists who study emotion, if any, take the view that every instance of an emotion category, such as anger, is identical to every other instance, sharing a set of necessary and sufficient features across situations, people and cultures. For example, Keltner and Cordaro (2017) recently wrote, “there is no one-to-one correspondence between a specific set of facial muscle actions or vocal cues and any and every experience of emotion” (p. 62). Yet there is considerable scientific debate about the *amount* of the within-category variation, the specific *features* that vary, the *causes* of the within-category variation, and *implications* of this variation for the nature of emotion (see Figure 1).

One popular scientific framework, referred to as the basic emotion approach, hypothesizes that instances of an emotion category are expressed with facial movements that vary, to some degree, around a typical set of movements (called a **prototype**) (for example, see Table 1). For example, it is hypothesized that in one instance, anger might be expressed with the expressive prototype (e.g., brows furrowed, eyes wide, lips tightened) plus additional facial movements, such as a widened mouth, whereas on other occasions, a facial movement in the prototype might be missing (e.g., anger might be expressed with narrowed eyes or without movement in the eyebrow region; for a discussion, see Box 2, in SOM). Nonetheless, the basic emotion approach still assumes that the core facial configuration – the prototype -- can be used to *diagnose* a person’s inner emotional state in much the same way that a fingerprint can be used to *uniquely* recognize a person. More substantial variation in expressions (e.g., smiling in anger, gasping with widened eyes in anger, and scowling not in anger, but in confusion or concentration) is typically explained as the result of some process that is independent of an emotion itself, such as **display rules**, emotion regulation strategies such as suppressing the expression, or culture-specific dialects (as proposed by various scientists, including Elfenbein, 2013, 2017; Ekman & Cordaro, 2011; Matsumoto, 1990; Matsumoto, Keltner, Shiota, Frank, & O’Sullivan, 2008; Tracy & Randles, 2011).

By contrast, other scientific frameworks propose that expressions of the same emotion category, such as anger, *substantially vary by design*, in a way that is tied to the immediate context, which includes the internal context (e.g., the person's metabolic condition, the past experiences that come to mind, etc.) and the outward context (e.g., whether a person is at work, at school, or at home, who else is present the broader cultural conditions, etc.), both of which vary in dynamic ways over time (see Box 2, SOM). These debates, while useful to scientists, provide little clear guidance for consumers of emotion research who are focused on the practical issue of whether various emotion categories are expressed with facial configurations of sufficient regularity and distinctiveness so that it is possible to read emotion in a person's face.

The common view of emotional expressions persist, too, because scientists' actions often don't follow their claims. Many scientists continue to design experiments, use stimuli and publish review papers that, ironically, leave readers with the impression that certain emotion categories each have a single, unique facial expression, even as those same scientists acknowledge that every emotion category can be expressed with a variable set of facial movements. Published studies typically test a one-to-one emotion-expression link (for examples, see the reference lists in Elfenbein & Ambady, 2002; Matsumoto, Keltner, Shiota, O'Sullivan & Frank, 2008; Keltner, Sauter, Tracy & Cowen, in press; also see most of the studies cited in this paper, e.g., Cordaro, Sun, Keltner, Kamble, Huddar, & McNeil, 2017). The exact facial configuration tested varies slightly from study to study, but a core facial configuration is still assumed (see Table 1 for examples), reinforcing rather than dispelling the common view that each emotion category is consistently and uniquely expressed with its own distinctive configuration of facial movements, which therefore can be used to diagnose its presence. Review articles reinforce this impression of a one-to-one mapping by including tables and figures that display a single, unique facial configuration for each emotion category, referred to as the expression, signal or display for that emotion (Figure 2 presents two recent examples).[6]

The common view of emotional expressions has also been imported into other scientific disciplines with an interest in understanding emotions, such as neuroscience and artificial intelligence (AI). For example, from a published paper on AI:

> "American psychologist Ekman noticed that some facial expressions corresponding to certain emotions are common for all the people independently of their gender, race, education, ethnicity, etc. He proposed the discrete emotional model using six universal emotions: *happiness*, *surprise*, *anger*, *disgust*, *sadness* and *fear."* (Brodny, Kolakowska, Landowska, Szwoch, Szwoch, & Wróbel, 2016, p. 1, italics in the original)

Similar examples come from our own papers. One paper series of papers focused on the brain structures involved in perceiving emotions from facial configurations (Adolphs, 2002; Adolphs et al., 1994) and the other focused on early life experiences (Pollak et al., 2000; Pollak & Kistler, 2002). These papers were framed in terms of "recognizing facial expressions of emotion" and exclusively presented participants with specific, posed photographs of scowling faces (the presumed facial expression for anger), wide-eyed gasping faces (the presumed facial expression for fear), and so on. Participants were shown faces of different individuals all posing the same facial configuration for each emotion category, ignoring the importance of context. One reason for this flawed approach to investigating the perception of emotion from faces was that then -- at the time these studies were conducted – as now, published experiments, review articles, and stimulus sets were dominated by the common view that certain emotion categories were signaled with an invariant set of facial configurations, referred to as "facial expressions of basic emotions."

In this paper, we review the scientific evidence that directly tests two beliefs that form the common view of emotional expressions: that certain emotion categories are each routinely expressed by a unique facial configuration and, correspondingly, that people can reliably infer someone else's emotional state from a set of facial movements. Our discussion is written for consumers of emotion research, whether they be scientists in other fields or non-scientists, who need not have deep knowledge of the various theories, debates, and broad range of findings in the science of emotion, with sufficient pointers to those discussions if they are of interest (see Box 2, SOM).

In discussing what this paper is about – the common view that a person's inner emotional state is revealed in facial movements -- it bears mentioning what this paper is *not* about: This paper is not a referendum on "basic emotion" view we briefly mentioned earlier in this section, proposed by the psychologist Paul Ekman and his colleagues, or any other research program or psychologist's view. Ekman's theoretical approach has been highly influential in research on emotion for much of the past 50 years. We often cite studies inspired by the basic emotion approach for this reason. In addition, the common view of emotional expressions is also most readily associated with a simplified version of basic emotion approach, as exemplified by the quotes above. Critiques of Ekman's basic emotion view (and related views) are numerous (e.g., Barrett, 2006a, 2007, 2011; Ortony & Turner, 1990; Russell, 1991, 1994, 1995), as are rejoinders that defend it (e.g., Ekman, 1992, 1994; Izard, 2007). Our paper steps back from this dialogue. We instead take as our focus the existing research on emotional expression and emotion perception and ask whether it is sufficiently strong to justify the way it is increasingly being used by those who consume it.

A Systematic Approach for Evaluating the Scientific Evidence

When you see someone smile and infer that the person is happy, you are making what is known as a **reverse inference**: you are assuming that the smile reveals something about the person's emotional state that you cannot access directly (see Figure 3). Reverse inference requires calculating a conditional probability: the probability that a person is in a particular emotion episode (such as happiness) given the observation of a unique set of facial muscle movements (such as a smile). The conditional probability is written as:

p[emotion category|a unique facial configuration])

for example,

p[happiness| a smiling facial configuration])

Reverse inferences about emotion are ubiquitous in everyday life – whenever you experience someone as emotional, your brain has performed a reverse inference, guessing at the cause of a facial movement when only having access to the movement itself. Every time an app on a phone or computer measures someone's facial muscle movements, identifies a facial configuration such as a frowning facial configuration, and proclaims that the target person is sad, that app has engaged in reverse inference, such as:

p[sadness| a frowning facial configuration])

Whenever a security agent infers anger from a scowl, the agent has assumed a strong likelihood for

p[anger| a scowling facial configuration])

Four criteria must be met to justify a reverse inference that a particular facial configuration expresses and therefore reveals a specific emotional state: **reliability**, **specificity**, **generalizability** and **validity** (explained in Table 2 and Figure 3). These criteria are commonly encountered in the field of psychological measurement and over the last several decades there has been an ongoing dialogue about thresholds for these criteria as they apply in production and perception studies, with some consensus emerging for the first three criteria (see Haidt &

Keltner, 1999). Only when a pattern of facial muscle movements strongly satisfies these four criteria can we justify calling it an "emotional expression." If any of these criteria are not met, then we should instead refer to a facial configuration with more neutral, descriptive terms without making unwarranted inferences, simply calling it a smile (rather than an expression of happiness), a frown (rather than an expression of sadness), a scowl (rather than an expression of anger), and so on.[7]

The Null Hypothesis and the Role of Context

Tests of reliability, specificity, generalizability and validity are almost always compared to what would be expected by sheer chance, if facial configurations (in studies of expression production) and inferences about facial configurations (in studies of emotion perception) occurred randomly with no relation to particular emotional states. In most studies, chance levels constitute the **null hypothesis**. An example of the null hypothesis for reliability is that people do not scowl when angry more frequently than would be expected by chance.[8] If people are observed to scowl more frequently when angry than they would by chance, then the null hypothesis can be rejected based on the reliability of the findings. We can also test the null hypothesis for specificity: If people scowl more frequently than they would by chance not only when angry but also when fearful, sad, confused, hungry, etc., then the null hypothesis for specificity is retained.[9]

In addition to testing hypotheses about reliability and specificity, tests of generalizability are becoming more common in the research literature, again using the null hypothesis. Questions about generalizability test whether a finding in one experiment is reproduced in other experiments in different contexts, using different experimental methods or sampling people from different populations. There are two crucial questions about generalizability when it comes to the production and perception of emotional expressions: Do the findings from a laboratory experiment generalize to observations in the real world? And, do the findings from studies that sample participants from Westernized, Educated, Industrialized, Rich and Democratic (WEIRD; Henrich, Heine, & Norenzayan, 2010) populations generalize to people who live in small-scale, remote communities?

Questions of validity are almost never addressed in production and perception studies. Even if reliable and specific facial movements are observed across generalizable circumstances, it is a difficult and unresolved question as to whether these facial movements can justify an inference about a person's emotion state. We have more to say about this later. In this paper, we evaluate the common view by reviewing evidence pertaining to the reliability, specificity, and generalizability of research findings from production and perception studies.

A focus on rejecting the null hypothesis, defined by what would be expected by chance alone, provides necessary but not sufficient support for the common view of emotional expressions. A slightly above chance co-occurrence of a facial configuration and instances of an emotion category, such as scowling in anger – for example, a correlation coefficient around r = .20 to .39 (adapted from Haidt & Keltner, 1999) -- suggests that a person *sometimes* scowls in anger, but not most or even much of the time. Weak evidence for reliability suggests that other factors not measured in the experiment are likely causing people to scowl during an instance of anger. It also suggests that people may express anger with facial configurations other than a scowl, possibly in reliable and predictable ways. Following common usage, we refer to these unmeasured factors collectively as **context**. A similar situation can be described for studies of emotion perception: when participants label a scowling facial configuration as "anger" in a weakly reliable way (between .20 and .39 percent of the time; Haidt & Keltner, 1999), then this suggests the possibility of unmeasured context effects.

In *principle*, context effects make it possible to test the common view by comparing it directly to an **alternative hypothesis** that a person's brain will be influenced by other causal factors (as opposed to comparing the findings to random chance). It is possible, for example, that a state of anger is expressed differently depending on various factors that can be studied, including the situational context (such as whether a person is at work, at school, or at home), social factors (such as who else is present in the situation and the relationship between the expresser and the perceiver), the person's internal physical context (based on how much sleep they had, how hungry they are, etc.), a person's internal mental context (such as the past experiences that come to mind or the evaluations they make), the temporal context (what just occurred a moment ago), differences between people (such as whether someone is male or female, warm or distant), and the cultural context, such as whether the expression is occurring in a culture that values the rights of individuals (vs. group cohesion), is open and allows for a variety of behaviors in a situation (vs. closed, having more rigid rules of conduct). Other theoretical approaches offer some of these specific alternative hypotheses (see Box 2 in SOM). In *practice*, however, experiments almost always test the common view against the null hypothesis for reliability and specificity and rarely test specific alternative hypotheses. When context is acknowledged and studied, it is usually examined as a factor that might moderate a common and universal emotional expression, preserving the core assumptions of the common view (e.g., Cordaro et al., 2017; for more discussion, see Box 3, SOM).

A Focus on Six Emotion Categories: Anger, Disgust, Fear, Happiness, Sadness and Surprise

Our critical examination of the research literature in this paper focuses primarily on testing the common view of facial expressions for six emotion categories -- anger, disgust, fear, happiness, sadness and surprise. We do not include a discussion of every emotion category ever studied in the science of emotion. We do not discuss the many emotion categories that exist in non-English speaking cultures, such as *gigil*, the irresistible urge to pinch or squeeze something cute, or *liget*, exuberant, collective aggression (for discussion of non-English emotion categories, see Mesquita & Frijda, 1992; Pavlenko, 2014; Russell, 1991). We do not discuss the various emotion categories that have been documented throughout history (e.g., Smith, 2016). Nor do we discuss every English emotion category for which a prototypical facial expression has been suggested. For example, recent studies motivated primarily by the basic emotion approach have suggested that there are "more than six distinct facial expressions …in fact, upwards of 20 multimodal expressions" (Keltner et al., in press, pg. 4), meaning that scientists have proposed a prototypic facial configuration as the facial expression for each of twenty or so emotion categories, including confusion, embarrassment, pride, sympathy, awe, and so on.

The reasons for our focus on six emotion categories are twofold. First, anger, disgust, fear, happiness, sadness and surprise categories anchor common beliefs about emotions and their expressions (as is evident from Box 4, in SOM) and therefore represent the clearest, strongest test of the common view. Second, these six emotion categories have been the primary focus of systematic research for almost a century and therefore provide the largest corpus of scientific evidence that can be evaluated. Unfortunately, the same cannot be said for any of other emotion categories in question. This is a particularly important point when considering the twenty plus emotion categories that are now the focus of research attention. A PsycInfo search for the term "facial expression" combined with "anger, disgust, fear, happiness, sadness, surprise" produced over 700 entries, but a similar search including "love, shame, contempt, hate, interest, distress, guilt" returned less than 70 entries (Duran & Fernandez-Dols, 2018). Almost all cross-cultural studies of emotion perception have focused on just anger, disgust, fear, happiness, sadness and surprise (plus or minus a few) and experiments that measure how people spontaneously move their faces to express instances of emotion categories other than these six remain rare. In particular, there are too few studies that measure spontaneous facial movements during episodes of other emotion categories (i.e., production studies) to conclude anything about reliability and specificity, and there are too few studies of how these additional emotion categories are perceived in small-scale, remote cultures to conclude anything about generalizability. In an era where the generalizability and robustness of psychological findings are under close scrutiny, it seemed prudent to focus on the emotion categories for which there are, by a factor of ten, the largest number of published experiments. Our discussion, which is based on a sample of six emotion categories, generalizes to emotion categories that have been studied, however.[10]

The proposed expressive facial configurations for each emotion category are presented in Figure 4, and the origin of these facial configurations is discussed in Box 4 in SOM. They originated with Charles Darwin, who stipulated (rather than discovered) that certain facial configurations are expressions of certain emotion categories, inspired by photographs taken by Duchenne and drawings made by the Scottish anatomist Charles Bell (Darwin, 1872). These stipulations largely form the basis of the common view of emotional expressions.

Producing Facial Expressions of Emotion: A Review of the Scientific Evidence

In this section, we first review the design of a typical experiment where emotions are induced and facial movements are measured. This review highlights several observations to keep in mind as we review the reliability, specificity and generalizability for expressions of anger, disgust, fear, happiness, sadness and surprise in a variety of populations, including adults in both urban and small-scale remote cultures, infants and children, and congenitally blind individuals. Our review is the most comprehensive to date and allows us to comment on whether the scientific findings generalize across different populations of individuals. The value of doing so becomes apparent when we observe how similar conclusions emerge from these research domains.

The Anatomy of a Typical Experiment Designed to Observe People's Facial Movements During Episodes of Emotion

In the typical expression production experiment, scientists expose participants to objects, images or events that they (the scientists) believe will evoke an instance of emotion. It's possible, in principle, to evoke a wide variety of instances for a given emotion category (e.g., Wilson-Mendenhall et al., 2015), but in practice, published studies evoke the most typical instances of each category, often elicited with a stimulus that is presented without context (e.g., a photograph, a short movie clip separated from the rest of the film, etc.). Scientists usually include some measure to verify that participants are in the expected emotional state (such as asking participants to describe how they feel by rating their experience against a set of emotion adjectives). They then observe participants' facial movements during the emotional episode and then quantify how well the measure of emotion predicts the observed facial movements. When done properly, this yields estimates of reliability and specificity, and in principle provides data to assess generalizability. There are limitations to assessing the validity of a facial configuration as an expression of emotion, as we explain below.

Measuring facial movements. Healthy humans have a common set of 17 facial muscle groups on each side of the face that contract and relax in patterns.[11] To create facial movements that are visible to the naked eye, facial muscles contract, changing the distance between facial features (Neth & Martinez, 2009) and shaping skin into folds and wrinkles on an underlying skeletal structure. Even when facial movements look the same to the naked eye, there may be differences in their execution under the skin. There are individual differences in mechanics of making a facial movement, including variation in the anatomical details (e.g., everyone has a slightly different configuration and relative size of the muscles, some people lack certain muscle components, etc.), in the neural control of those muscles (Cattaneo & Pavesi, 2014; Hutto & Vattoth, 2015; Muri, 2015), and in the underlying skeletal structure of the face (discussed in Box 5, in SOM).

There are three common procedures for measuring facial movements in a scientific experiment. The most sensitive, objective measure of facial movements detects the electrical activity from actual muscular contractions, called facial electromyography (again, see Box 5, in SOM). This is a **perceiver-independent** way of assessing facial movements that detects muscle contractions that are not necessarily visible to the naked eye (Tassinary & Cacioppo 1992). Facial EMG's utility is unfortunately offset by its impracticality: facial EMG requires placing electrodes on a participant's face, which can cause skin abrasions. In addition, a person can typically tolerate only a few electrodes on the face at a time. At the writing of this paper, there were relatively few published papers using facial EMG (we identified 123 studies), the overwhelming majority of which sparsely sampled the face, measuring the electrical signals for only a small number of muscles (between one to six); none of the studies measured naturalistic facial movements as they occur outside the lab, in everyday life. As a consequence, we focus our discussion on two other measurement methods: a **perceiver-dependent** method that describes visible facial movements, called facial actions, which uses human coders who indicate the presence or absence of a facial movement while viewing video recordings of participants, and automated methods for detecting of facial actions from photographs or videos.

Measuring facial movements with human coders. The Facial Action Coding System, or FACS (Ekman et al., 2002), is a systematic approach to describe what a face looks like when facial movements have occurred. FACS codes describe the presence and intensity of facial movements. Importantly, FACS is purely descriptive and is therefore *agnostic about whether those movements might express emotions or any other mental event*.[12] Human coders train for many weeks to reliably identify specific movements called "action units" or AUs. Each AU is hypothesized to correspond to the contraction of a distinct facial muscle or a distinct grouping of muscles that is visible as a specific facial movement. For example, the raising of the inner corners of the eyebrows (contracting the *frontalis* muscle *pars medialis)* corresponds to AU 1. Lowering of the inner corners of the brows (activation of the corrugator supercilii, depressor glabellae and depressor supercilii) corresponds to AU 4. AUs are scored and analyzed as independent elements, but the underlying anatomy of many facial muscles constrains them so they cannot move independently of one another, generating dependencies between AUs (e.g., see Hao, Wang, Peng, & Ji, 2018). Facial action units (AU) and their corresponding list of facial muscles can be found in Table 3. Expert FACS coders approach inter-rater reliabilities of .80 for individual AUs (Jeni, Cohn, & De la Torre, 2013). The first version of FACS (Ekman & Friesen, 1978) was largely based on the work of Swedish anatomist Carl-Herman Hjortsjö who catalogued the facial configurations described by Duchenne (Hjortsjö, 1969). In addition to the updated versions of FACS (Ekman et al., 2002), other facial coding systems have been devised for human infants (Izard et al., 1995; Oster, 2003), chimpanzees (Vick et al., 2007), and macaque monkeys (Parr et al., 2010).[13] Figure 4 displays the common FACS codes for the configurations of facial movements that have been proposed as the expression of anger, disgust, fear, happiness, sadness and surprise.

Measuring facial movements with automated algorithms. Human coders require time-consuming, intensive training and practice before they can reliably assign AU codes. After training, it is a slow process to code photographs or videos frame by frame making human FACS coding impractical to use on facial movements as they occur in everyday life. Large inventories of naturalistic photographs and videos, which have been curated only fairly recently (Benitez-Quiroz et al., 2016), would require decades to manually code. This problem is addressed by automated FACS coding systems using computer vision algorithms (Martinez & Du, 2012; Martinez, 2017; Valstar et al., 2017).[14] Recently developed computer vision systems have automated the coding of some (but not all) facial AUs (e.g., Benitez-Quiroz et al., *in press*; Benitez-Quiroz et al., 2017b; Chu et al., 2017; Corneanu et al., 2016; Essa & Pentland, 1997; Martinez, 2017a; Martinez & Du, 2012; Valstar et al., 2017; see Box 6, SOM) making it more feasible to observe facial movements as they occur in everyday life, at least in principle (see Box 7, SOM). Automated FACS coding is accurate (>90%) when compared to the AU codes from expert human coders, provided that the images were captured under ideal laboratory conditions, where faces are viewed from the front, are well illuminated, are not occluded, and are posed in a controlled way (Benitez-Quiroz et al., 2016). Under ideal conditions, accuracy is highest (~ 99%) when algorithms are tested and trained on images from the same database (Benitez-Quiroz et al., 2016). The best of these algorithms works quite well when trained and tested on images from different databases (~90%), as long as the images are all taken in ideal conditions (Benitez-Quiroz et al., 2016). Accuracy (compared to human FACS coding) decreases substantially more when coding facial actions in still images or in video frames taken in everyday life where conditions are unconstrained and facial configurations are not stereotypical (e.g.,Yitzhak et al., 2017).[15] For example, 38 automated FACS coding algorithms were recently trained on one million images (the 2017 EmotioNet Challenge; Benitez-Quiroz et al., 2017a) and evaluated against separate test images which were FACS coded by experts.[16] In these less constrained conditions, accuracy dropped below 83% and a combined measure of precision and recall (a measure called F_1, ranging from zero to one) was below .65 (Benitez-Quiroz et al., 2017a).[17] These results indicate that current algorithms are not accurate enough in their detection of facial AUs to fully substitute for expert coders when describing facial movements in everyday life. Nonetheless, these algorithms offer a distinct practical advantage because they can be used in conjunction with human coders to speed up the study of facial configurations in millions of images in the wild. It is likely that automated methods will continue to improve as better and more robust algorithms are developed and as more diverse face images become available.

Measuring an emotional state. Once an approach has been chosen for measuring facial movements, a clear test of the common view of emotional expressions depends on having valid measures that reliably and specifically characterize the instances of each emotion category in a generalizable way, to which the measurements of facial muscle movements can be compared. The methods that scientists use to assess people's emotional states vary in their dependence on **human inference**, however, which raises questions about the validity of the measures.

Relatively objective measures of an emotional instance. The more objective end of the measurement spectrum includes dynamic changes in the autonomic nervous system (ANS), such as cardiovascular, respiratory or perspiration changes (measured as variations in skin conductance), and dynamic changes in the central nervous system, such as changes in blood flow or electrical activity in the brain. These measures are thought to be more objective because the measurements themselves (the numbers) do not require a human judgment (i.e., the measurements are perceiver-independent). Only the interpretation of the measurements (their psychological meaning) requires human inference. For example, a human observer does not judge whether skin conductance or neural activity increases or decreases; human judgment only comes into play when the measurements are interpreted for the emotional meaning.

Currently, there are no objective measures, either singly or as a pattern, that reliability and uniquely identify one emotion category from another in a replicable way. Statistical summaries of hundreds of experiments, called **meta-analyses**, show for example, that currently there is no relationship between an emotion category, such as anger, and a single, specific set of physical changes in ANS that accompany the instances of that category, even probabilistically (the most comprehensive study published to date is Siegel et al., 2018, but for earlier studies see Cacioppo et al., 2000; Stemmler, 2004; also see Box 8, SOM). In anger, for example, blood pressure can go up, go down, or stay the same (i.e., changes in blood pressure are not consistently associated with anger). And a rise in blood pressure is not unique to instances of anger; it also can occur during a range of other emotional episodes (i.e., changes in blood pressure do not specifically occur in anger and only in anger). [18]Individual studies often find patterns of ANS measures that distinguish an instance of one emotion category from another, but those patterns don't replicate and instead vary across studies, even when studies use the same methods and stimuli, and sample from the same population of participants (e.g., compare findings from Kragel & LaBar, 2013 with Stephens, Christie, & Friedman, 2010). Similar within-category variation is routinely observed for changes in neural activity measured with brain imaging (Lindquist et al., 2012) and single neuron recordings (Guillory & Bujarski, 2014). For example, pattern classification studies discover multivariate patterns of activity across the brain for emotion categories such as anger, sadness, fear, and so on, but these patterns do not replicate from study to study (e.g., Kragel & LaBar, 2015; Saarimäki et al., 2016; Wager et al., 2015; for a discussion, see Clark-Polner et al., 2017). This observed variation does not imply that biological variability during emotional episodes is random, but rather that it may be context-dependent (e.g., yellow and green zones of Figure 1). It may also be the case that current biological measures are simply insufficiently sensitive or comprehensive enough to capture situated variation in a precise way. If this is so, then such variation should be considered *unexplained*, rather than random.

There is a difficult circularity built into these studies that is worth pointing out, and that we encounter again a few paragraphs down: Scientists must use some criterion for identifying when instances of an emotion category are present in the first place (so as to draw conclusions about whether or not emotion categories can be distinguished by different patterns of physical measurements).[19] In most studies that attempt to find bodily or neural "signatures" of emotions, the criterion is a subjective one, either reported by the participants or provided by the scientist, which introduces problems of its own, as we discuss in the next section.

Subjective measures of an emotional instance. Without objective measures to identify the emotional state of a participant, scientists typically rely on the relatively more subjective measures that anchor the other end of the measurement spectrum. The subjective judgments can come from the participants (who complete self-report measures), from other observers (who infer emotion in the participants), or from the scientists themselves (who use a variety of criteria, including commonsense, to infer the presence of an emotional episode). These are all examples of **perceiver-dependent** measurements because the measurements themselves, as well as their interpretation, directly rely on human inference.

Scientists often rely on their own judgments and intuitions to stipulate when an emotion is present or absent in participants (as Charles Darwin did). For example, snakes and spiders are said to evoke fear. So are situations that involve escaping from a predator. Sometimes scientists stipulate that certain actions indicate the presence of fear, such as freezing or fleeing or even attacking in defense. The conclusions that scientists draw about emotions depends on the validity of their initial assumptions. It is noteworthy that when it comes to emotions, scientists use exactly the same categories as non-scientists, which may give us cause for concern, as forewarned by William James (James, 1890, 1894)[20]

Inferences about emotional episodes can also come from other people, for example independent samples of study participants, who categorize the situations in which facial movements are observed. Scientists can ask observers to infer when participants are emotional by having them judge subjects' behavior or tone of voice; for example, see our discussion of Camras et al. (2007) discussed in the section on infants and children, below.

A third common strategy to identify the emotional state of participants is to simply ask them what they are experiencing. Their self-reports of emotional experience then become the criteria for deciding whether an emotional episode is present or absent. Self-reports are often considered imperfect measures of emotion because they depend on subjective judgements and beliefs and require translation into words. In addition, a person can be experiencing an emotional event yet be unaware of it and therefore unable to report on it (i.e., a person can be *conscious but unaware* of their experience and unable to report it), or may be unable to express how they feel using emotion words, a condition known as alexithymia. Despite questions about their validity, self-reports are the most common measure of emotion that scientists compare to facial AUs.

Human inference and assessing the presence of an emotional state. At this point, it should be obvious that any measure of an emotional state, to which measurements of facial muscle movements can be compared, itself requires some degree of human inference; what varies is the amount of inference that is required. Herein lies a problem: To properly test the hypothesis that certain facial movements reliably and specifically express emotion, scientists (ironically) must first make a reverse inference that an emotional event is occurring – that is, they *infer* the emotional instance by observing changes in the body, brain, and behavior (e.g., only if blood pressure consistently and uniquely rises in anger can a rise in blood pressure be used as a marker of anger). Or they infer (a reverse inference) that an event or object evokes an instance of a specific emotion category (e.g., an electric shock elicits fear but not irritation, curiosity, or uncertainty). These reverse inferences are scientifically sound *only* if measures of emotion reliably, specifically and validly characterize the instances of the emotion category. So, any clear, scientific test of the common view of emotional expressions rests on a set of more basic inferences about whether an emotional episode is present or absent, and any conclusions that come from such a test are only as sound as those basic inferences.

If all measures of emotion (to which measurements of facial muscle movements are compared) rest on human judgment to some degree, then, in principle, this prevents a scientist from being sure that an emotional state is present, which in turn limits the validity of any experiment designed to test whether a facial configuration validly expresses a specific emotion category. All face-emotion associations that are observed in an experiment reflect **human consensus**, i.e., the degree of **agreement** between self-judgments (of the participants), expert-judgments (of the scientist), and/or judgments of other observers (of perceivers who are asked to infer emotion in the participants). These types of agreement are often incorrectly referred to as **accuracy**. We touch on this point again when we discuss studies that test whether certain facial configurations are routinely perceived as expressions of anger, disgust, fear, and so on.

Testing the common view of emotional expressions: Interpreting the scientific observations. If a specific facial configuration *reliably* expresses instances of a certain emotion category in any given experiment, then we would expect measurements of the face (e.g., facial AU codes) to co-occur with measurements that indicate that participants are in the target emotional state. In principle, those measures might be more objective, such as ANS changes during an emotional event, or they might be more subjective, deriving from the scientist, from other perceivers who make judgments about the study participants, or from the participants themselves. In practice, however, most experiments compare facial movements to subjective measures of emotion -- a scientist's judgment about which emotions are evoked by a particular stimulus, perceivers judgments about participants' emotional states, or participants' self-reports of emotional experience -- because ANS and other more objective measurements do not themselves distinguish one emotion category from another in a reliable and specific way. For example, in an experiment, scientists might ask: Do the AUs that create a scowling facial configuration co-occur with self-reports of feeling angry? Do the AUs that create a pouting facial configuration co-occur with perceiver's judgments that participants are sad? Do the AUs that create a wide-eyed gasping facial configuration co-occur when people are exposed to an electric shock? And so on. If such observations suggest that a configuration of muscle movements is reliably observed during episodes of a given emotion category, then those movements are said to express the emotion in question. As we will see, many studies show that some facial configurations occur more often than random chance, but are not observed with a high degree of reliability (according to the criteria from Haidt & Keltner (1999), outlined in Table 2 and Figure 3).

If a specific facial configuration specifically (i.e., *uniquely*) expresses instances of a certain emotion category in any given experiment, then we would expect to observe little co-occurrence between measurements of the face and measurements indicating the presence of emotional instances from *other* categories, except what would be expected by chance (again, see Table 2 and Figure 3). For example, in an experiment, scientists might ask: do the AUs that create a scowling facial configuration co-occur with self-reports of feeling sad, confused, or social motives such as dominance? Do the AUs that create a pouting facial configuration co-occur with perceiver's judgments that participants are angry or afraid? Do the AUs that create a wide-eyed gasping facial configuration co-occur when people are exposed to a competitor whom they are trying to scare? And so on.

If a configuration of facial movements is observed in instances of a certain emotion category in a reliable, specific way within an experiment, so that we can infer that the movements are expressing an instance of the emotion in that study as hypothesized, then scientists can safely infer that the facial movements in question are *an* expression of that emotion category's instances in that situation. One more step is required before we can infer that the facial configuration is *the* expression of that emotion: we must observe a similar pattern of facial configuration-emotion co-occurrences across different experiments, to some extent generalizing across the specific measures and methods used and the participants and contexts sampled. If the facial configuration-emotion co-occurrences *replicate* across experiments that sample people from the same culture, then the facial configuration in question can be reasonably be referred to as *an* emotional expression only in that culture; e.g., if a scowling facial configuration co-occurs with measures of anger (and only anger) across most studies conducted on adult participants in the US who are free from illness, then it is reasonable to refer to a scowl as an expression of anger in the US. If facial configuration-emotion co-occurrences generalize across cultures – that is, replicate across experiments that sample a variety of instances of that emotion category in people from different cultures -- then the facial configuration in question can be said to **universally express** the emotion category in question.

Studies of Healthy Adults from the U.S. and Other Developed Nations

We now review the scientific evidence from studies that document how people spontaneously move their facial muscles during instances of anger, disgust, fear, happiness, sadness and surprise, and how they pose their faces when asked to indicate how they express each emotion category. We examine evidence gathered in the lab and in naturalistic settings, sampling healthy adults who live in a variety of cultural contexts. To evaluate the reliability, specificity and generalizability of the scientific findings, we adapted criteria set out by Haidt & Keltner (1999), as discussed in Table 2.

Spontaneous facial movements in laboratory studies. A meta-analysis was recently conducted to test the hypothesis that the facial configuration in Figure 4 co-occur, as hypothesized, with specific emotion categories (Duran et al., 2017). This analysis was published in a book chapter. Thirty-seven published articles reported on how people moved their faces when exposed to objects or events that evoke emotion. Most studies included in the meta-analysis were conducted in the laboratory. The findings from these experiments were statistically summarized to assess the reliability of facial movements as expressions of emotion (see Figure 5). In all emotion categories tested, other than fear, participants moved their facial muscles into the expected configuration more consistently than what we would expect by chance. Consistency levels were weak, however, indicating that the proposed facial configurations in Figure 4 have limited reliability (and to some extent, limited generalizability; i.e., a scowling facial configuration is an expression of anger, but not *the* expression of anger. More often than not, people moved their faces in ways that were not consistent with the hypotheses of the common view. An expanded version of this meta-analysis (Duran & Fernandez-Dols, 2018) analyzed 89 effect sizes from 47 studies totaling 3599 participants, with similar results: the hypothesized facial configurations were observed, with average effect sizes of r = .32 (for the average correlation between the intensity of a facial configuration and a measure of emotion, with correlations for specific emotion categories ranging from .25 to .38, corresponding to weak evidence of reliability) and proportion = .19 (for the average proportion of the times that a facial configuration was observed during an emotional event, with proportions for specific emotion categories ranging from .15 to .25, interpreted as no evidence to weak evidence of reliability).[21]

No overall assessment of specificity was reported in either the original or the expanded meta-analysis because most published studies do not report the **false positive rate** (i.e., the frequency with which a facial AU is observed when an instance of the hypothesized emotion category was not present; see Figure 3). Nonetheless, some striking examples of specificity failures have been documented in the scientific literature. For example, a certain smile, called a “Duchenne” smile, is defined in terms of facial muscle contractions (i.e., in terms of **facial morphology**): it involves movement of the *orbiculari oculis* which raises the cheeks and causes wrinkles at the outer corners of the eyes in addition to movement of the *zygomatic major* which raises the corners of the lips into a smile. A Duchenne smile is thought to be a spontaneous expression of authentic happiness. Research shows that a Duchenne smile can be intentionally produced when people are not happy, however (Gunnery & Hall, 2014; Gunnery et al., 2013; also see Krumhuber & Manstead, 2009), consistent with evidence that Duchenne smiles often occur when people are signaling submission or affiliation rather than solely reflecting *happiness* (Rychlowska et al., 2017).

Spontaneous facial movements in naturalistic settings. Studies of facial configuration-emotion category associations in naturalistic settings tend to yield similar results to studies that were conducted in more controlled laboratory settings (Fernandez-Dols, 2017; Fernandez-Dols & Crivelli, 2013). Some studies observe that people express emotions in real world settings by spontaneously making the facial muscle movements proposed in Figure 4, but such observations do not replicate well across studies (e.g., compare Matsumoto & Willingham, 2006 vs. Crivelli, Carrera and Fernandez-Dols, 2015; Rosenberg & Ekman, 1994 vs. Fernandez-Dols, Sanchez, Carrera, & Ruiz-Belda, 1997). For example, two field studies of winning judo fighters recently demonstrated that so-called "Duchenne" smiles were better predicted by whether an athlete was interacting with an audience than the degree of happiness reported after winning their matches (Crivelli, Carrera, & Fernandez-Dols, 2015). Only eight of the 55 winning fighters produced a "Duchenne" smile in Study 1; all occurred during a social interaction. Only 25 out of 119 winning fighters produced a "Duchenne" smile in Study 2, documenting, at best, weak evidence for reliability.

Posed facial movements. Another source of evidence comes from asking participants sampled from various cultures to deliberately pose the facial configurations that they *believe* they use to express emotions. In these studies, participants are given a single emotion word or a single, brief statement to describe each emotion category and then asked to freely pose the expression that they believe they make. In this way, they directly examine common beliefs about emotional expressions. For example, one study provided college students from Canada and Gabon (in Central Africa) with dictionary definitions for ten emotion categories. After practicing in front of a mirror, participants posed the facial configurations so that “their friends would be able to understand easily what they feel” and their poses were FACS coded (Elfenbein et al., 2007, p. 134). Similarly, a recent study asked college students in China, India, Japan, Korea, and the US, to pose the facial movements they believe they make when expressing each of 22 emotion categories (Cordaro, Sun, Keltner, Kamble, Huddar & McNeil, 2017). Participants heard a brief scenario describing an event that might cause anger (“You have been insulted, and you are very angry about it”) and then were instructed to pose a facial (and non-verbal, vocal) expression of emotion, as if the events in the scenario were happening to them. Experimenters were present in the testing room as participants posed their responses. Both studies found moderate to strong evidence for a cross-cultural, common expressive pose for anger, fear, and surprise categories, and weak to moderate evidence for the happiness category, with cultural variation around those common poses; the findings were weaker for disgust and sadness categories (Figure 6).

Neither study compared participants’ posed expressions to observations of how they *actually* moved their faces when expressing emotion. Nonetheless, a quick comparison of the findings from both studies and the proportions of spontaneous facial movements made during emotional events (from the Duran et al. (2017) meta-analysis) makes it clear that posed and spontaneous movements differ, sometimes quite substantially (again, see Figure 6). When people pose a facial configuration that they believe expresses an emotion category, they make facial movements that more reliably agree with the hypothesized facial configurations in Figure 6. The same cannot be said of people's spontaneous facial movements during actual emotional episodes, however (for convergent evidence, see Motley & Camden, 1988; Namba et al., 2016). One possible interpretation of these findings is that posed and spontaneous facial muscle configurations correspond to distinct communication systems. Indeed, there is some evidence that volitional and involuntary facial movements are controlled by different circuits in the skeletomotor system (Rinn, 1984). Another factor that may contribute to the discrepancy between posed and spontaneous facial movements is that people's beliefs about their own behavior often reflect their stereotypes or beliefs and do not necessarily correspond to how they actually behave in real life (see Robinson & Clore, 2002).

Summary. Our review of the available evidence thus far is summarized in the first through third data rows in Table 4. The hypothesized facial configurations presented in Figure 4 spontaneously occur with weak reliability during instances of the predicted emotion category, suggesting that they sometimes serve to express the predicted emotion. Furthermore, the specificity of each facial configuration as an expression of a specific emotion category is *largely unknown* (because it is typically not reported in many studies). In our view, this pattern of findings is most compatible with the interpretation that hypothesized facial configurations are not made reliably or specifically enough to use them to infer a person's emotional state. We are not suggesting that facial movements are meaningless and devoid of information. Instead, the data suggest that the meaning of any set of facial movements may be much more variable and context-dependent than hypothesized by the common view.

Studies of Healthy Adults Living in Small-Scale, Remote Cultures

The emotion categories that are at the heart of common view– anger, disgust, fear, happiness, sadness and surprise -- were derived from modern US English (Wierzbicka, 2014) and their proposed expressions (in Figure 4) derive from observations of people who live in urbanized, Western settings. Nonetheless, it is hypothesized that these are facial configurations evolved as emotion-specific expressions to signal socially-relevant emotional information (Shariff & Tracy, 2011) in the challenging situations that originated in our hunting and gathering hominin ancestors who lived on the African savannah during the Pleistocene era (Pinker, 1997; Tooby & Cosmides, 1990). It is further hypothesized that these facial configurations should therefore be observed during instances of the predicted emotion categories with strong reliability and specificity in people around the world, although the facial movements might be slightly modified by culture (Cordaro et al., 2017; Ekman, 1972). The strongest test of these hypotheses would be to sample participants who live in remote parts of the world with relatively little exposure to western cultural norms, practices and values (Norenzayan & Heine, 2005; Henrich et al., 2010) and observe their facial movements during emotional episodes.[22] In our evaluation of the evidence, we continued to use the criteria summarized by Haidt & Keltner (1999; see Table 2).

Spontaneous facial movements in naturalistic settings. Our review of scientific studies that systematically measure the spontaneous facial movements in people of small-scale, remote cultures is brief by necessity: there aren't any. At the time of publication, we were unable to identify even a single published report or manuscript registered on open-access, pre-print services that measured facial muscle movements in people of remote cultures as they experienced emotional events. Scientists have almost exclusively observed how people *label* facial configurations as emotional expressions (i.e., they study emotion perception, not production) to test the hypothesis that certain facial configurations evolved to express certain emotion categories in a reliable, specific and generalizable (i.e., universal) manner. Later in the paper we return to this issue and discuss the findings from these emotion perception studies.

There are nonetheless several descriptive reports that provide support for the common view of universal emotional expressions (similar to what Valente et al., 2017 refer to as an "observational approach"). For example, the US psychologist Paul Ekman and colleagues curated an archive of photographs of the Fore hunter-gatherers taken during his visits to Papua New Guinea in the 1960s (Ekman,1980). The photographs were taken as people went about their daily activities in the small hamlets of the eastern highlands of Papua New Guinea. Ekman used his knowledge of the situation in which each photograph was taken to assign each facial configuration to an emotion category, leading him to conclude that the Fore expressed emotions with the proposed facial configurations shown in Figure 4. Yet different scientific methods yielded a contrasting conclusion. When Trobriand Islanders living in Papua New Guinea were asked to infer emotions in facial configurations by labeling these photographs in their native language, both by freely offering words and by choosing the best fitting emotion word from a list of nine choices, they did not label the facial configurations as proposed by Ekman and colleagues at above chance levels (Crivelli et al., 2017).[23] In fact, the proposed fear expression -- the wide-eyed gasping face -- is actually interpreted as an expression of threat (intent to harm) and anger by the Maori of New Zealand and in the Trobriand Islanders in remote Papua New Guinea (Crivelli & Fridlund, 2016).

A compendium of spontaneous human behavior published by the Austrian ethologist Irenäus Eibl-Eibesfeldt (Eibl-Eibesfeldt, 1989) is sometimes cited as evidence for the hypothesis that certain facial movements are universal signals for specific emotion categories. No systematic coding procedure was used in his investigations, however. Upon close examination, Eibl-Eibesfeldt's detailed descriptions appear to be more consistent with the studies of people from more industrialized cultures that we reviewed above: people move their faces in a variety of ways during episodes belonging to the same emotion category. For example, as reported by Eibl-Eibesfeldt, a rapid eyebrow raise (called an eyebrow flash) is thought to express friendly recognition in some, but not all, cultures. This movement would be coded with FACS AU 1 (inner brow raise) and AU 2 (outer brow raise) that are part of the proposed expressions for surprise and fear (Ekman et al., 1983), sympathy (Haidt & Keltner, 1999) and awe (Shiota et al., 2003). Even Eibl-Eibesfeldt acknowledged that eyebrow flashes were not unique expressions of specific emotion categories, writing that they also served as a greeting, to invite social contact, as a sign of thanks, an initiation of flirting, and a general indication of "yes" in Samoans and other Polynesians, in the Eipo and Trobriand islanders in Papua New Guinea, and in the Yanomami of South America. In Japan, eyebrow flashes are considered an impolite way for adults to greet one another. In the US and Europe, an eyebrow flash was observed when friends greet one another, but not strangers.

Posed facial movements. One study read a brief emotion story to people who live in the remote Fore culture of Papua New Guinea and asked each person to "show how his face would appear" if he was the person described in the emotion stories (Ekman, 1972, p. 273; sample size was not reported). Videotapes of nine participants were shown to 34 US college students who were asked to judge which emotion was being expressed. US participants were asked to infer the emotional meaning of the facial poses by choosing an emotion word from six choices provided by the experimenter (called a choice-from-array task, see Table 5). Participants inferred the intended emotional meaning above chance guessing for smiling (happiness, 73%), frowning (sadness, 68%), scowling (anger, 51%), and nose-wrinkling (disgust, 46%), but not for surprise and fear (27% and 18% respectively).

Summary. Our review of the available evidence from expression production studies in small-scale, remote cultures is inconclusive because there are no systematic, controlled observations that examine how people who live in these cultural contexts spontaneously move their facial muscles during emotional episodes. The evidence that does exist suggests that common beliefs about emotion may share some similarities across urban and small-scale cultural contexts, but more research is needed before any interpretations are warranted. These findings are summarized in the fourth and fifth data rows of Table 4.

Studies of Healthy Infants and Children

The facial movements of infants and young children provide a valuable way to test common beliefs about emotional expressions because, unlike older children and adults, babies cannot exert voluntary control over their spontaneous expressive behaviors, meaning that they are unable to deliberately mask or portray instances of emotion in accordance with social demands. As a general rule, infants understand far more about the world than what they can easily convey through their physical actions, making it difficult for experiments to distinguish between what infants *understand*, which often exceeds what they can actually *do*. Experiments must use human inference to determine when an infant is in an emotional state, as is the case in studies of adults (see *Human inference and assessing the presence of an emotional state*). The presence (or absence) of an instance of emotion is inferred (i.e., stipulated), either by a scientist (who exposes a child to something that is presumed to evoke an emotion episode) or by adult “raters” who infer the emotional meaning of the evoking situation or the child’s body movements and vocalizations (see *Subjective measures of an emotional instance*). In the latter cases, inferences are measured by asking research participants to label the situation or the child's emotional state by choosing an emotion word or image from a small set of options, a task known as choice-from-array. We address the strengths and weaknesses of choice-from-array tasks (see Table 6) and the potential risk of confirmatory bias with the use of such methods (see *Some observations on interpreting the data*, below).

With such a strong reliance on human inference, there is a risk that scientists will implicitly confound the measurements made in an experiment with their interpretation of those measurements, in effect over-interpreting infant behavior as reflecting a specific aspect of an emotional event, in part because these young research participants cannot speak for themselves. Some early and influential studies confound the observation of facial movements with their interpreted emotional meaning, leading to the conclusions that babies as young as 7-months of age were capable of producing an expression of anger when, in fact, it is more scientifically correct to say that the babies were scowling. For example, in one study, infants' facial movements were coded as they were given a cookie, and then the cookie was taken away and placed out of reach although still clearly visible. The babies appeared to scowl when the cookie was removed and not when it was in their mouths (Stenberg, Campos & Emde, 1983). It is certainly possible that this repeated giving and taking away of the treat angered the infants, but the babies might also have been confused or just generally distressed. Without some independent evidence to indicate that a state of anger was induced, we cannot confidently conclude that certain facial movements in an infant reliably express a specific instance of emotion.

The Stenberg et al. study illustrates some of the design issues that have historically been of concern in many studies with infants. First, emotion-inducing situations are often defined with commonsense intuitions rather than objective evidence (e.g., an infant is assumed to become angry when a cookie is taken away). In fact, it is difficult to know how any individual infant at any point in time will construct and react to such an event. Second, when an infant produces a facial movement, a common assumption is used to infer its emotional meaning without additional measures or controls (e.g., when a scowling facial configuration is observed, it is assumed to necessarily be an expression of infant anger, even if there are no data to confirm that a scowl is specific to instances of anger in an infant). In fact, years later, Campos and his team revised their earlier interpretation of their findings as their research program progressed, later concluding that the facial movements in question (infants lowering and drawing together their brows, staring straight ahead, or pressing their lips together) were more generally associated with unpleasantness and distress, and were not reliable expressions of anger (e.g., Camras, Oster et al., 2007).

The inference problem is particularly poignant when fetuses are studied. For example, a study that used 4-D ultrasonography observed 20-week-old fetuses knitting their brows and described the facial movements as expressions of *distress* (Dondi et al., 2014). Yet the fetuses were producing these facial movements during situations when fetal distress was unlikely. The brow-knitting was observed during noninvasive ultrasound scanning that did not involve perturbation of the fetus and the pregnant women were at rest. Furthermore, the scans were brief in duration and the facial movements were interspersed with other movements that are typically not thought to express negative emotions, such as smiling and mouthing. This is an example of making a scientific inference about an emotion occurring based solely on the facial movements without converging evidence that the organism in question (a fetus) was in a distressed state.

Doing so highlights the common but unsound assumption that certain facial movements reliably index instances of the same emotion category.

The study of expression production in infants and children must deal with other design challenges, in addition to the reliance on human inference, that are shared by experiments employing adult participants. In particular, most experiments observe facial movements in a restricted range of laboratory settings rather than in the wide variety of situations that naturally occur in everyday life. The frequent use of only a single stimulus or event to observe facial movements for each emotion category limits the opportunity to discover whether the expression of an emotion category vary systematically with context.

Even with these design considerations, the scientific findings from studies of infants and children parallel those that we encountered from studies on adults: lack of reliability and specificity in facial muscle movements is the norm, not the exception (again, according to the Haidt & Keltner (1999) criteria in Table 2). Although some older studies concluded that infants produce invariant emotional expressions (e.g., Izard, Hembree, Dougherty, & Spirrizi, 1983; Izard et al., 1987; Izard et al., 1995; Lewis, Ramsay & Sullivan, 2006), these conclusions have been largely overturned by more recent work and in many cases have been reinterpreted and revised by the authors themselves (e.g., Lewis, Ramsay & Sullivan, 2006)..

Facial movement in fetuses, infants and young children. The most detailed research on facial movements in fetuses and newborns has focused on smiles. Human fetuses lower their brows (AU4), raise their cheeks (AU6), wrinkle their noses (AU9), crease their nasolabia (AU11), pull the corners of their lips (AU12), show their tongues (AU19), part their lips (AU25), and stretch their mouths (AU27) -- all of which have been implicated, to some degree, in adult laughter. Infants sometimes produce facial movements that resemble adult laughter when they are in distress and pain (Dondi et al., 2014; Hata et al., 2013; Reissland et al., 2011; Reissland, Francis, & Mason, 2013; Yan et al., 2006). Within 24 hours of birth, infants raise their cheek muscles in response to being touched (Cecchini et al., 2011). But these movements are not specific to smiling; neonates also raise their cheeks (contract the zygomatic muscle) during rapid eye movement (REM) sleep, when drowsy, and during active sleep (Dondi et al., 2007). A neonatal smile with raised cheeks is caused by brainstem activation (Rinn, 1984), reflecting internally generated arousal rather than expressing or communicating an emotion or even a more general feeling of pleasure (Emde & Koenig, 1969; Sroufe, 1996; Wolff, 1987). So, it remains unclear whether fetal or neonatal facial muscle movements have any relationship to specific emotional episodes, as well as more generally to pleasant feelings or to other social meanings (Messinger, 2002).

In fact, it's not clear that fetal and neonatal facial movements always have a psychological meaning (consistent with a behavioral ecology view of facial movements; Fridlund, 2017). Newborns appear to produce some combinations of facial movements for muscular reasons. For example, infants produce facial movements associated with the proposed expression for "surprise" (open mouth and raised eyebrows) in situations that are unsurprising, just because opening the mouth necessarily raises their eyebrows; conversely, infants do not consistently show the proposed expressive configuration for surprise in contexts that are likely to be surprising (Camras, 1992; Camras et al., 2017). The facial movement that is part of the proposed expression for *sadness* (brows oblique and drawn together) occurs when infants attempt to lift their heads to direct their gaze (Michel, Camras, & Sullivan, 1992).

In addition, newborns produce many facial movements that co-occur with fussiness, distress, focused attention, and distaste (Oster, 2005). Newborns react to being given sweet versus sour liquids; for example, newborns make a nose-wrinkle movement, which is part of the proposed expressive configuration for disgust, when given a sour liquid (Granchrow et al., 1983). However, other studies show that newborns also make this facial movement when given sweet, salty, sour and bitter tastes (e.g., Rosenstein & Oster, 1988). Still other studies show that nose-wrinkling does not always occur when infants taste lemon juice (i.e., when that facial movement is expected; Bennett et al., 2002). More generally, infants rarely produce consistent facial movements that cleanly map onto any single emotion category. Instead, infants produce a variety of facial configurations, indicating a lack of emotional specificity (Matias & Cohn, 1993).

There are further examples that illustrate how infant facial movements lack strong reliability and specificity. In a study of 11-month old babies from the US, China and Japan, infants saw a toy gorilla head that growled (to induce fear) or their arms were restrained (to induce anger; Camras et al., 2007). Observers judged the infants to be fearful or angry based on their body movements; yet, the infants produced the same facial movements in the two situations.[24] In another study, one-year-old infants were videotaped in situations where they were tickled (to elicit joy), tasted sour flavors (to elicit disgust), watched a jack-in-the box (to elicit surprise), had their arm restrained (to elicit anger), and were approached by a masked stranger (to elicit fear) (Bennett, Bendarsky, and Lewis, 2002). Infants whose arms were restrained (to purportedly induce an instance of anger) produced the facial actions associated with the proposed facial configuration for an anger expression only 24 percent of the time (low reliability), and instead 80 infants (54%) produced the facial actions proposed as the expression of surprise, 37 infants (25%) produced the facial actions proposed as the expression of joy, 29 infants (19%) produced the facial actions proposed as the expression of fear, and 28 (18%) produced the facial actions proposed as the expression of sadness. This dramatic lack of specificity was observed for all emotion categories studied. An equal number of babies produced facial movements that are proposed as the expressions of joy, surprise, anger, disgust, and fear categories when a sour liquid was placed on infants' tongues to elicit disgust. When infants faced a masked stranger, only 20 (13%) produced facial movements that correspond to the proposed expression for fear, compared to 56 infants (37%) who produced facial actions associated with the proposed expression for instances of joy.[25]

Taken together, these findings suggest that infant facial movements may be associated with the **affective features** of experience, such as distress or arousal, as originally described by Bridges (1932), or communicate a desire to approach or avoid something (e.g., Lewis, Sullivan & Kim, 2015). Affective features such as valence (ranging from pleasantness to distress) and arousal (ranging from activated to quiescent) are continuous properties of consciousness, just as approach and avoidance are continuous properties of action. These affective features are shared by many instances of different emotion categories, as well as with mental events that are not considered emotional (as discussed in Box 9, in SOM) but are still effective and important for infants.[26] Over time, infants likely learn to differentiate mental events with simple affective features into episodes of emotion with additional psychological features that are specific to their socio-cultural contexts, making them maximally effective at eliciting needed responses from their caregivers (Barrett, 2017a; Holodynski & Friedlmeier 2006; Weiss & Nurcombe, 1992; Witherington et al., 2008).

The affective meaning of an infant's facial movements may, in fact, be the very properties that make these movements so salient to adult observers. When infants move their lips, open their mouths, or constrict their eyes, adults view infants as feeling more positively or negatively depending upon the context (Bolzani et al., 2005). Infant expressions thus do have a reliable link to instrumental effects in the adults who observe them – playing an important role in parent-infant interaction, attachment and the beginnings of social communication (Atzil et al., 2018; Feldman, 2016). For example, if an infant cries with narrowed eyes, adults rate that infant's emotion as more negative or having an unwanted experience or needing help, but if the infant makes that same eye movement while smiling, adults interpret the infant as experiencing more positive emotion. These data consistently point to the usefulness of facial movements in the communication of arousal and valence (properties of affect; Box 9, SOM). Even when episodes of more specific emotions start to emerge, we don't yet have evidence that facial movements map reliably and regularly to a specific emotion category.

Young children begin to produce adult-like facial configurations *after* the first year of life. Even then, however, children's facial movements continue to lack strong reliability and specificity (Bennett et al., 2002; Camras & Shutter, 2010; Matias & Cohn, 1993; Oster, 2005). Examples of a wide-eyed gasping facial configuration, proposed as the expression of fear (see Figure 4), have rarely been observed or reported in young infants (Witherington et al., 2010). Nor do infants reliably produce a scowling facial configuration, proposed as the expression of anger (again, see Figure 4). Infants scowl when they cry or are about to cry (Camras, Fatani, Fraumeni & Shuster, 2016). A frown (mouth corner depression, AU15) is not reliably and specifically observed when infants are frustrated (Lewis & Sullivan, 2014; Sullivan et al., 2003). A smile (cheek raising and lip corner pulling, AU6 and AU12) is not reliably observed when infants are in visually engaging or mastery situations, or even when they are in pleasant social interactions (Messinger, 2002).

Experiments that observe young children's facial movements in naturalistic settings find largely the same results as those conducted in controlled laboratory settings. For example, one study trained ethnographic videographers to record a family's daily activities over four days (Sears et al., 2014). Coders judged whether or not the child from each participating family made a scowling facial configuration (referred to as an expression of anger), a frowning facial configuration (referred to as an expression of sadness), and so on, for the six (presumed) emotion categories included in the study -- happiness, sadness, surprise, disgust, fear, and anger. During instances that were coded as anger (defined as situations that included verbal disagreements/sibling bickering, requests for compliance and/or reprimands from parents, parent refusal of child requests, during homework, and sibling provocation), a variety of facial movements were observed, including frowns, furrowed brows, and eye-rolls, as well as a variety of vocalizations, including shouts and whining, and both nonaggressive and aggressive physical behaviors. Perhaps the most telling observations for our purposes is that expressions of anger were more often vocal than facial. During anger situations, children raised their voices 42% of the time, followed by whining about 21% of the time. By contrast, children made scowling facial configurations only 16.2% of the time.[27] Yet even during anger situations, the facial movements were predominantly frowning, which can be part of many different proposed facial configurations. The authors reasoned that children engage in specific behaviors to obtain specific goals, and that behaviors such as whining are more likely to attract attention and possibly change parental behavior than will a facial movement. Indeed, it is easier for parents to ignore a negative facial expression than a whining child in the room! Similar findings for low reliability and specificity of the facial configurations presented in Figure 4 were recently observed in a naturalistic study that videotaped seven to nine-year old children and their mothers discussing a conflict during their visit to the laboratory related to homework, chores, bedtime or interactions with siblings (Castro et al., 2017).

Summary. Newborns and infants react to the world around them with facial movements. There is not yet sufficient evidence, however, to conclude that these facial movements reliably and specifically express the instances of any specific emotion category (findings summarized in Table 4). When considered alongside vocalizations and body movements, there is consistent evidence that infant facial movements reliably signal distress, interest and arousal, and perhaps serve as a call for help and comfort. In young children, instances of the same emotion category appear to be expressed by a variety of different muscle movements, and the same muscle movements occur during instances of various emotion categories, and even during non-emotional instances. It may be the case that reliability and specificity emerges through learning and development (see Box 10, in SOM), but this remains an open question that awaits future research.

Studies of Congenitally Blind Individuals

Another source of evidence to test the common view comes from observations of facial movements in people who were born blind. The assumption is that people who are blind cannot learn, by watching others, which facial muscles to move when expressing emotion. Based on this assumption, several studies have claimed to find evidence that congenitally blind individuals express emotions with the hypothesized facial configurations in Figure 4 (e.g., blind athletes show expressions that are reliably interpreted as shame and pride, Tracy & Matsumoto, 2008; see also Matsumoto & Willingham, 2009). People who are born blind learn through other **sensory modalities**, however (for a review, see Bedny & Saxe, 2012), and therefore can learn whatever regularities exist between emotional states and words for facial movements from hearing descriptions in conversation, in books and movies, and by direct instruction.[28] As an example of such learning, Olympic athletes who win medals smile only when they know they are watched by other people, such as when they are on the podium facing the audience; in other situations, such as while waiting behind the podium or while on the podium facing away from people but towards a flag, they did not smile (but presumably were still very happy; Fernandez-Dols et al., 1995). Such findings are consistent with the behavioral ecology view of facial expressions, Fridlund, 1991, 2017) and with more recent sociological evidence that smiles are social cues that can communicate different social messages depending on the cultural context (Martin, Rychlowska, Wood and Niedenthal, 2017).

The limitations that apply to studies of emotional expressions in sighted individuals, reviewed throughout this paper, are even more applicable to scientific studies of emotional expressions in the blind.[29] Participants are given pre-determined emotion categories that shape their possible responses, and facial movements are often quantified by human judges who have their own biases when making commonsense judgments (e.g., Galati et al., 1997; Galati et al., 2001; Valente et al., 2017). In addition, people who are blind make additional, often unusual movements of the head and the eyes (Chiesa et al., 2015). For example, people who are blind from birth often move their head in unusual ways to better hear objects or echoes. These unusual movements might interfere with or contaminate expressive facial movements. More importantly, they reveal whether a participant is blind or sighted, and this knowledge can bias human raters who are judging the presence or absence of facial movements in emotional situations.

Helpful insights about the facial expressions of congenitally blind individuals comes from a recent review (Valente et al., 2017) that surveyed 21 studies published between 1932 and 2015. These studies observe how blind participants move their faces during instances of emotion and then compared those movements both to the proposed expressive forms in Figure 4 and to the facial movements of sighted people. Both spontaneous facial movements and posed movements were tested. Eight older studies (published between 1932-1977) reported that congenitally blind individuals spontaneously expressed emotions with the proposed facial configurations in Figure 4, but Valente et al. (correctly) questioned the objectivity of these studies because the data were largely based on subjective impressions offered by researchers or their assistants. The 13 studies published between 1980 and 2015 were better designed: they videotaped participants' facial movements and described them using a formal facial coding system like FACS for adults or a similar coding system for children. These studies are too few in number and have insufficient sample sizes to conduct a formal meta-analysis, but taken together they suggest that, in general, congenitally blind individuals *spontaneously* moved their faces in similar ways to sighted individuals during instances of emotion: both groups expressed instances of anger, disgust, fear, happiness, sadness or surprise with the proposed expressive configurations (or their individual AUs) in Figure 4 with either weak reliability or no reliability at all, and neither group produced any of the configurations with any specificity (e.g., Galati et al., 2001; Galati et al., 2003; Galati et al., 1997). The lack of specificity is not surprising given that, upon closer inspection, several of the studies discussed in Valente et al. (2017) compared emotion categories that systematically differ in their prototypical affective properties, contrasting facial movements in pleasant vs. unpleasant circumstances (e.g., Cole et al., 1989), or observed facial movements only in pleasant circumstances without distinguishing the facial AUs for the happiness category vs. other positive emotion categories (e.g., Chiesa et al., 2015), such that their findings cannot be interpreted unambiguously as evidence pertaining to emotional expressions, per se.

While congenitally blind and sighted individuals were similar to one another in the variety of their spontaneous facial movements, they differed in their posed facial configurations. After listening to descriptions of situations that were supposed to elicit an instance of anger, sadness, fear, disgust, surprise, and happiness, sited participants posed their faces with the proposed expressive forms for the negative emotion categories in Figure 4 at higher levels of reliability and specificity than did blind participants (Galati et al., 1997; Roch-Levecq, 2006). These findings suggest that congenitally blind individuals have different beliefs about emotional expressions or that their knowledge of social rules for producing those configurations on command differs from those of sighted individuals.

Summary. The evidence from studies of blind individuals is consistent with the other scientific evidence reviewed so far (Table 4). Even in the absence of visual experience, blind individuals, like sighted individuals, develop the ability to spontaneously make a variety of facial movements to express emotion, and those movements do not reliably and specifically configure as proposed by the common view of emotion (depicted in Figure 4). Learning to voluntarily pose the proposed expressions in Figure 4 does seem to covary with vision, however, further emphasizing that posed and spontaneous expressions should be treated as different phenomena. Further scientific attention is warranted to examine how congenitally blind individuals learn, via other sensory modalities, to express emotions.

Summary of Scientific Evidence on the Production of Facial Expressions

The scientific findings we have reviewed thus far – dealing with how people actually move their faces during emotional events – does not strongly support the common view that people reliably and specifically express instances of emotion categories with spontaneous facial configurations that resemble those proposed in Figure 4. Adults around the world, infants and children and congenitally blind individuals all show much more variability than commonly hypothesized. Studies of posed expressions further suggest that particular facial movements are linked to particular emotions more by consensus and beliefs, rather than by scientific evidence for "emotion expression." Consequently, the commonly used phrases such as "emotional facial expression," "emotional expression" or "emotional display" are misleading. More neutral phrases that assume less, such as "facial configuration" or "pattern of facial movements" or even "facial actions," should be used instead.

We next turn our attention to the question of whether people reliably and specifically infer certain emotions from certain patterns of facial movements, shifting our focus from studies of production to studies of perception. It has been long assumed that emotion perception provide an indirect way of testing the common view of emotion production, because facial expressions, when they are assumed to be displays of internal emotional states, are thought to have co-evolved with the ability to recognize and read them (Ekman, Friesen, & Ellsworth, 1972). For example, Shariff and Tracy (2011) have suggested that emotional expression (production) and emotion perception likely co-evolved as an integrated signaling system (for additional discussion, see Jack, Sun, Delis, Garrod, & Schyns, 2016).[30] In the next section, we review the scientific evidence on emotion perception.

Perceiving Emotions from Facial Movements: A Review of the Scientific Evidence

For over a century, an active line of research has directly examined whether people reliably and specifically infer emotional meaning in the facial configurations presented in Figure 4. Most of these studies are interpreted as evidence for people's ability to ***recognize*** *or* ***decode*** emotion in facial configurations, on the assumption that the configurations broadcast or signal emotional information to be recognized or detected. This is yet another example of confusing what is known and what is being tested. A more correct interpretation is that these studies indicate whether people reliably and specifically *infer* or *judge* emotion in those facial configurations. This pervasive confusion in the scientific literature may explain why very few studies have actually investigated *the processes by which* people detect the onset and offset of facial movements and infer emotions in those movements (i.e., few studies consider the mechanisms by which people infer emotional states from detecting and perceiving facial movements) (for discussion, see Martinez, 2017a, 2017b). In this section, we first review the design of typical emotion perception experiments that are used to test the common view that emotions can be reliably and specifically "read out" from facial movements. We also examine whether people infer emotions from the facial movements in dynamic, computer-generated faces, a class of studies that offer a more **data-driven** way to study emotion perception, and in virtual humans, which provides the opportunity for a more **implicit approach** to studying emotion perception.

The Anatomy of a Typical Experiment Designed to Observe Whether People Reliably and Specifically Infer Emotion in Facial Movements

For a person – a perceiver -- to infer that another person is in an emotional state by looking at that person's facial movements, the perceiver must have many competencies. People move their faces continuously (i.e., real human faces are never still), so a perceiver must *notice* or **detect** the relevant facial movements in question and **discriminate** them from other facial movements (that is, the perceiver must be able to set a perceptual boundary to know when the movements begin and end, and, for example, that a scowl is different from a sneer). The perceiver must be able to **identify** (or segment) the movements as an ensemble or pattern (i.e., bind them together and distinguish them from other movements that are normally inferred to be irrelevant). And the perceiver must be able to infer similarities and differences between different instances of facial movements, as specified by the task (e.g., categorize a group of facial movements as instances expressing anger, fear, etc.). This categorization might involve merely labeling the facial movements, referred to as **action identification** (describing how a face is moving, such as smiling) or it might involve inferring that a particular mental state caused the actions, referred to as **mental inference** or mentalizing (inferring why the action is performed, such as a state of happiness; Vallacher & Wegner, 1987). In principle, the categorization could also involve inferring a situational cause for the actions, but in practice, this question is rarely investigated in studies of emotion perception. The overwhelming majority of studies ask participants to make mental inferences, although as we discuss later in this section, there appears to be important cultural variation in whether emotions are perceived as situated actions vs. as mental states that cause actions.

The use of posed configurations of facial movements in assessments of emotion perception. The majority of the experiments that study emotion perception ask participants to infer emotion in facial configurations that are posed by actors who are not in an emotional state when the photos were taken or by computer-generated humans who have no actual emotional state. As a consequence, it is not possible to assess the accuracy (i.e., validity) of perceivers' emotional inferences and, correspondingly, data from emotion perception studies cannot be interpreted as support for the validity of common beliefs about emotional expressions. As is the case in studies of expression production, it is more appropriate to interpret participants' responses in terms of their agreement (or consensus) with common beliefs. Even more serious is the fact that the proposed expressive facial configurations in Figure 4 do not capture the wider range of muscle movements that are observed when people express instances of these emotion categories. A recent study that mined over seven million images from the internet (for method, see Box 7 in SOM; Srinivasan & Martinez, 2018) identified multiple facial configurations associated with the same emotion category label and their synonyms --17 distinct facial configurations were associated with the word “happiness,” five with “anger,” four with “sadness,” four with “surprised,” two with “fear,” and one with “disgust.” The different facial configurations associated with each emotion word were more than mere variations on a universal core expression – they were distinctive sets of facial movements.[31]

Measuring emotion perception. The typical emotion perception experiment takes one of several forms, summarized in Table 5. **Choice-from-array** tasks, in which participants are asked to match photos of facial configurations and emotion words (with or without brief stories), have dominated the study of emotion perception since the 1970s. For example, a meta-analysis of emotion perception studies published in 2002 summarized 87 studies, 83 (95%) of which exclusively used a choice-from-array response method (Elfenbein & Ambady, 2002). This method has been widely criticized for over two decades, however, because they limit the possibility of observing evidence that could disconfirm the common view. Participants are strongly constrained in how they can infer meaning in a facial configuration, such as a photograph of a scowling facial configuration, since their choices are constrained to the options provided in the experiment (usually a small number of emotion words). In fact, the preponderance of choice-from-array tasks in the scientific study of emotion perception has been identified as one important factor that has helped perpetuate and sustain the common view (Russell, 1994). Other tasks exist for assessing emotion perception (see Table 5), including those that use a **free-labeling** method, where participants are asked to freely nominate words to label photographs of posed facial configurations, rather than choosing a word from a small set of predefined options. For example, upon viewing a scowling configuration, participants might offer words like “angry,” “sad,” “confused,” “hungry,” or even “wanting to avoid a social interaction.” By allowing participants more freedom in how they infer meaning in a facial configuration, free-labeling makes it equally possible to observe evidence that could either support or disconfirm the common view.

Recent innovations in measuring emotion perception use computer generated faces or heads rather than photographs of posed human faces. One method, called **reverse correlation,** measures participants’ internal model of emotional expressions (i.e., their mental representations of which facial configurations are likely to express instances of emotion) by observing how participants label an avatar head that displays random combinations of animated facial action units (Yu et al., 2012; for a review, see Jack et al., 2018; Jack & Schyns, 2017). As each pattern appears (on a given test trial), participants infer its emotional meaning by choosing an emotion label from a set of options (a choice-from-array response). After thousands of trials, researchers estimate the statistical relationship between the dynamic patterns of facial movements and each emotion word (e.g., disgust) to reveal participants’ beliefs about which facial configurations are likely to express different emotion categories.

A second approach using computer-generated faces would have participants interact with more fully developed **virtual humans** (Rickel, Marsella, et al., 2003), also known as Embodied Conversational Agents (Cassell et al., 2000). Virtual humans are software-based artifacts that look like and act like people (for examples, see Figure 7). They are similar to characters in video games in their surface appearance and are designed to interact face-to-face with humans using the same verbal and nonverbal behavior that people use to interact with one another. The underlying technologies used to realize virtual humans vary considerably in approach and capability, but most virtual human models can be programmed to make context-sensitive, dynamic facial actions that, when in a person, would typically communicate emotional information to other people (see Box 11 in SOM for discussion). The majority of the scientific studies with virtual humans were not designed to test whether human participants infer specific emotional meaning in a virtual human's facial movements, but their design makes them useful for studying when and how facial movements take on meaning as emotional expressions: Unlike all the other ways of assessing emotion perception discussed so far, which ask participants to make explicit inferences about the emotional cause of facial configurations, interactions with virtual humans allow scientists to study how a participant **implicitly** infers emotional meaning during social interactions.

Testing the common view from observations of whether certain facial configurations are reliably and specifically perceived as expressions of certain emotion categories. Traditionally, in most experiments, if participants reliably infer an emotional state from a facial configuration (e.g., inferring anger from a scowling facial configuration) at levels that are greater than what would be expected by chance, then this is taken as evidence that people **recognize** an emotional state in its facial display. It is more scientifically correct, however to interpret this as evidence that people *infer* an emotional state (i.e., they consistently make a reverse inference) unless the inference has been verified as valid (i.e., the person in the photograph is, indeed, in the expected emotional state). Only when reverse inferences are observed in a reliable and specific way within an experiment can scientists reasonably infer that participants are perceiving an instance of a certain emotion category in a certain facial configuration; technically, the inference holds only for emotion perception as it occurs in the particular situations contained in the experiment (because situations are never randomly sampled). If the emotion perception evidence *replicates* across experiments that sample people from the same culture, then the interpretation can be generalized to emotion perceptions in that culture. Only when the findings generalize across cultures – that is, replicate across experiments that sample people from different cultures -- is it reasonable to conclude that people universally infer a specific emotional state when perceiving as specific facial configuration. These findings might also be interpreted as evidence about the reliability and specificity of *producing* emotional expressions if the co-evolution assumption is valid (i.e., that emotional expressions and their perception co-evolved as an integrated signaling system; Ekman et al., 1972; Jack et al., 2016; Shariff & Tracy, 2011).

Studies of Healthy Adults From the U.S. and Other Developed Nations

Studies that measure emotion perception with choice-from-array tasks. The most recent meta-analysis of emotion perception studies was published in 2002 (Elfenbein & Ambady, 2002). It statistically summarized 87 experiments in which over 22,000 participants from over 20 cultures around the world inferred emotional meaning in facial configurations and other stimuli (such as posed vocalizations). The majority of participants were sampled from larger-scale or developed countries, including Argentina, Brazil. Canada, Chile, China, England, Estonia, Ethiopia, France, Germany, Greece, Indonesia, Ireland, Israel, Italy, Japan, Malaysia, Mexico, the Netherlands, Scotland, Singapore, Sweden, Switzerland, Turkey, the US, Zambia and various Caribbean countries. The majority of studies (95%) used posed facial configurations; only four studies had participants label spontaneous facial movements, a dramatic example of the challenges facing validity that we discussed earlier. All but four studies used a choice-from-array response method to measure emotion inferences, a good example of the challenges facing hypothesis disconfirmation that we discussed earlier.

The results of the meta-analysis, presented in Figure 8a, reveal that perceivers inferred emotions in the facial configurations in Figure 4 in line with the common view, well above chance levels (using the criteria set out by Haidt & Keltner (1999), presented in Table 2).. Results provided strong evidence that, when participants are viewing posed facial configurations made by people from their own culture, they reliably perceived the expected emotion in those configurations: scowling facial configurations were perceived as anger expressions, wide-eyed facial configurations were perceived as fear expressions, and so on, for all six emotion categories. Moderate levels of reliability were observed when perceivers were labeling facial configurations posed by people from other cultures; this difference in reliability between same- and cross-culture differences is referred to as an **ingroup advantage** (see Box 12, in SOM). The majority of emotion perception studies do not report whether the hypothesized facial configurations are perceived with any specificity (e.g., how likely was a scowl to be perceived as expressing an instance of emotion categories other than anger, or as an instance of a mental category that is not considered emotional). Without information about specificity, no firm conclusions can be drawn about the emotional meaning of the facial configurations in Figure 4, especially for the translational purpose of inferring someone's emotional state from their facial comportment in real life.

Nonetheless, most of the studies cited in the meta-analysis interpret their reliability findings alone as evidence for the reverse inference of inferring anger from a scowling face, disgust from a nose-wrinkled face, fear from a wide-eyed gasping face, and so on. Such findings may explain why many scientists who study emotion, when surveyed, indicated that they believe compelling evidence exists for the hypothesis that certain emotion categories are each expressed with a unique, universal facial configuration (see Ekman, 2016) and interpret variation in emotional expressions to be caused by cultural learning that modifies what are presumed to be inborn universal expressive patterns (e.g., Cordaro et al., 2017; Ekman, 1972; Elfenbein, 2013). Cultural learning has also been hypothesized to modify how people “decode” facial configurations during emotion perception (Buck, 1984).

Studies that measure emotion perception with free-labeling tasks. As we foreshadowed, experimental methods that place fewer constraints on participants’ inferences in experiments that measure emotion perception (Table 5) provide considerably less support for the common view of emotional expressions. In the least constrained experimental task, called free-labeling, perceivers freely volunteer a word (emotion or otherwise) that they believe best captures the meaning in a facial configuration rather than choosing from a small set of experimenter-chosen options. In urban samples, participants who freely-label facial configurations produce the expected emotion labels with weak reliability (when labeling spontaneously produced facial configurations) to moderate reliability (when labeling posed facial configurations), and usually reveal weak specificity when it is assessed at all (for examples and discussion, see Russell, 1994; also see Naab & Russell, 2007). For example, when participants from many countries where English, Spanish, Mandarin Chinese, Farsi, Arabic and Russian is spoken as a first language were then asked to freely provide emotion words to label each of 35 facial configurations that had been cross-culturally identified (Srinivasan & Martinez, 2018), their labels provided evidence of a moderately reliable one-to-one correspondence between any facial configuration and emotion category, but there was no evidence of specificity (see Figure 8b).[32] Multiple facial configurations were associated with the same emotion category label (e.g., 17 different facial configurations were associated with the expression of happiness, five with anger, four with sadness, four with surprise, two with fear, and one with disgust). This many-to-many mapping is inconsistent with the common view that the facial configurations in Figure 4 are universally recognized as expressing the hypothesized emotion category, and they give evidence of variation that is far beyond what is proposed by the basic emotion view. Some of this variability may come from variability across different cultures and languages, but there is variability even within a single culture and language. Evidence of this many-to-many mapping is apparent in free-labeling tasks in small-scale, remote samples as well (Gendron et al., 2018), which we discuss in the next section.

Studies that measure emotion perception with the reverse correlation method. Using a choice-from-array response method with the reverse correlation method is an inductive way to learn people's beliefs about which facial configurations express an emotion category (for a review, see Jack et al., 2018; Jack & Schyns, 2017). In such studies, participants view thousands of random combinations of AUs that are computer generated on an avatar head and label each one by choosing an emotion word from a set of pre-defined options. All of the facial configurations labeled with the same emotion word (e.g., anger) are then statistically combined for each participant to estimate a belief about which facial movements express the corresponding emotion category. One recent study using the reverse correlation method with U.K. and Chinese participants found evidence of both variation in the facial movements that were judged to express a single emotion category, as well as similarity in the facial movements that were judged to express different categories (Jack et al., 2016). The study first identified groupings of emotion words that are widely discussed in the scientific literature (which, we should note, is dominated by English), corresponding to 30 English words grouped into eight emotion categories for the U.K. sample (happy/excited/love, pride, surprise, fear, contempt/disgust, anger, sad and shame/embarrassed) and 52 Chinese words grouped into twelve categories in the Chinese sample (joyful/excitement, pleasant surprise, great surprise/amazement, shock/alarm, fear, disgust, anger, sad, embarrassment, shame, pride, and despise). The reverse correlation method revealed 62 separate facial configurations: the same emotion category in a given culture was associated with multiple models of facial movements because synonyms of the same emotion category were associated with distinctive models of facial movements. Amidst this variability, Jack and colleagues also found that these 62 separate facial configurations could be summarized as four prototypes which are presented in Table 6, along with the corresponding emotion words that they were frequently associated with. Each prototype was described with a unique set of affective features (combinations of valence, arousal and dominance). When the four estimated configurations are compared with the common view presented in Figure 4, along with the basic emotion hypotheses listed in Table 1, there are some striking similarities: Configuration 1 most closely resembles the proposed expression for happiness, configuration 2 is similar to a combination of the proposed expressions for fear and anger, configuration 3 most closely resembles the proposed expression for surprise, and configuration 4 is similar to a combination of the proposed expressions for disgust and anger. [33] Taken together, these findings suggests that, at the most general level of description, participants' beliefs about emotional expressions (i.e., their internal models of which facial movements expressed which emotions) were consistent with the common view (indeed, they could be taken to constitute part of the common view), but when examined in finer detail with more granularity, participants' also believe that there is substantial within-category variation in the facial movements that express instances of the same emotion category. This finding suggests that the way the common view is often described in reviews, depicted in the media, and used in many applications, does not in fact do justice to people's more detailed beliefs about variability in facial expressions.

Studies that implicitly assess emotion perception during interactions with virtual humans. Designers typically study how a virtual human's expressive movements influence an interaction with a human participant. Much of the early research modeling expressive movements in virtual humans focused on endowing them with the facial expressions proposed in Figure 4. A number of studies have endowed virtual humans with blends of these configurations (Bui et al., 2004; Arya et al. 2009). Designers are also inspired by other people's beliefs about how emotions are expressed. Actors, for example, have been asked to pose facial configurations that they believe express emotions, which are then processed by graphical and machine learning algorithms to craft the relation between emotional states and expressive movements (Alexander et al, 2009). In another study, human subjects used a specially designed software tool to craft animations of facial movements that they believed express certain mental categories, including emotion categories. Then, other human subjects judged the crafted facial configurations (Ochs et al., 2010). Increasingly, data-driven methods are used that place people in emotion-eliciting conditions, capture the facial and body motion and then synthesize animations from those captured motions (Niewadomski et al., 2015; Ding et al., 2014, Wang, Marsella, & Hawkins, 2008).

In general, studies with virtual humans nicely show how the situational context influences how people infer the meaning of facial movements (de Melo et al., 2014). For example, in a game that allowed competition and cooperation (Prisoner's Dilemma, Pruitt & Kimmel, 1977), a virtual human who smiled after making a competitive move evoked more competitive, less cooperative responses from human participants compared to a virtual human using an identical strategy in the game (tit-for-tat) but that smiled after cooperating. Virtual humans who make a verbal comment about a film that is inconsistent with their facial movements, such as saying they enjoyed the film but grimacing that was quickly followed by a smile, were perceived as less reliable, trustworthy and credible (Rehm & Andre, 2005).

The dynamics of the facial actions, including the relative timing, speed and duration of the individual facial actions, as well as the sequence of facial muscle movements over time, offer information over and above the mere presence or absence of the movements themselves and have an important influence on how human perceivers interpret facial movements (e.g., Ambadar et al., 2009; Keltner 1995; Jack & Schyns, 2017, Krumhuber et al. 2013) and how much they trust a virtual human during a social interaction (Krumhuber et al., 2009). Research with virtual humans has shown that the dynamics of facial muscle movements are critical for them to be perceived as emotional expressions (Niewiadomski et al, 2015; Ochs et al., 2010). These findings are consistent with research showing that the temporal dynamics carry information about the emotional meaning of facial movements that are made by real humans (e.g., Kamachi et al., 2001; Sato & Yoshikawa, 2004; Krumhuber & Kappas, 2005; for a review, see Krumhuber, Kappas & Manstead, 2013).[34]

Summary. Whether or not people can reliably perceive emotions in the expressive configurations of Figure 4, as predicted by the common view, depends on how participants are asked to report or register their inferences (see Table 4). Hundreds of experiments have asked participants to infer the emotional meaning of posed, exaggerated facial configurations like those presented in Figure 4 by choosing a single emotion word from a small number of options offered by scientists, called choice-from-array-tasks. This experimental approach tends to generate moderate to strong evidence that people reliably label scowling facial configurations as angry, frowning facial configurations as sad, and so on for all six emotion categories that anchor the common view. Choice-from-array tasks severely limit the possibility of observing evidence that can disconfirm the common view of emotional expressions, however, because they restrict participants' options for inferring the psychological meaning of facial configurations by offering them a limited set of emotion labels. (As we discuss below, when people are provided with labels other than angry, sad, afraid, as so on, they routinely choose them; e.g., Carroll & Russell, 1996; also see Crivelli et al., 2017). Additionally, the specificity of those judgments is largely unreported. Scientists often go further and interpret the reliability findings from these studies as evidence that scowls are expressions of anger, frowns are expressions of sadness, and so on. This logic is not sound, however, because most of these studies ask participants to infer emotion in posed, static faces which are likely limited in their validity (i.e., people posing facial configurations like those depicted in Figure 4 are unlikely to be in the hypothesized emotional state). Furthermore, other ways of assessing emotion perception, such as the reverse correlation method and free-labeling tasks, find much weaker evidence for reliability and/or specificity of emotion inferences. Instead, they suggest that what people actually infer and believe about facial movements incorporates considerable variability: In short, the common view depicted in many reviews, summaries, the media, and used in numerous applications is not an accurate reflection of what people in fact believe about facial expressions of emotion, when probed in more detail. In the next section, we discuss scientific evidence from studies of emotion perception in small-scale remote cultures, which further undermines the common view.

Studies of Healthy Adults Living in Small-Scale, Remote Cultures

A growing number of studies examine emotion perception in people from remote, non-industrialized groups. A more in-depth review of these studies can be found in Gendron et al. (2018). Our goal here is to summarize the trends found in this line of research (see Table 7).

Studies that measure emotion perception with choice-from-array tasks. During the period from 1969 to 1975, somewhere between five and eight small-scale samples from remote cultures in the South Pacific were studied with choice-from-array tasks to investigate whether participants perceived emotional expression in facial movements in a similar way when compared to people from the US and other industrialized countries of the Western world (see Figure 9a). Our uncertainty in the number of samples stems from reporting inconsistencies in the published record (see note to Table 7). We present the findings here according to how the original authors reported their findings, despite the inconsistencies. Five samples performed choice-from-array tasks, three in which participants chose a photographed facial configuration to match one brief vignette that described each emotion category (Ekman, 1972; Ekman & Friesen, 1971; Sorenson, 1975) and two in which they chose a photograph to match an emotion word (Ekman, Sorenson, & Friesen, 1969). All five samples performing some version of a choice-from-array task provided strong evidence in support of cross-cultural reliability of emotion perception in small-scale societies. Evidence for specificity was not reported**.** Until 2008, *all claims* that anger, sadness, fear, disgust, happiness and surprise are universally recognized (and therefore are universally expressed) were largely based on three papers (two of them peer reviewed) reporting on four samples (Ekman, 1972; Ekman & Friesen, 1971; Ekman et al., 1969).[35]

Since 2008, 10 verifiably separate experiments observing emotional inferences in small-scale societies have been published or submitted for publication. These studies include a greater diversity of social and ecological contexts, including sampling five small-scale societies across Africa and the South Pacific (see Figure 9b) who were tested with a greater diversity of research methods listed in Table 5, including tasks that allow for the possibility of observing cross-cultural variation in emotion perception and therefore the possibility of disconfirming the common view. Six samples registered their emotion inferences using a choice-from-array task, in which participants were given an emotion word and asked to choose the posed facial configuration that best matched it or vice versa (Crivelli, Jarillo et al., 2016; Crivelli, Russell et al., 2016; Crivelli et al., 2017, Study 2; Gendron et al., 2018, Study 2; Tracy & Robins, 2008). Only one study (Tracy & Robins, 2008) reported that participants selected an emotion word to match the facial configurations similar to those in Figure 4 more reliably than what would be expected by chance, and effects ranged from weak (anger and fear) to strong (happiness) with surprise and disgust falling in the moderate range.[36] Information about the specificity of emotion inferences was not reported. A close examination of the evidence from four studies by Crivelli and colleagues suggest weak to moderate levels of reliability for inferring happiness in smiling facial configurations (all four studies), sadness in frowning facial configurations (all four studies), fear in gasping, wide-eyed facial configurations (three studies), anger in scowling facial configurations (two studies) and disgust in nose-wrinkled facial configurations (three studies). A detailed breakdown of findings can be found in Box 13, in SOM. None of the studies found specificity for any facial configuration, however, except that smiling was reported as unique to happiness, but that finding did not replicate across samples.[37]

The final study using a choice-from-array task with people from a small-scale, remote culture is important because it involves the Hadza hunter-gatherers of Tanzania (Gendron et al., 2018, Study 2). [38] The Hadza are a high-value sample for two reasons. First, universal and innate emotional expressions are hypothesized to have evolved to solve the recurring fitness challenges of hunting and gathering in small groups on the African savanna (Pinker, 1997; Shariff & Tracy, 2011; Tooby & Cosmides, 2008); the Hadza offer a rare opportunity to study foragers who are currently living in an ecosystem that is thought to be similar to that of our Paleolithic ancestors.[39] Second, the population is rapidly disappearing (http://www.sciencemag.org/news/2018/05/farmers-tourists-and-cattle-threaten-wipe-out-some-world-s-last-hunter-gatherers). Prior to this study, the Hadza had not participated in any studies of emotion perception, although they have been the subject of social cognition research more broadly (H. C. Barrett et al., 2016; Bryant et al., 2016). After listening to a brief story about a typical instance of anger, disgust, fear, happiness, sadness and surprise, Hadza participants chose the expected facial configuration more often than chance if the target and foil were distinguished by the affective property referred to as valence (i.e., a smiling configuration depicting a pleasant state vs. a scowling configuration depicting an unpleasant state, consistent with anthropological studies of emotion (Russell, 1991), linguistic studies (Osgood, May & Miron, 1975) and findings from other recent studies of participants from small-scale societies, such as the Himba (Gendron et al., 2014a, b) and the Trobriand Islanders (Crivelli, Jarillo et al., 2014). (Also see Srinivasan & Martinez, 2018, described in Box 7, who showed that perceivers can reliably infer valence but not arousal in facial configurations). In addition, Hadza participants who had some contact with people from other cultures -- they had some formal schooling or could speak Swahili which is not their native language – were more consistently able to choose the common facial configuration than were those with no formal schooling who spoke minimal Swahili (for a similar finding with Fore participants in a free labeling study, see Table 2 in Sorenson, 1975). Of the 27 Hadza participants who had minimal contact with other cultures, only 12 reliably chose the wide-eyed gasping facial configuration to match the fear story at above chance levels. (Compare this finding to the observation that the hypothesized universal expression for fear – a wide-eyed gasping facial configuration – is understood as an aggressive, threatening display by Trobriand Islanders; Crivelli, Jarillo & Fridlund, 2016, 2017; Crivelli, Russell et al., 2016).

Studies that measure emotion perception with free-labeling tasks. During the period from 1969 to 1975, between one and three small-scale samples from remote cultures in the South Pacific were studied with free-labeling to investigate emotion perception (reported in Sorenson, 1975; see Table 7). From 2008 onward, two additional studies were conducted, one using spontaneous facial configurations (Crivelli et al., 2017, Study 1) and the other using posed facial configurations (Gendron et al., 2018, Study 2). Overall, all five studies provide no evidence that the facial configurations in Figure 4 evolved to specifically express certain emotion categories. The three free-labeling studies reported in Sorenson (1975) produced variable results. The only replicable finding appears to be that participants labeled smiling facial configurations uniquely as happiness in all studies (as the only pleasant emotion category tested). The two newer free-labeling studies both indicated that participants only rarely spontaneously labeled facial configurations with the expected emotion labels (or their synonyms) above chance levels. Trobriand Islanders did not label the proposed facial configurations for happiness, sadness, anger, surprise or disgust with the expected emotion labels (or their synonyms) at above chance levels (although they did label the faces consistently with other words; Crivelli et al., 2017, Study 1). Hadza participants labeled smiling and scowling facial configurations at above chance levels as happiness (44%) and anger (65%), respectively (Gendron et al., 2018, Study 2). The word "anger" was not used to uniquely label scowling facial configurations, however, and was frequently applied to frowning, nose-wrinkled and gasping facial configurations.

Facial movements carry meaningful information, even if they do not reliably and specifically display internal emotional states. The more recent studies of people living in small-scale, remote cultures suggest two interesting observations that are worthy of note. First, even though people may not routinely infer anger from scowls, sadness from frowns, and so on, they do reliably infer other social meanings for those facial configurations, because facial movements often carry important information about a person's inner state, such as their social motives (Crivelli et al., 2016, 2017; Rychlowska et al., 2015; Wood et al., 2016; Yik & Russell, 1999; for a discussion, see Fridlund, 2017; Martin et al., 2017). For example, as we mentioned earlier, Trobriand Islanders consistently labeled wide-eyed gasping faces (the proposed expressive facial configuration for the fear category) as signaling an intent to attack (i.e., a threat; for additional evidence in carvings and masks in a variety of cultures, including Maori, !Kung Bushmen, Himba, Eipo, see Crivelli, Jarillo, & Fridlund, 2016, 2017).

Second, people do not always infer internal psychological states (emotions or otherwise) from facial movements. People who live in non-western cultural contexts, including Himba and Hadza participants, are more likely to assume that other people's minds are not accessible to them, a phenomenon called **opacity of mind** in anthropology (Danziger, 2006; Robbins & Rumsey, 2008). Instead, facial movements are perceived as actions that predict future actions in certain situations (e.g., a wide-eyed gasping face is labeled as "looking" (Crivelli et al., 2017; Gendron et al., 2014a; Gendron et al. 2018). Similar observations were unavailable for the earlier studies conducted by Ekman, Friesen and Sorenson because, according to Sorenson (1975), they directed participants to provide emotion terms. When participants spontaneously offered an action label (e.g. "she is just looking") or a social evaluation (e.g., "he is ugly", or "he is stupid"), they were asked to provide an "affect term." Findings like these suggest that there may be profound cultural variation in the type of inferences human perceivers make when looking at other human faces in general, an observation that has been raised by a number of anthropologists and historians.

A note on interpreting the data. To properly interpret the scientific evidence, it's crucial to consider the constraints placed on participants by the experimental tasks they are asked to complete, summarized in Table 5. In most urban and in some remote samples, experiments using choice-from-array tasks produce evidence supporting the common view: Participants reliably label scowling facial configurations as angry, smiling facial configurations as happy, and so on. (We don't yet know whether perceivers are uniquely labeling each facial configuration as a specific emotion because most studies don't report that information.) It has been known for almost a century that choice-from-array tasks help participants obtain a level of reliability in their emotion perceptions that are not routinely seen in studies using methods that allow participants to respond more freely, and this is one reason they were chosen for use in the first place (for a discussion, see Gendron & Barrett, 2009, 2017; Russell, 1994; Widen & Russell, 2013). When participants are offered words for happiness, fear, surprise, anger, sadness, and disgust to register their inferences for a scowling facial configuration, they are prevented from judging a face as expressing other emotion categories (such as confusion or embarrassment), non-emotional mental states (e.g., a social motive, such as rejection or avoidance), or physical events (such as pain, illness or gas), thus inflating reliability rates within the task. When people are provided with other options, they routinely choose them. For example, participants label scowling faces as "determined" or "puzzled," wide-eyed faces as "hopeful" and gasping faces as "pained" when they are provided with stories about those emotions rather than with stories of anger, surprise and fear (Carroll & Russell, 1996; also see Crivelli et al., 2017). The problem is not with the choice-from-array task per se – it is more with failing to consider alternative explanations for the observations in an experiment and therefore drawing unwarranted conclusions from the data.

Choice-from-array tasks may do more than just limit response options, making it difficult to disconfirm commonsense beliefs. The emotion words provided during the task may actually encourage people to see anger in scowls, sadness in pouts, and so on, or to learn associations between a word (such as "anger") and a facial configuration (such as a scowl) during the experiment (e.g., Gendron et al., 2015; Hoemann, Crittenden, Ruark, Gendron, & Barrett, in press). The potency of words is discussed in Box 14, in SOM.

Summary. The pattern of findings from the studies conducted with remote samples replicates and underscores the pattern observed in samples of participants from larger, more urban cultural contexts: Asking perceivers to infer an emotion by matching a facial configuration to an emotion word selected from a small array of options, or telling participants a brief story about a typical instance of an emotion category and asking them to pick a facial configuration from an array of two or three photos, generally inflates agreement rates, producing evidence that is more likely to support the hypothesis of reliable emotion perception when compared to data coming from less constrained response methods such as free labeling (see Table 4). This is particularly true for studies that include only one pleasant emotion category, i.e., happiness, where all foils differ from the target in valence, and therefore the robust reliability and specificity for inferring happiness from smiling in these studies may be the result of participants engaging in valence perception rather than emotion perception, per se. Studies that use less constrained tasks that are designed to more freely discover how people perceive emotion instead yield evidence that generally fails to find support for the common view. Less constrained studies suggest that perceivers infer more than one emotion category from the same facial configuration, infer the same emotion category in a variety of different configurations and often disagree about the set of emotion categories that they infer. Cultural variation in emotion perception is consistent with the variation we observed in the first section of this paper when we reviewed studies of emotional expression production (again, see Table 4), and is even consistent with the basic of face perception, which itself is determined by experience and cultural factors (Caldara, 2016).

Studies of Healthy Infants and Children

Some scientists concur with the common view that infants can read specific instances of emotion in faces from birth (Haviland & Lelwica, 1987; Izard, Woodburn & Finlon, 2010; Leppänen & Nelson, 2009; Walker-Andrews, 2005). However, it is difficult to ascertain whether infants and young children possess the various capacities required to perceive emotion per se: simply detecting and discriminating facial movements is not the same as categorizing them to infer their emotional meaning. This is because it is challenging to design well-controlled experiments that do a good job of distinguishing these two capacities. Infants are preverbal, so scientists use other measurement techniques, such as the amount of time an infant looks at a stimulus, to infer whether infants can discriminate one facial configuration from another, and ultimately, whether infants categorize those configurations as emotionally meaningful (for a brief explanation, see Box 15, in SOM). This approach introduces several possible confounds because of the stimuli used in the experiments: infants and children are typically shown photographs of the proposed expressive forms that are similar to those presented in Figure 4 (e.g., Leppanen et al, 2009; Peltola et al., 2008). Infants are more familiar with some of these configuration than with others (e.g., most infants are more familiar with smiling faces than with scowls or frowns) and familiarity is known to influence perception (see Box 15, in SOM), making it difficult to know which features of a face are holding an infant's attention (familiarity or novelty) and which might be the basis of categorization in terms of emotional meaning. The configurations proposed for each emotion category also differ in their perceptual features (e.g., the proposed expressions for fear and surprise contain widened eyes whereas the proposed expression for sadness does not), contributing more ambiguity to the interpretation of findings. For example, when an infant discriminates smiling and scowling facial configurations, it is tempting to infer that that the child is discriminating expressions of anger and happiness when in fact that target of discrimination is the presence or absence of teeth in a photograph (Caron, Caron & Myers, 1985). Moreover, the facial configurations in question are usually made from exaggerated facial movements that are not typical of the expressive variation that children actually observe in their everyday lives (Grossman, 2010). Furthermore, unlike adults, infants may have had little or no experience with viewing *photographs* of anything, including heads of people with no bodies and no context.

The most important and pervasive confound in developmental studies of emotion perception is that most studies are not designed to test whether infants and children discriminate facial configurations according to their emotional meaning or whether they are discriminating affective features (pleasant vs. unpleasant; high arousal vs. low arousal) (see Box 9, SOM). Often, a facial configuration that is intended to depict a pleasant instance of emotion (smiling in happiness) is compared to one that is intended to depict an unpleasant instance of emotion (e.g., scowling in anger, frowning in sadness or gasping in fear), or these configurations are compared to a neutral face at rest (e.g., Leppänen, Richmond, Vogel-Farley, & Nelson, 2009; Leppänen, Moulson, Nelson & Vogel-Farley, 2007; Montague & Walker-Andrews, 2001). (This problem is similar to the one encountered earlier in our discussion of emotion perception studies in adults from small scale societies, in which perceptions of valence can be confused with perceptions of emotion categories). For example, in one study, 16-18 month olds preferred toys paired with smiling faces and avoided toys paired with scowling and gasping faces (Martin et al. 2014); this type of study cannot distinguish between whether infants are differentiating pleasant from unpleasant, approach vs. avoidance, or something about a specific emotion. Another study (Soken & Pick, 1999) reported that seven-month-olds distinguish sadness and anger when looking at faces, but only when the faces were paired with vocalizations. What is unclear is the extent to which the level of arousal or activation conveyed in the acoustic signals were most salient to infants. A recent study suggested that 10-month-old infants can differentiate between the high arousal, unpleasant scowling and nose-wrinkled facial configurations that are proposed as expressions of anger and disgust, suggesting that they can categorize these two facial configurations separately (Ruba et al., 2017). Yet, the scowling and nose-wrinkled facial configurations also differed in the properties besides their proposed emotional meaning: scowling faces showed no teeth, but nose-wrinkled faces were toothy, and it is well known that infants use perceptual features such as “toothiness” to categorize faces (see Caron et al., 1985). If an infant looks longer at a (pleasant) smiling facial configuration after viewing several (unpleasant) scowling faces, this does not necessarily mean that the infant has discriminated and understands “happiness” from “anger”; the infant might have discriminated positive from negative, affective from neutral, familiar from novel, the presence of teeth from the absence, less eye sclera from more, or even different amounts of contrast in the photographs. In the future, experiments must be designed to rule out the possibility that infants are categorizing facial configurations into different groupings based on factors other than emotion to provide a sound basis to infer that infants are processing specific emotional meaning.

As a consequence of these confounds, there is still much to learn about the developmental course of emotion perception abilities. By three months of age, infants can distinguish the facial features (the morphology) in the proposed expressive configurations for happiness, surprise, and anger, and, by seven months, they can discriminate the features in proposed expressive configurations for fear, sadness, and interest. Left uncertain is whether, beyond just discriminating between the mere appearance of particular facial features, infants also understand the emotional meaning that is typically inferred from those features. By seven months of age, infants can reliably infer whether someone is feeling pleasant or unpleasant when facial configurations are accompanied by sensory information from the voice (Flom & Bahrick, 2007; Walker-Andrews & Dickson, 1997). Only a handful of studies have attempted to test whether infants can infer emotional meaning in facial configurations rather than just discriminating between faces with different physical appearances, but they report conflicting results (Schwartz et al., 1985; Serrano et al., 1992). One promising future direction involves measuring the electrical signals (event related potentials, or ERPs) in infant brains as they view the proposed expressive configurations for anger and fear categories (e.g., Kobiella et al., 2008; Hoehl & Striano, 2008). Both of these studies reported differential brain responses to the proposed facial configurations for anger and fear, but their findings did not replicate one another (and for certain measurements, they observed opposing effects; for a broader review, see Grossmann, 2015).

Studies that measure a child's ability to use an adult caregiver's facial movements to resolve ambiguous or threatening situations, referred to as **social referencing**, have been interpreted as evidence of emotion perception in infants. One-year-olds use social referencing to stay in close physical proximity to a caregiver who is expressing negative affect, while infants are more likely to approach novel objects if the caregiver expresses positive affect (Carver & Vaccaro, 2007; Moses et al., 2001; Saarni et al., 2006). Similar results emerge from the caregiver's tone of voice (Hertenstein & Campos, 2004; Mumme, Fernald, Herrera, 1996). In fact, by 14 months of age, the positive or negative tone of a caregiver's voice influence what an infant will touch even more so than will a caregiver's facial movements or the content of what the adult is actually saying (Valliant-Molina & Bahrick, 2012; Vaish & Striano, 2004). These studies clearly suggest that infants can infer the valenced meaning of facial movements, at least when made by live (as opposed to virtual) people who they are familiar with. But, again, these data do not help resolve what, if anything, infants infer about the emotional meaning of facial movements.

Learning to perceive emotions. Children grow in emotionally rich social environments, making it difficult to run experiments that are capable of testing the common view of emotion perception while also taking into account the possible roles for learning and social experience. Nonetheless, several themes have emerged in the scientific literature, all of which suggest a clear role for learning and context in children's developing emotion perception capacities.

One hypothesis that continues to be strongly supported by experiments is that children's capacity to infer emotional meaning in facial movements depends on context (the conditions surrounding the face that may convey information about a face's meaning). For example, emotion concept learning, as a potent source of internal context, shapes emotion perception capacity (discussed in Boxes 10 and 16 in SOM). There are also developmental changes in how people use context to shape their emotional inferences about facial movements. Children as young as 19 months old can detect facial movements that are emotionally incongruent with a context (Walle & Campos, 2014). For example, when presented with adult facial configurations that are placed on bodies posing an emotional context (e.g., a scowling facial configuration placed on a body holding a soiled diaper), children (aged four, eight, and twelve) moved their eyes back and forth between faces and bodies when deciding how to label the emotional meaning of the faces, whereas adult participants directed their gaze (and overt visual attention) to the face alone, judging its emotional meaning in a way that was independent of the bodily context (Leitzke & Pollak, 2016). The youngest children were equally likely to label the scene based on face or context. The results of this experiment suggest that younger children devote greater attention to contextual information and actively cross-reference facial and contextual cues, presumably to better learn about and understand the emotional meaning those cues.[40]

Another important source of context that shapes the development of emotion perception in children involves the broader environment in which children grow. Children who grow up in neglectful or abusive environments, where their emotional interactions with caregivers are highly atypical, have a different developmental trajectory than do those growing in more consistently nurturing environments (Bick & Nelson, 2016; Pollak, 2015). Parents from these high-risk families produce unclear or context-inconsistent expressions of emotion (Shackman et al., 2010). Neglected children (who do not receive sufficient social feedback) show delays in perceiving emotions in the ways that adults do (Camras et al., 2006; Pollak et al., 2000), whereas children who are physically abused learn to preferentially attend to and identify facial movements that are associated with threat, such as a scowling facial configuration (Briggs-Gowan et al., 2015; Cicchetti & Curtis, 2005; da Silva Ferreira, Crippa, & de Lima Osório, 2014; Pollak, Vardi, Putzer Bechner, & Curtin, 2005; Shackman & Pollak, 2014; Shackman, Shackman, & Pollak, 2007). Abused children require less perceptual information to infer anger in a scowling configuration (Pollak & Sinha, 2002) and more reliably track the trajectory of facial muscle activations that signal threat (Pollak, Messner, Kistler & Cohn, 2009). Children raised in physically abusive environments also more readily infer anger and threat in ambiguous facial configurations (Pollak & Kistler, 2002) and then require more effortful control to disengage their attention from signs of threat (Pollak & Tolley-Schell, 2003) when compared to children who have not been maltreated. This close attention to scowling faces with knitted eyebrows shapes how abused children understand what facial movements mean. For example, one study found that five-year-old abused children tended to believe that almost any kind of interpersonal situation could result in an adult becoming angry; by contrast, most non-abused children understand that anger is likely in particular interpersonal circumstances (Perlman et al., 2008).

By three years of age, North American children not only start to show reliability in their emotion perceptions but they also begin to show evidence of specificity. They understand that facial movements do not necessarily map on to emotional states, and how someone really feels can be faked or masked. Moreover, they know what facial movements are expected in a particular context and try to produce them despite their feelings. For example, the "disappointing gift" experiments developed by psychologist Pamela Cole and her colleagues demonstrate this well. In one study, preschool-aged children were told they would be rewarded with a gift after they completed a task. Later, children received a beautifully wrapped package that contained a disappointing item, such as a broken pair of cheap sunglasses. When facing a smiling unfamiliar adult who has presented them with a gift, children forced themselves to smile (lip corner pull, cheek raise, and brow raise) and to thank the experimenter. Yet, while the children were smiling, they often kept their eyes focused, down, slumped their shoulders, and made negative statements about the object, indicating that they did not, in fact, feel positive about the situation (Cole, 1986). Moreover, there was no difference in the behavioral responses of visually impaired children when receiving a disappointing gift (Cole, Jenkins, & Shott, 1989). Studies like this one provide a more implicit way of assessing children's knowledge about emotion perception (i.e., it illustrates the inferences that children expect others to make from their own facial movements).

Summary. There is currently no clear evidence to support the hypothesis that infants and young children reliability and specifically infer emotion in the proposed expressive configurations for anger, disgust, fear, happiness, sadness and surprise categories (presented in Figure 4; findings summarized in Table 4). A more plausible interpretation of the existing evidence is that young infants infer affective meaning such as valence and arousal in facial configurations. Data from infants and young children obtained using a variety of methods further suggests that emotion perception abilities emerge and are shaped through learning in a social environment. These findings are consistent with evidence that the human face may be evolutionarily privileged to communicate importance or salience. But it is not clear that the expressive configurations proposed for specific emotion categories are similarly privileged in this way.

Summary of Scientific Evidence on the Perception of Emotion in Faces

The scientific findings on perception studies generally replicate those from production studies in failing to strongly support the common view. The one exception to this overall pattern of findings is seen in studies that ask participants match a posed face to an emotion word or scenario. This method produces evidence to support the common view, even when it is applied to completely novel emotion categories with made up expressive cues, opening up interesting questions about the psychological potency of the elements that make up choice-from-array designs (such as the emotion words embedded in the task or the choice of foils on a given trial). These findings reinforce our earlier conclusion that terms like “facial configuration” or “pattern of facial movements” or even “facial actions” are preferred to more loaded terms like “emotional facial expression,” “emotional expression” or “emotional display,” which can be, at best misleading, and at worst, incorrect.

Summary and Recommendations

Evaluation of the Empirical Evidence

The common view that humans around the world reliably produce and recognize certain emotions in specific configurations of facial movements continues to echo within the science of emotion, even as scientists increasingly acknowledge that anger, sadness, happiness and other emotion categories are more variable in their facial expressions. This entrenched common view does more than guide the practice of science. It influences public understanding of emotion, and hence education, clinical practice, and applications in industry. Indeed, it reaches into almost every facet of modern life, including emoticons and movies. Nonetheless, there is insufficient evidence to support it. People do express instances of anger, disgust, fear, happiness, sadness and surprise with the hypothesized facial configurations presented in Figure 4 at above chance levels, suggesting that those facial configurations sometimes serve as expressions of emotion as proposed. However, the reliability of this finding is weak, and there is evidence that the strength of support for the common view varies systematically with the research methods used. The strongest support for the common view -- found in data from urban, industrialized or developed samples completing choice-from-array tasks -- does not show robust generalizability. Evidence for specificity is lacking in almost all research domains. A summary of the scientific evidence is presented in Table 4.

The research findings do not imply that people move their faces randomly or that the configurations in Figure 4 have no psychological meaning. Instead, they reveal that the facial configurations in question are not “fingerprints” or diagnostic displays that reliability and specifically signal particular emotional states regardless of context, person and culture. It is not possible to confidently infer happiness from just a smile, anger from a scowl, or sadness from a frown, as numerous technology tries to do when applying what they mistakenly believe to be the scientific facts.

Instead, the available evidence from different populations and research domains – infants and children, adults living in industrialized countries and in remote cultures, and even individuals who are congenitally blind -- overwhelmingly points to a different conclusion: when facial movements do express emotional states, they are considerably more variable and dependent on context than the common beliefs allows. There appear to be many-to-many mappings between facial configurations and emotion categories (e.g., anger is expressed with a broader range of facial movements than just a scowl and scowls express more than anger). A scowling facial configuration may be *an* expression of anger in the sense of being a part of anger in a given instance. But a scowling facial configuration is not *the* expression of anger in any generalizable or universal way. Scowling facial configurations and the others in Figure 4 belong to a much large repertoire of facial movements that express more than one emotion category, and also non-emotional inner states, in a way that is tailored to specific situations and cultural contexts. The face is a powerful tool for social communication (Jack & Schyns, 2017). Facial movements, like reflexive and voluntary motor movements (Barrett & Finlay, in press), are strongly context-dependent. Recent evidence suggests the people's categories for emotions are flexible and responsive to the types and frequencies of facial movements they are exposed to in their environments (Plate, Wood, Woodard, & Pollak, in press).

The degree of variation suggested by the published evidence also goes well beyond the hypothesis that the facial configurations in Figure 4 are prototypes or *typical* expressions, and that any observed variation are merely the result of cultural accents, display rules, suppression or other regulatory strategies, differences in induction methods, measurement error, or stochastic noise (as proposed by various scientists, including Elfenbein, 2013, 2017; Ekman & Cordaro, 2011; Levenson, 2011; Matsumoto, 1990; Roseman, 2011; Tracy & Randles, 2011). Instead, the facial configurations in Figure 4 are best thought of as Western gestures, symbols or **stereotypes** that fail to capture the rich variety with which people spontaneously move their faces to express emotions in everyday life. A stereotype is not a prototype. The distinction is an important one, because a prototype is the most frequent or typical instance of a category (Murphy, 2002), whereas a stereotype is an oversimplified belief that is taken as generally more applicable than it actually is.

The conclusion that emotional expressions are more variable and context-dependent than commonly assumed is also mirrored by the evidence from physiological changes (such as heart rate and skin conductance measures, Box 8, SOM) and even in evidence on the brain basis of emotion (Clark-Polner et al., 2017). The task of science is to systematically document these context-dependent patterns, as well as understand the mechanisms that cause them, so that we can explain and predict them. Clearly, the face is a rich source of information that plays a crucial role in guiding social interaction. Facial movements, when measured in a high dimensional dynamic context, may serve the diagnostic purpose that many consumers of emotion science are looking for (where context can be a cultural context, a specific situation, a person's learning history or momentary physiological state, or even the temporal context of what just took place a moment ago; Barrett et al., 2011; Gendron et al., 2013).

A Note on the Scientific Literature

Our review identified several broad problems that lurk within the scientific research on facial expressions and that may cause considerable misunderstanding and confusion for consumers of this research. First, statistical standards are commonly adopted that don't translate well for applying emotion research to other domains, applied or scientific. Showing that people frown when sad or scowl when angry with greater statistical reliability than would be expected by chance may be a scientific finding that warrants publication in a peer-reviewed journal, but above-chance responding is often low in absolute terms, making broad conclusions impossible, particularly for translation to domains of life where a person's outcomes can be influenced by what emotional meaning perceivers infer. Making inferences based on statistical reliability without concern for specificity and generalizability is similarly problematic. Second, even studies that surmount these common shortcomings often have a mismatch between what is claimed in their conclusions (or in what others claim in reviews or citations of those primary research papers), and what inferences can, in fact, be supported by the results. Third, and relatedly, this mismatch often results from problems in how studies are designed—the particular stimuli used, the tasks used, and the statistical analyses are critically important and constrain what can be observed and inferred in the first place. Fourth, published research on emotional expressions and emotion perception often confounds the measurements made in an experiment with the interpretation of the data, referring without sufficient justification to facial movements as "emotional displays," "emotional expressions" or even "facial expressions," rather than "facial configurations," "facial movements" or "facial actions"; referring to people "detecting" or "recognizing" emotion rather than "perceiving" or "inferring" an emotional state based on some set of cues (facial movements, vocal acoustics, body posture, etc.); and referring to "accuracy" rather than "agreement" or "consensus."

A Note on Other Emotion Categories

Our conclusions most directly challenge what we have termed the "common view": that a scowling facial configuration is *the* expression of anger, a nose-wrinkled facial configuration *the* expression of disgust, a gasping facial configuration *the* expression of fear, a smiling facial configuration *the* expression of happiness, a frowning facial configuration *the* expression of sadness, and that a startled facial configuration is *the* expression of surprise. By necessity, we focused on our review of evidence on these six emotion categories, rather than the more than twenty emotion categories that are currently being studied, because studies on these six are far more numerous than for other emotion categories. Nonetheless, some scientists claim that these other emotion categories each have distinctive, universal expressions, facial or otherwise, that is modified or accented by culture (e.g., Cordaro et al., 2017; Keltner et al., in press). In our view, such claims rest on evidence that is subject to the same critique as we offered for the research that we reviewed in detail here. In short, even though our review focused on the six emotion categories that are sometimes referred to as "basic emotions," our observations and conclusions generalize to studies of other emotion categories that use similar methods.

Recommendations for Consumers of Emotion Research on Applying the Scientific Findings

Presently, many consumers of emotion research assume that certain questions about emotional expressions have been answered satisfactorily when in fact this is not the case. Technology companies, for example, are spending millions of research dollars to build devices to read emotions from faces, erroneously taking the common view as the one that is scientifically best supported. A more accurate description, however, is that their technology detects facial movements, not emotional expressions.[41] Corporations like Amazon are exploring virtual human technology to interface with consumers. Virtual humans are used to educate children, train physicians, train the military as well as infer psychological disorders and perhaps eventually even be used to offer treatments. At the moment, the science of emotion is ill-equipped to support these initiatives. Emotional expressions are more variable and context-dependent than originally assumed, and most of the published research was not designed to probe this variation and characterize this context-dependence. As a consequence, right now, the scientific evidence offers less actionable guidance to consumers than is commonly assumed.

In fact, our review of the scientific evidence indicates that very little about how and why certain facial movements express instances of emotion is actually known at a level of detail that such conclusions could be used in important, real-world applications. To help consumers navigate the science of emotion, we offer some tips for how to read experiments and other scientific papers (Table 8).

More generally, companies may well be fundamentally asking the wrong question. Attempts to simply “read out” people’s internal states from an analysis of their facial movements alone, without considering various aspects of context are at best incomplete, and at worst entirely lack validity, no matter how sophisticated the computational algorithms. These technology developments are powerful tools to investigate the expression and perception of emotions, as we discuss below. Right now, however, it is premature to use this technology to reach conclusions about what people feel based on their facial movements--which brings us to recommendations for future research.

Recommendations for Future Scientific Research

Specific, concrete recommendations for future research to capitalize on the opportunity offered by current challenges can be found in Table 9, but we highlight a few general points here. Foremost, the expressive stereotypes that summarize the common view, like those depicted in Figure 4, are ubiquitous in published research, but it’s time to move beyond a science of stereotypes to develop a science of how people actually move their faces to express emotion and the processes by which those movements carry information about emotion to someone else (a perceiver). (See Box 16 in SOM for a discussion of information theory as applied to emotional communication). The stereotypes of Figure 4 must be replaced by a thriving scientific effort to observe and describe the lexicon of context-sensitive ways in which people move their facial muscles to express emotion, and the discovery of when and how people infer emotions in other people's facial movements.

New research on emotion should consider sampling individuals deeply, with high dimensional measurements, across many different situations, times of day, etc.: a big data approach to learn the expressive repertoires of individual people. In the ideal case, videos of people in natural situations could be quantified by automated algorithms for various physical features such as facial movements, posture, gait, and tone of voice. To this we could add the sampling of other physical features such as ambulatory monitoring of autonomic nervous system changes to sample the internal milieu of people's bodies as they dynamically change over time, ambulatory eye-tracking to assess gaze and attention, ambulatory brain imaging such as EEG and optical brain imaging (fNIRs). The failure to find reliable "fingerprints" for emotion categories stems, at least in part, from the same reason there are no reliable facial movements to express these categories: approaches have ignored meaningful variability due to context. There is also blue tooth technology to capture the physical spaces people inhabit (which can be quantified for various structural and social descriptive features such as how much light and noise they are exposed to), whether they are with another person, how that person reacts, and so on. Rich, multimodal observations could, in principle, be available from videos, which when time-synched with the other physical measurements, could be extremely useful in understanding the conditions for when certain facial movements are made and what they might mean in a given context. Naturally, big data in the absence of hypotheses is not necessarily helpful.

People could be offered the opportunity to annotate their videos with subjective ratings of the features that describe their experiences (whether or not they are identified as emotions). Candidate features are affective properties such as valence and arousal (see Box 9 in SOM). The features might also be appraisals as descriptions of how a situation is experienced (Barrett, Mesquita et al., 2007) and have the potential to add to the high dimensional characterization of what causes facial movements and what they mean.[42] Such an approach introduces various technical and modeling challenges, but this sort of deeply inductive approach is now within reach.

Another opportunity for high dimensional sampling involves interactions with virtual humans. Because virtual humans can realize contingent behavior in rich social interactions under strict and precise experimental control, they can provide a richer, more natural context in which to study emotional expressions and emotion perception than is true for traditional laboratory studies, while not losing the experimental control that limits the causal inferences from ethological studies.

To date, this potential has not been exploited to explore the reliability and specificity in context-sensitive relations between facial movements and mental states. As we noted earlier, most of the systems are now designed to teach people a variety of skills, where the goal is not to assess how well participants perceive emotions in facial movements under realistic, socially ambiguous conditions, but instead to program expressive behaviors into virtual humans that will motivate people to learn the needed skills. In these experiments, the psychological realism of facial movements is often secondary to the primary goals of the experiment. A scientist might even program a virtual human with behavior or appearance that is un-natural or infeasible for a human (i.e., that are supernormal) so that a participant can unambiguously interpret and be influenced by the agent's actions (Tinbergen, 1953; D. Barrett, 2007).

Nonetheless, the scientific approach of observing people as they interact with artificial humans holds great promise for understanding the dynamics and mechanisms of emotion perception and may get us closer to understanding human emotion perception in everyday life. Virtual humans are vivid. Unlike more passive approaches to evoking emotion such as viewing videos or images of facial configurations, a virtual human engages a human participant in a direct, social interaction to elicit perceptual judgments that are either directly reported or inferred from behaviors measured in the participant. Virtual humans are also highly controllable, allowing for more precise experimentation (Blascovich et al., 2002). A virtual human's facial movements and other details can be repeated across participants offering the potential for robust and replicable observations. Numerous studies have demonstrated that humans are influenced by them (e.g., Baylor & Kim, 2008; Krumhuber et al, 2007; McCall et al., 2009). For example, human learners are more engaged by virtual agents who move their faces (and modulate their voices), leading them (the real humans) to increased sense of self-efficacy (Kim, Baylor, & Shen, 2007). As a consequence, virtual humans potentially "allow for the study of emotion in a rich virtual ecology, a form of synthetic in vivo experimentation" (Marsella & Gratch, 2016). When combined with the high dimensional sampling we described earlier, there is the potential to *revolutionize* our understanding of emotional expressions by asking *different questions* than those encouraged by common views. Automated algorithms using data captured from videos offer substantial improvements with a data-driven, unsupervised approach. The result could be the robust descriptions about the context-sensitive nature of emotional expressions that is currently missing, and that would set the stage for a more mechanistic, causal account of emotions and their expressions.

An ethology of emotions and their expressions can also be pursued in the lab. Experiments can go beyond a study of how people move their faces in a single situation chosen to be most typical of a given emotion category. Most studies to date have been designed to observe facial movements in only the most typical situations. Future studies should examine emotional expression and perception across a range of situations that vary systematically in their physical, psychological, and social features, and aim to understand both the various ways that humans acquire the skills to express and perceive emotion, as well as the conditions that can impair the development of these processes.

The shift towards more context-sensitive scientific studies of emotion has already begun (see Box 3 in SOM), but it currently falls short of what we are recommending. Non-scientists (and some scientists) still anchor on the common view and only slowly shift away from it (Tversky & Kahneman, 1974; Wilson et al., 1996). The pervasiveness of the common view supports strong convictions about what it is that faces signal, and people often continue to hold to those convictions even when they are demonstrably wrong (Barrett, 2017a; Todorov, 2017). Such convictions reflect cultural beliefs and stereotypes, however. This state of affairs is not unique to the science of emotional expression or to the science of emotion more generally (Kuhn, 1962).

In our view, the scientific path forward begins with the explicit acknowledgement that we know much less than we thought we did, providing an opportunity to cultivate the spirit of discovery with renewed vigor and take scientific discovery in a new direction (Firestein, 2016). With this context of discovery comes the sobering realization that those of us who cultivate the science of emotion and the consumers who use this research should seriously question the assumptions of the common view and step back from what we thought we knew about reading emotions in faces. Understanding how best to infer someone's emotional state or predict someone's future actions from their facial movements awaits the outcomes of future research.

End Notes

[1] English does not contain gender-neutral pronouns. As a consequence, we alternate between male and female pronouns.
[2] Decades of research in social psychology shows that humans automatically try to predict other people's behavior by inferring a mental state – this is called mental state inference or mentalizing, such as when inferring someone's emotional state (e.g., for a review, see Gilbert, 1998). This research suggests that inference and prediction are not separate steps (Smith & DeCoster, 2000).
[3] Bolded words appear in the glossary.
[4] To be clear, teaching children how to infer emotions in others is not a problem because this skill is related to efficient communication with others. The question is whether children are being taught information that is scientifically valid and generalizable.
[5] As of November 10, 2018, a website for the Detego Group indicated that "The methods developed (*sic*) by Paul Ekman are based on 40 years of research and are being taught to the FBI, CIA, Scotland Yard and more forensics specialists around the world" (http://www.detegogroup.eu/paul-ekman-introduction/?lang=en).
[6] This empirical emphasis is largely consistent with scientists' explicit reports of what they believe, according to a recent survey from 2014. Two-hundred and forty eight scientists who published peer-reviewed papers on the topic of emotion were asked about their views on what the scientific evidence shows. Of the 149 (60%) who responded, 119 (80%) indicated that they believed compelling evidence exists for the hypothesis that certain emotion categories are expressed with universal facial configurations or vocal signals (Ekman, 2016); no questions about variability were included in the survey.
[7] In social psychology, this is the distinction between identifying an action and making an inference about the mental cause of the action (Gilbert, 1998; Vallacher & Wegner, 1987).
[8] This corresponds to the null hypothesis for the true positive (in Figure 3).
[9] To test the specificity hypothesis, we test something called the false positive: that people frequently scowl when not angry, meaning that they scowl more frequently than chance when fearful, sad, confused, hungry, etc. (see Figure 3). Retaining the null hypothesis for the false positive, that people do not scowl more frequently than they would by chance when fearful, sad, confused, hungry, etc., is equivalent to rejecting the null hypothesis (i.e., finding support for) the specificity hypothesis. Rejecting the null hypothesis for the false positive, because people scowl when fearful, sad, confused, hungry, etc., in addition to when angry, is evidence of no specificity (i.e., retaining the null hypothesis for the test of specificity).
[10] Our decision to focus on the anger, disgust, fear, happiness, sadness and surprise categories was reinforced by two observations. First, consider a recent poll that asked scientists about their beliefs (Ekman, 2016). Two-hundred and forty eight scientists who published peer-reviewed papers on the topic of emotion were given a list of 18 emotion labels and were asked to indicate which, according to available empirical evidence, have been established as biological categories with universal expressions. Of the 149 (60%) who responded,

> "*There was high agreement about five emotions … : anger (91%), fear (90%), disgust (86%), sadness (80%), and happiness (76%)*. Shame, surprise, and embarrassment were endorsed by 40%–50%. Other emotions, currently under study by various investigators drew substantially less support: guilt (37%), contempt (34%), love (32%), awe (31%), pain (28%), envy (28%), compassion (20%), pride (9%), and gratitude (6%)." (Ekman, 2016, p. 32, italics added).

Second, there is no smoking gun in the published research on these additional emotion categories – that is, there are no scientific findings related to the production or perception of facial expressions for those emotion categories that thus far challenge the general conclusions of this paper. Simply put: regardless of how few or how many emotion categories we evaluated, the findings are the same.
[11] Different number of facial muscles are reported in various sources depending on how muscles are grouped or divided.
[12] From http://erikarosenberg.com/facs/ :"scientists often refer to a set of actions that occur on the face simultaneously as "facial events," rather than calling them facial expressions. It is more descriptive. The word "expression" suggests that something from the inside becomes observable on the outside. Yet not every facial behavior expresses an internal state – most probably do not."
[13] see https://how-emotions-are-made.com/notes/facial_action_coding

[14] Box 6 in SOM presents a summary of computer vision algorithms for automatically detecting facial actions.
[15] Changes in illumination and face orientation are currently major hurdles.
[16] Thirty-eight groups, each with their own face reading algorithm, announced their intention to participate in the challenge (Benitez-Quiroz et al., 2017a). Groups tuned their algorithms on the set of training images that were provided two weeks before the challenge deadline. Final evaluations were done on the testing set only. Of the original 38 groups, only four submitted results before the challenge ended.
[17] These accuracy levels might be considered an upper estimate because of the characteristics of the training and test image databases. The methods for choosing the database are described in Benitez-Quiroz et. al 2016), although we provide a few important details here: Note, however, that a number of images are posed and professional taken. Some facial configurations are exaggerated. Under these idealized circumstances, manual verification of these faces was estimated at 81% accuracy.
[18] It is also possible that an individual person has a variety of probabilistic physical changes that reliably and specifically occur during the instances of a single emotion category, but for a number of reasons this hypothesis has not yet been scientifically tested. Specific studies to address this question would be very helpful.
[19] There are ways to get around this circularity by using unsupervised, data-driven methods to discover categories, but to date, studies have used supervised approaches where categories are prescribed by human inference.
[20] By relying on their own beliefs, scientists are using human consensus to identify when an emotional episode is occurring and which emotion category it belongs to (i.e., when they agree that fear or some other emotion is present, then it is said to be present). It's important to realize that every single experiment dealing with emotion to date relies on human inference in this way. Consensus inferences are made in many areas of science. In physics and astronomy consensus emerges from expert scientists whose beliefs and assumptions often challenge the common sense view, such as in the case of quantum mechanics, dark matter, and black holes. In other areas of psychology, consensus is used to define many categories, such as memory and attention, as well as psychiatric categories, such schizophrenia and autism. Even defining depression as a mental vs. a physical illness is a matter of consensus rather than objective ground truth. But it is noteworthy that when it comes to emotions, scientists use exactly the same categories as non-scientists, which may give us cause for concern (as forewarned by William James; James, 1890, 1894). For example, compare the findings in Box 8 with the recent survey of scientists who study emotion (Ekman, 2016): 88 out of 149 scientists responded continue to believe that certain emotion categories have universal physiological markers, despite meta-analyses showing otherwise.
[21] These meta-analytic findings are consistent with an earlier summary published by Matsumoto et al. (2008): of the 14 studies using rigorous FACS coding by human experts, only five reported that participants spontaneously displayed some or all of the hypothesized AUs during emotions. This is in contrast to the nine studies using the less reliable EM-FACS coding, all of which reported support. These findings suggest that some type of perceptual bias creeps in when observers make configural judgments of whether an AU is present or not (e.g., indicating whether or not a participant is smiling, or displaying "happiness") than when AUs are coded independently, one at a time.
[22] Remote, small-scale cultures are not untouched by western influences. All cultures have some minimal contact with western cultures (and this was also the case for the seminal papers published by Ekman and his colleagues in the 1970s; Gendron & Crivelli, 2017).
[23] The Trobriand Islanders are a different ethnic group than the Fore; Trobrianders are subsistence fisherman and horticulturalists living in a small archipelago of islands located 200km from the mainland (the origin of the original Fore who were photographed). As Crivelli et al. make clear in their paper, these findings are a within-nation rather than a within-culture comparison.
[24] The value of this particular study is that the researchers not only coded infants' facial movements but also measured a range of concurrent movements that could support inferences about the infants' feelings of pleasantness, unpleasantness and level of arousal, termed affect (see Box 9), including increased respiration, withdrawal/leaning away with the body, stilling/freezing, struggling, turning toward the mother, extreme withdrawal, hiding of their faces, squirming, self-stimulation, looking toward mother, pointing at the object, doing a "double-take," and banging on the table.
[25] Bennett et al. (2002) note that when they observed facial actions were thought to be associated with more than one emotion category (e.g., when an infant produced a facial configuration that was a combination of scowling (anger) and pouting (sadness), they interpreted the expression using the facial actions in only the upper region of the face, which indicates that infants' facial movements were even more variable than reported in the data tables. A footnote in the paper further indicates that infants produced facial movements that were interpreted to reflect "interest" across

all of the eliciting situations, but these facial actions were not included in any data analyses (Bennett et al., 2002, footnote 1). Any facial configuration that included AUs stipulated as interest and AUs for another emotion category was coded as an expression of the other emotion category.

[26] Also, it is not clear that children find sour foods disgusting (e.g., Stein, Ottenberg, & Roulet, 1958; Rozin et al., 1986).Young children appear to be *attracted* to many things that adults find disgusting, whereas by the age of five, children have more adult-like behavioral responses and reject them (Rozin et al., 1986). For a discussion of how disgust is learned, see Widen & Russell (2013).

[27] In another naturalistic study, videos of children aged four through seven were downloaded from the internet and FACS coded (Camras et al., 2018). The children were playing "the scary maze game": a child solves maze after maze of increasing difficulty, only to encounter a screaming, demonic girl from the movie The Exorcist (filmed in 1973). The game is generally thought to evoke an instance of fear (hence the name "scary"), but it may also evoke surprise as the scary stimulus makes a sudden unexpected appearance. Children only produced the wide-eyed, gasping configuration (the proposed facial expression of fear) and/or a startled configuration (the proposed facial expression of surprise) with weak reliability (38% and 10%, respectively).

[28] By analogy, people who have been blind since birth learn color concepts and the relation between these concepts, such as "red," "blue," and "green" are similar to those of sighted people (e.g., congenitally blind individuals understand the US concept for "blue" is more similar to "green" than to "red"; Shepard & Cooper, 1992). The structure of brain regions in visual cortex that represent visual concepts are also virtually indistinguishable in sighted and congenitally blind individuals (Koster-Hale et al., 2014; Wang et al., 2015).

[29] The onset and severity of blindness varies hugely across studies. Even a small amount of visual experience in infancy or early childhood will influence brain development and provide experiences for learning about emotions (see earlier section on emotion concept development in infants). Helen Keller, for example, could see and hear until she was 19 months old, providing some initial scaffolding for her later ability to communicate.

[30] For example, recently, Ekman (2017) wrote, "Another challenge to the findings of universality came from the anthropologist, Margaret Mead …. Establishing that posed expressions are universal, she said, does not necessarily mean that spontaneous expressions are universal. I replied (Ekman, 1977) that it seemed illogical to presume that people can readily interpret posed facial expressions if they had not seen those facial expressions and experienced them in actual social life" (Ekman, 2017, p. 46).

[31] While these findings are instructive, they likely provide a lower limit of the possible real world variation in the facial configurations that express the varied instances of a given emotion category. After all, the internet is a curated version of reality and some frequent facial configurations are likely missing because they are rarely uploaded to the internet. Similarly, some configurations commonly found on the internet might not be commonly observed in the real world.

[32] Compare these findings to those from a study that mined images from the internet using a similar but narrower approach, and who had two raters use a choice-from-array method to label the images (Mollahosseini et al., 2016).

[33] Configuration 3 also resembles people's beliefs about the configurations that express fear and awe (i.e., the "international core patterns" reported by Cordaro et al. 2017).

[34] More generally, participants are more likely to perceive the intended emotion in the hypothesized facial configurations of Figure 4 when they are displayed on dynamically moving, synthetic faces (Wehrle, Kaiser, Schmidt, & Scherer, 2000), in video footage of posed facial muscle movements (e.g., Ambadar, Schooler & Cohn, 2005; Cunningham & Wallraven, 2009), and even in point-light displays of motion created by facial muscle movements (Bassili, 1979). This "dynamic advantage" sometimes disappears when participants are viewing real human faces (e.g., Fiorentini & Viviani, 2011; Gold, Barker, et al., 2013; Miles & Johnston, 2007; Nelson & Russell, 2011).

[35] Ekman & Friesen (1971) was chosen as one of the forty studies that changed psychology (Hicks, 2012) and, along with Ekman et al. (1969) is routinely discussed in introductory psychology textbooks.

[36] Dioula participants from Burkina Faso in West Africa showed strong reliability for labeling smiling facial configurations as happiness, moderate reliability for labeling frowning facial configurations as sadness, startled facial configurations as surprise, and nose-wrinkled facial configurations as disgust, and weak reliability for labeling scowling facial configurations as anger and wide-eyed gasping facial configurations as fear.

[37] For example, a sample of Trobriand Islanders, who are subsistence horticulturalists and fishermen living in the Trobriand Islands of Papua New Guinea, labeled a scowling facial configuration as anger with above chance reliability (.29% of the time), but also labeled that facial configuration more frequently with "feels like avoiding a

social interaction" (.50% of the time) (Crivelli, Russell et al., 2017, Study 2). In fact, the wide-eyed, gasping facial configuration that is thought to be the expression for fear (Figure 4) is understood as an expression of aggression or threat in the Trobriand culture (Crivelli, Jarillo & Fridlund, 2016, 2017; Crivelli, Russell et al., 2016). Trobrianders uniquely labeled smiling facial configurations as happiness across two studies but this finding did not replicate in a third sample nor in a sample of Mwani participants who are subsistence fisherman living on Matemo Island in Mozambigue, Africa.

[38] The ancestors of the Hadza are thought to have been continuously practicing a hunting and gathering lifestyle for at least the past 50,000 years in their current region of East Africa. Furthermore, Hadza social structure, mobility, residential patterns, and language have thus far remained largely buffered from their interactions with other ethnic groups (Apicella & Crittenden, 2016; Crittenden & Marlowe, 2008) which have been sustained for at least the past 100 years (Jones, 2016)

[39] The wide-eyed gasping stereotype for fear is thought to have evolved for enhanced sensory sampling that supports efficient threat detection (Susskind et al., 2008). Similarly, the nose-wrinkle stereotype for disgust is thought to have evolved in order to limit exposure to noxious stimuli (Chapman & Anderson, 2012; Chapman, Kim, Susskind, & Anderson, 2009).

[40] Interestingly, adult perceivers may have overtly looked at the postures less, but other evidence with the same stimuli suggest that different body contexts influenced how adult participants visually scanned the exact same facial configurations; Aviezer et al., 2008). At the other end of the age spectrum, older adults are also more influenced by context when inferring emotional meaning in facial configurations as compared to young adults (Ngo & Isaacowitz, 2015).

[41] Some applications will not be affected by context because they are not aiming to use facial movements to infer an individual's underlying emotional state. These initiatives have very specific applications in mind. For example, detecting pain in patients (Apple), driver drowsiness (Google), creating virtual facial expression stickers or animojis from one's own facial poses (Facebook, iPhone X), or Alibaba's "smile to pay."

[42] The word "appraisal" has two meanings in the science of emotion. Here, appraisals simply to refer to the descriptive features of how a situation is experienced, such as novelty, goal relevance, etc., without any inference about how those experiential features are caused (e.g., Clore & Ortony, 2008; Ortony & Clore, 2013). The other meaning of appraisal refers to the mechanisms that cause the experiential features as components of emotion (e.g., the component process model of emotion, in which appraisals are considered evaluative "checks" that the human mind uses in a serial fashion; e.g., Scherer, Mortillaro & Mehu, 2017). There is very little evidence that appraisals are, in fact, causal in nature (for a discussion, Parkinson, 1997). In some studies, for example, participants are presented with a written scenario that is assumed to automatically trigger a specific sequence of appraisal checks (i.e., cognitive evaluations), which in turn is hypothesized to produce a specific pattern of facial muscle movements. Notice that the main causal mechanisms here – appraisal checks – are not measured directly but are inferred to have occurred. In other studies, participants are asked to explicitly report on the appraisals they experience, on the assumption that the corresponding "checks" are active. Emerging scientific evidence links appraisals, as descriptive features, to facial movements, although the evidence to date suggests that these relationships are not as consistent as specific as hypothesized (a summary of this research program can be found in in Scherer, Mortillaro & Mehu, 2017).

Glossary

Accuracy: Extent to which a participant's performance corresponds to the intended performance on an experimental task. Critically, this requires proper experimental task design, so that the intended correct performance is **perceiver-independent**, and not subject to the whims of the experimenter.

Affect: A general property of experience that has at least two features: pleasantness or unpleasantness (valence) and degree of arousal. Affect is part of every waking moment of life and is not specific to instances of emotion, although all emotional experiences have affect at their core.

Appraisal: Scientists use the word "appraisal" either to describe how a situation is experienced (e.g., a situation is experienced as novel) or to refer to a literal cognitive mechanism that causes those features of experience (e.g., an evaluation or judgment of whether or not a situation is novel).

Approach/avoidance: A fundamental dimension of motivated behavior. It is different from valence, which is a dimension of experience rather than of behavior.

Category/Categorization: The psychological grouping of a collection of objects, people or events that are perceived to be similar in some way. May be done consciously or unconsciously. May be explicit (as when applying a verbal label to instances of the grouping) or implicit (treating instances the same way or behaving towards them in the same way).

Choice-from-array tasks: Any judgment task that asks research participants to pick a correct answer from a small selection of options provided by the experimenter. For example, in the study of emotion perception, participants are often shown a posed facial configuration depicting an emotional expression (e.g., a scowl), along with a small selection of emotion words (e.g., "angry," "sad," "happy") and asked to pick the word that best describes the face.

Common view: In this paper, the most predominant view about how emotions are related to facial movements. While difficult to quantify, we characterize it through examples, e.g., an internet Google search (Box 1, SOM). The common view holds that (a) certain emotions categories reliably cause specific patterns of facial muscle movements, and (b) specific configurations of facial muscle movements are diagnostic of certain emotions categories. See Figure 4.

Conditional probability: The probability that an event "X' will occur given that another event "Y" has already occurred, or p(X/Y). If "X" is a frown and "Y" is sadness, then p(frown/sadness) is the conditional probability that a person will frown when sad. See also **forward inference, reverse inference.**

Configural (vs featural) perception of a face: The visual analysis of something, like a face, that is holistic, meaning that the face is visually analyzed as a gestalt (or whole unit) that incorporates features and their relations. Featural processing means that individual features are perceived independently, without reference to one another.

Confirmation bias: The tendency to search for, remember, or believe evidence that is consistent with one's existing beliefs or theories, in favor of evidence inconsistent with one's beliefs or theories.

Congenitally blind: People who are born without vision. In the literature, there is considerable heterogeneity, with some people being truly blind from the moment they are born, but others having severe visual impairments short of complete blindness or becoming blind in infancy. If the cause is peripheral (in the eyes rather than the brain), such individuals may still be able to think and imagine very similarly to sighted individuals.

Consistency: An outcome that does not vary greatly across time, context, or different individuals (see forward inference). Consistency is not accuracy (a group of people can consistently believe something that is wrong).

Discrimination: In psychophysics, to judge that two stimuli are different from one another; separate from identifying what they are (identification) or what they mean (recognition).

Ecological validity: Refers to the extent to which the findings of a research study are able to be **generalized** to real-life settings; the extent to which an experimental protocol captures valid aspects of the real world (related to **in-the-wild**).

Emotional episode: A window of time during which there is an **emotional instance**. Often, but not always, accompanied by an experience of emotion , and sometimes, but not always, involves an **emotional expression.**

Emotional expression: a facial configuration, bodily movement, or vocal expression that reliability and specifically communicates an emotional state. Many so-called emotional expressions are in fact errors of **reverse inference** on the part of perceivers (e.g., an actor crying when not sad).

Emotional granularity: experiencing or perceiving emotions according to many different **categories** (e.g., low granularity = angry, sad, and afraid are all synonyms of "unpleasant;" high granularity = "frustration," "irritation," and "rage" are all distinct from one another and from "anger".

Emotional instance (or instance of emotion): An event categorized as an emotion. For example, an instance of anger is the categorization of an **emotional episode** of anger. In cognitive science, an instance is called a "token" and the category is called a "type". So, an instance of anger is a token of the category anger. (see Emotional episode).

Face inferiority effect: A phenomenon observed in emotion perception studies of toddlers and young children. They have difficulty inferring the causes for emotions depicted in facial movements alone when compared to inferring the causes of emotions depicted with stories or words.

Facial affect coding system (FACS): A system to describe and quantify visible human facial movements.

Facial configuration: A pattern of visible contractions of multiple muscles in the face; the production analog to configural perception of faces. Configurations can be described objectively (e.g., with FACS coding). Not synonymous with "facial expression", which requires an inference about how the facial configurations were caused.

Facial expression: A **facial configuration** that someone infers is expressing an internal state. Facial expressions of emotion are configurations that perceivers **reverse infer** to have been caused by an internal emotion state; they are thus **perceiver-dependent**.

Facial movement: A **facial configuration** that is objectively described in a **perceiver-independent** way. This description is agnostic about whether the movement expresses an emotion and does not use **reverse inference**. **FACS** coding is an example.

Forced-choice task: An experimental task in which a participant must choose between options provided by the experimenter.

Forward inference: Inferring an effect from knowing its cause. An example would be the **conditional probability** of observing a frown given we know somebody is angry, p(frown|anger).

Free labeling: An experimental task that is not **forced-choice**, but in which the participant generates words of her/his choosing.

Generalization: The replication of research findings across different settings, samples, or methods. Generalizability can be weak, for instance if a finding replicates to a limited extent, or strong, if it replicates across very different methods and cultures.

Habituation task: In a habituation task, infants are repeatedly shown objects or images that belong to the same category. When subsequently shown a novel stimulus (one that is not experienced as similar to the others), infants look longer at it. Used to infer how infants categorize stimuli.

In the wild: In the real world (vs. in the lab). Related to **ecological validity**.

In-group advantage: In sociology and social psychology, an in-group is the social group of which a person psychologically feels they are a member; typically, people have more visual experience and familiarity with in-group members. In-group advantage refers to the often superior ability to perceive faces or voices from one's in-group, as compared to from an out-group.

Mental inference/mentalizing: Assigning a mental cause to actions; also sometimes referred to as "theory of mind". The reverse inference of attributing emotions from seeing facial movements is an example of mentalizing.

Meta-analysis: A method for statistically combining findings from many studies.

Multimodal: Combining information from more than one of the senses (e.g., vision and audition).

Null hypothesis: The hypothesis or default position that there is no relationship between dependent and independent variables. The probability of observing results that support the null hypothesis is chance level, i.e. what would obtain if observations are random, or permuted. Consequently, if the null hypothesis is true, the distribution of p-values is uniform (every possible outcome has an equal chance).

Perceiver-dependent: Interpretation of an observation that depends on human judgment. Perceiver dependency can produce conclusions that are **consistent** across people but not accurate or **valid**.

Perceiver-independent: An observation that does not depend on human judgment. Although some philosophers argue that all observations require some human judgment, there are degrees of dependency. Judging that a flower vase is rectangular or oval is relatively perceiver-independent, whereas judging whether it looks nice is perceiver-dependent.

Percent agreement: A measure of agreement between raters; high agreement produces high inter-subject **consistency**. Percent agreement is not the same as percent accuracy, since the former is more **perceiver-dependent** than the latter.

Perceptual matching task: An experimental task that requires research participants to judge two stimuli, such as two facial configurations, as similar or different. This only requires **discrimination**, not **categorization**, recognition, or naming.

Priors: Background beliefs. In the context of Bayes's Theorem, the belief that a hypothesis is true depends not just on the evidence presented but also on the strength of prior beliefs. If a person has a strong prior, this may result in a **confirmation bias**.

Prototype: The most frequent or most typical instance of a category. Distinct from **stereotype**: A group of people may have a **perceiver-dependent** stereotype that is an inaccurate representation of the prototype.

Recognition: Acknowledging something's existence (which is confirmed to exist by perceiver-independent means). Contrast with perception (which involves inference and interpretation).

Replication: The extent to which new experiments come to the same conclusions as a previous study. Strong replications **generalize** well: similar conclusions are obtained even when the new experiments use different subject samples, stimuli, or contexts.

Reverse correlation: A psychophysical, data-driven technique for deriving a representation of something (e.g., an image of a facial configuration) by averaging across a large number of judgments.

Reverse inference: Inferring a cause from having observed its purported effect. For instance, inferring that a scowl means someone is angry (the conditional probability, p(anger|frown)). In general, reverse inference is poorly constrained, since multiple causes are usually compatible with any observation.

Sensory modalities: The different senses: vision, hearing, etc.

Specificity: Research conclusions that include positive as well negative statements. For instance, concluding that a frown signals anger but that other facial movements do not signal anger, and that a frown does not signal emotions other than anger. High specificity helps make

reverse inference valid. Ideally, research conclusions feature both high specificity for some domains, and high generalizability for others (e.g., that a frown signals only anger, but does so across all people and cultures).

Statistical learning: Detecting statistical regularities from an environment; learning to recognize patterns.

Stereotype: a widely held but inaccurate belief about a person or category.

Universal: Something that is common or shared by all humans. The source of this commonality (innate or learned) is a separate issue. If an effect is universal, it generalizes across cultures.

Validity: Whether an observed variable actually measures what is claimed. E.g., whether a facial movement indicates an emotion (construct validity), or is specific for a particular emotion (discriminative validity).

Acknowledgements

We thank Jose-Miguel Fernandez-Dols and James Russell for providing us with a copy of their edited volume before its publication. This article benefited from specific discussions with Linda Camras, Pamela Cole, Maria Gendron, Alan Fridlund, and Tuan Le Mau. We are grateful to Jennifer Fugate for her assistance with constructing Figure 4, and in particular for her FACS coding the facial configurations presented in Figure 4. We also send thanks to Sheri Widen who provided assistance on Box Figure 10-1. Many thanks also to Linda Camras, Vanessa Castro, Carlos Crivelli and David Cordaro for providing us with additional details of their published experiments. And thanks to Jeff Cohn for his guidance on acceptable inter-rater reliability levels for FACS coding. We are also deeply grateful to those friends and colleagues who invested their time and efforts in commenting on an earlier draft of this paper (although the summaries of published works and conclusions drawn from them reflect only the views of the authors): Linda Camras, Maria Gendron, Rachael Jack, Judy Hall, Ajay Satpute, and Batja Mesquita. The paper was supported by grants to from the U.S. Army Research Institute for the Behavioral and Social Sciences (W911NF-16-1-019), the National Cancer Institute (U01 CA193632) and the National Institute of Mental Health (R01 MH113234 and R01 MH109464) to LFB; the National Science Foundation (CMMI 1638234) to LFB and SM; NIMH grant 2P50MH094258 to RA; the National Institutes of Health (R01-DC-014498, R01-EY-020834) and the Human Frontier Science Program (RGP0036/2016) to AMM; NEI grant R56 EY020834 to AMM and LFB; the National Institute of Mental Health (R01 MH61285), the National Institute of Child Health and Human Development (U54 HD090256) and a James McKeen Cattell Fund Fellowship to SDP; and the Air Force Office of Scientific Research (FA9550-14-1-0364) to SM. The views, opinions, and/or findings contained in this paper are those of the authors and shall not be

construed as an official U.S. Department of the Army position, policy, or decision, unless so designated by other documents.

References

Adolphs, R. (2002). Neural mechanisms for recognizing emotions. *Current Opinion in Neurobiology 12*, 169-178.

Adolphs, R. (2017). How should neuroscience study emotions? by distinguishing emotion states, concepts, and experiences. *Soc Cogn Affect Neurosci, 12*(1), 24-31. doi:10.1093/scan/nsw153

Adolphs, R., Tranel, D., Damasio, H., and Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature*, *372*, 669-672.

`Ali Arya, Steve DiPaola, and Avi Parush, "Perceptually Valid facial Expressions for Character-Based Applications," International Journal of Computer Games Technology, vol. 2009, Article ID 462315, 13 pages, 2009.

Ambadar, Z., Schooler, J. W., & Cohn, J. F. (2005). Deciphering the enigmatic face: The importance of facial dynamics in interpreting subtle facial expressions. *Psychological Science, 16*, 403-410.

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.

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

Apicella, C. L., & Crittenden, A. N. (2016). Hunter-Gatherer families and parenting *The Handbook of Evolutionary Psychology*: John Wiley & Sons, Inc.

Arcaro, M. J., Schade, P. F., Vincent, J. L., Ponce, C. R., & Livingstone, M.S. (2017). Seeing faces is necessary for face-domain formation. *Nature Neuroscience*, *20*, 1404-1412.

Atzil, S., Gao, W., Fradkin, I., & Barrett, L. F**.** (2018). Growing a social brain. *Nature Human Behavior*, 2, 624–636.

Atzil, S., Touroutoglou, A., Rudy, T., Salcedo, S., Feldman, R., Hooker, J., Dickerson, B. C., Catana, C. & Barrett, L. F. (2017). Dopamine in the medial amygdala network mediates human bonding. *Proceedings of the National Academy of Sciences, 114, 2361-2366.*

Aviezer, H., & Hassin, R. R. (2017). Inherently ambiguous: An argument for contextualized emotion perception. In J-M. Fernandez-Dols & J. A. Russell (Eds.), *The science of facial expression* (p. 333-349). NY: Oxford U Press.

Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., . . . Bentin, S. (2008). Angry, disgusted, or afraid? Studies on the malleability of emotion perception. *Psychological Science, 19*(7), 724-732.

Aviezer, H., Trope, Y., & Todorov, A. (2012). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science, 338*(6111), 1225-1229. doi:10.1126/science.1224313

Bach, D. R., & Dayan, P. (2017). Algorithms for survival: a comparative perspective on emotions. *Nature Reviews Neuroscience, 18*, 311. doi:10.1038/nrn.2017.35

Bandes, S. A. (2014). Remorse, demeanor, and the consequences of misinterpretation. *Journal of Law, Religion and State*, *3(2)*, 170-199. doi:10.1163/22124810-00302004

Baron-Cohen S., Golan O., Wheelwright S., Hill J. J. Mind Reading: the interactive guide to emotions. London, UK: Jessica Kingsley Limited.; 2004.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., and Plumb, I. (2001). The "reading the mind in the eyes" test revised version: a study with normal adults, and adults with asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry 42*, 241-251

Barratt, D., Rédei, A. C., Innes-Ker, A., & van de Weijer, J. (2016). Does the Kuleshov effect really exist? Revisiting a classic film experiment on facial expressions and emotional contexts. *Perception, 45,* 847–874.

Barrett, Deirdre (2007). "Supernormal Stimuli--Why Birds Are Cuckoo". *Waistland: The (R)Evolutionary Science Behind Our Weight and Fitness Crisis*. New York, New York: W.W. Norton & Co.

Barrett, H. C., Bolyanatz, A., Crittenden, A. N., Fessler, D. M., Fitzpatrick, S., Gurven, M., . . . Pisor, A. (2016). Small-scale societies exhibit fundamental variation in the role of intentions in moral judgment. *Proceedings of the National Academy of Sciences, 113*(17), 4688-4693.

Barrett, L. F. (2004). Feelings or words? Understanding the content in self-report ratings of emotional experience. *Journal of Personality and Social Psychology, 87,* 266-281. PMC1351136

Barrett, L. F. (2006a). Are Emotions Natural Kinds? *Perspectives on Psychological Science, 1,* 28-58.

Barrett, L.F. (2006b). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review, 10*, 20-46.

Barrett, L. F., Lindquist, K., Bliss-Moreau, E., Duncan, S., Gendron, M., Mize, J., & Brennan, L. (2007). Of mice and men: Natural kinds of emotion in the mammalian brain? *Perspectives on Psychological Science*, 2, 297-312

Barrett, L. F. (2012). Emotions are real. *Emotion, 12*, 413-429

Barrett, L. F. (2013). Psychological construction: A Darwinian approach to the science of emotion. *Emotion Review, 5*, 379-389.

Barrett, L. F. (2011). Was Darwin wrong about emotional expressions? *Current Directions in Psychological Science*, *20, 400*-406.

Barrett, L. F. (2017a). *How emotions are made: The secret life of the brain*. New York: Houghton-Mifflin-Harcourt.

Barrett, L. F. (2017b). The theory of constructed emotion: An active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience, 12,* 1-23. https://doi.org/10.1093/scan/nsw154

Barrett, L. F. (2017c). Functionalism cannot save the classical view of emotion. *Social Cognitive and Affective Neuroscience, 12,* 34-36. https://doi.org/10.1093/scan/nsw156

Barrett, L. F., & Bliss-Moreau, E. (2009). Affect as a psychological primitive. *Advances in Experimental Social Psychology, 41*, 167-218.

Barrett, L. F**.** & Finlay, B. L. (in press). Concepts, goals and the control of survival-related behaviors. *Current Opinion in the Behavioral Sciences*.

Barrett, L. F. & Gendron, M. (2016). The importance of context: Three corrections to Cordaro, Keltner, Tshering, Wangchuk, & Flynn (2016). *Emotion, 16*, 803-806.

Barrett, L. F., & Satpute, A. B. (2017). Historical pitfalls and new directions in the neuroscience of emotion. *Neuroscience Letters*.

Barrett, L. F., Mesquita, B., Ochsner, K. N., & Gross, J. J. (2007). The experience of emotion. *Annual Review of Psychology, 58*, 373-403. PMC1934613

Barrett, L.F., Mesquita, B., & Gendron, M. (2011). Context in emotion perception. *Current Directions in Psychological Science, 20*, 286-290.

Barrett, L. F., Wilson-Mendenhall, C. D., & Barsalou, L. W. (2015). The conceptual act theory: A road map. Chapter in L. F. Barrett and J. A. Russell (Eds.), *The psychological construction of emotion* (p. 83-110). New York: Guilford.

Barsalou, L. W. (1983). Ad hoc categories. *Memory and Cognition, 11*(3), 211-227.

Barsalou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *J Exp Psychol Learn Mem Cogn, 11*(4), 629-654.

Barsalou, L.W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 617–45.

Barsalou, L. W., Simmons, W. K., Barbey, A. K., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences, 7(*2), 84-91.

Bassili, J. N. (1979). Emotion recognition: The role of facial movement and the relative importance of upper and lower areas of the face. *Journal of Personality and Social Psychology, 37,* 2049-2058.

Baylor, A. L., & Kim, S. (2008). *The Effects of Agent Nonverbal Communication on Procedural and Attitudinal Learning Outcomes*. Paper presented at the International Conference on Intelligent Virtual Agents, Tokyo.

Bedny, M., Saxe, R. (2012). Insights into the origins of knowledge from the cognitive neuroscience of blindness. *Cognitive Neuropsychology 29*: 56-84.

Bell C. (1806). *Essays on the Anatomy of Expression in Painting*. London: Longman, Hurst,Rees & Orme.

Benitez-Quiroz, C. F., Srinivasan, R., & Martinez, A. M. (2016). *Emotionet: An accurate, real-time algorithm for the automatic annotation of a million facial expressions in the wild*. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 5562-5570).

Benitez-Quiroz, F., Srinivasan, R., & Martinez, A. M. (2019). Discriminant Functional Learning of Color Features for the Recognition of Facial Action Units and their Intensities. *IEEE transactions on pattern analysis and machine intelligence*.

Benitez-Quiroz, C. F., Srinivasan, R., Feng, Q., Wang, Y., & Martinez, A. M. (2017a). *EmotioNet Challenge: Recognition of facial expressions of emotion in the wild*. arXiv preprint arXiv:1703.01210.

Benitez-Quiroz, C. F., Wang, Y., & Martinez, A. M. (2017b). *Recognition of action units in the wild with deep nets and a new global-Local loss*. In Proceedings of the International Conference on Computer Vision.

Benitez-Quiroz, F., Srinivasan, R., & Martinez, A. M. (*in press*). *Discriminant Functional Learning of Color Features for the Recognition of Facial Action Units and their Intensities*. IEEE transactions on pattern analysis and machine intelligence.

Bennett, D. S., Bendersky, M., & Lewis, M. (2002). Facial expressivity at 4 months: A context by expression analysis. *Infancy, 3*(1), 97-113.

Berggren, S., Sue Fletcher-Watson, Nina Milenkovic, Peter B. Marschik, Sven Bölte & Ulf Jonsson. (2018). Emotion recognition training in autism spectrum disorder: A systematic review of challenges related to generalizability. *Developmental Neurorehabilitation, 21*,141-154.

Bick, J., & Nelson, C. A. (2016). Early adverse experiences and the developing brain. *Neuropsychopharmacology, 41*(1), 177-196.

Blascovich, J., Loomis, J., Beall, A., Swinth, K., Hoyt, C., & Bailenson, J. N. (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry, 13*, 103-124.

Bolzani Dinehart, L. H., Messinger, D. S., Acosta, S. I., Cassel, T., Ambadar, Z., & Cohn, J. (2005). Adult perceptions of positive and negative infant emotional expressions. *Infancy, 8*(3), 279-303.

Boucher, J. D., & Carlson, G. E. (1980). Recognition of facial Expression in three cultures. *Journal of cross-cultural psychology, 11*(3), 263-280.

Bridges, C. B. (1932). The suppressors of purple. *Molecular and General Genetics MGG, 60*(1), 207-218.

Briggs-Gowan, MJ, Pollak, SD, Grasso, D, Voss, J, Mian, ND, Zobel, E, McCarthy, KJ, Wakschlag, LS, & Pine, DS. (2015). Attention bias and anxiety in young children exposed to family violence. *Journal of Child Psychology and Psychiatry*, *56*:11, 1194–1201.

Brodny, G., Kolakowska, A., Landowska, A., Szwoch, M., Szwoch, W., & Wróbel, M. R. (2016). *Comparison of selected off-the-shelf solutions for emotion recognition based on facial expressions*. Paper presented at the 9th International Conference on Human System Interaction, Portsmouth, UK, July 6-8, 2016.

Bryant, G. A., & Barrett, H. C. (2008). Vocal emotion recognition across disparate cultures. *Journal of Cognition and Culture*, *8*, 135–148.

Bryant, G. A., Fessler, D. M., Fusaroli, R., Clint, E., Aarøe, L., Apicella, C. L., . . . Chavez, B. (2016). Detecting affiliation in colaughter across 24 societies. *Proceedings of the National Academy of Sciences, 113*(17), 4682-4687.

Buck, R. (1984). *The communication of emotion*. New York, NY: Guilford Press.

Bui, D., Heylen, D., Poel, M. & Nijholt, A. (2004) Combination of facial movements on a 3D talking head. *Proceedings Computer Graphics International*, IEEE, pp. 284–291.

Cacioppo, J. T., Berntson, G. G., Larsen, J. H., Poehlmann, K. M., & Ito, T. A. (2000). The psychophysiology of emotion. In R. Lewis & J. M. Haviland-Jones (Eds.), *The handbook of emotion* (2 ed., pp. 173-191). New York: Guilford Press.

Cain, J. (2000). *The way I feel*. Parenting Press.

Calbi, M., Heimann, K., Barratt, D., Siri, F., Umilta, M. A., & Gallese, V. (2017). How Context Influences Our Perception of Emotional Faces: A Behavioral Study on the Kuleshov Effect. *Front Psychol, 8*, 1684. doi:10.3389/fpsyg.2017.01684

Caldara, R. (2017). Culture reveals a flexible system for face processing. *Current Directions in Psychological Science, 26*, 249-255.

Calvo, M.G., Gutierrez-Garcia, A., Fernandez-Martin, A., and Nummenmaa, L. (2014). Recognition of facial expressions of emotion is related to their frequency in everyday life. *Journal of Nonverbal Behavior*, *38*, 549-567.

Campos, J. J., Campos, R. G., & Barrett, K. C. (1989). Emergent themes in the study of emotional development and emotion regulation. *Developmental Psychology, 25*(3), 394.

Campos, J. J., Mumme, D., Kermoian, R., & Campos, R. G. (1994). A functionalist perspective on the nature of emotion. *Japanese Journal of Research on Emotions, 2*(1), 1-20.

Camras, L. A. (1991). Conceptualizing early infant affect: View II and Reply. In K. Strongman (Ed.), *International review of studies on emotion* (pp. 16–28, 33–36). New York: Wiley

Camras, L. A. (1992). Expressive development and basic emotions. *Cognition & Emotion, 6*(3-4), 269-283.

Camras, L. A., & Allison, K. (1985). Children's understanding of emotional facial expressions and verbal labels. *Journal of Nonverbal Behavior, 9*(2), 84-94.

Camras, L. A., & Shutter, J. M. (2010). Emotional facial expressions in infancy. *Emotion Review, 2*(2), 120-129.

Camras, L. A., Castro, V. L., Halberstadt, A. G., & Shuster, M. M. (2017). Spontaneously produced facial expressions in infants and children. In J-M. Fernandez-Dols & J. A. Russell (Eds.), *The science of facial expression* (p. 279-296). NY: Oxford U Press.

Camras, Fatani, Fraumeni & Shuster, 2016. The development of facial expressions. In Barrett, L. F., Lewis, M., & Haviland-Jones, J. M. (Eds.). (2016). *Handbook of Emotions*, 4th Edition (255-271). New York: Guilford.

Camras, L. A., Oster, H., Ujiie, T., Campos, J. J., Bakeman, R., & Meng, Z. (2007). Do infants show distinct negative facial expressions for fear and anger? Emotional expression in 11-month-old European American, Chinese, and Japanese infants. *Infancy, 11*(2), 131-155.

Camras, L. A., Perlman, S. B., Fries, A. B. W., & Pollak, S. D. (2006). Post-institutionalized Chinese and Eastern European children: Heterogeneity in the development of emotion understanding. *International Journal of Behavioral Development, 30*(3), 193-199.

Caron, R. F., Caron, A. J., & Myers, R. A. (1985). Do infants see emotional expressions in static faces? *Child Development, 56,* 1552-1560.

Carrera-Levillain P, Fernandez-Dols J. M. (1994). Neutral faces in context: Their emotional meaning and their function. *Journal of Nonverbal Behavior, 18,* 281-299.

Carroll, J. M., & Russell, J. A. (1996). Do facial expressions signal specific emotions? Judging emotion from the face in context. *Journal of Personality and Social Psychology, 70*(2), 205-218.

Carver, L. J., & Vaccaro, B. G. (2007). 12-month-old infants allocate increased neural resources to stimuli associated with negative adult emotion. *Developmental Psychology, 43(1)*, 54.

Cassell, J., et al., Eds. (2000). Embodied Conversational Agents. Cambridge, MA, MIT Press.

Cassia, V. M., Turati, C. & Simion, F. (2004). Can a nonspecific bias toward top-heavy patterns explain newborns' face preference? *Psychological Science* **15**, 379-383.

Castro, V. L., Camras, L. A., Halberstadt, A. G., & Shuster, M. (2017). Children's Prototypic facial Expressions During Emotion-Eliciting Conversations With Their Mothers. *Emotion*. doi:10.1037/emo0000354

Cattaneo, L., Pavesi, G. (2014). "The facial motor system." *Neuroscience and Biobehavioral Reviews, 38*, 135-159.

Cecchini, M., Baroni, E., Di Vito, C., & Lai, C. (2011). Smiling in newborns during communicative wake and active sleep. *Infant Behavior and Development, 34*(3), 417-423.

Ceulemans, E., Kuppens, P., & Mechelen, I. V. (2012). Capturing the Structure of Distinct Types of Individual Differences in the Situation-specific Experience of Emotions: The Case of Anger. *European Journal of Personality, 26*(5), 484-495.

Chapman, H.A. & Anderson, A.K. (2013). Things rank and gross in nature: a review and synthesis of moral disgust. *Psychological Bulletin, 139*, 300-27

Chapman, H. A., Kim, D. A., Susskind, J. M., & Anderson, A. K. (2009). In bad taste: Evidence for the oral origins of moral disgust. *Science, 323*(5918), 1222-1226.

Chiesa, S., Galati, D., & Schmidt, S. (2015). Communicative interactions between visually impaired mothers and their sighted children: analysis of gaze, facial expressions, voice and physical contacts. *Child: Care, Health and Development*, 41, 6, 1040–1046.

Chu, W. S., De la Torre, F., & Cohn, J. F. (2017). Selective transfer machine for personalized facial expression analysis. *IEEE transactions on pattern analysis and machine intelligence, 39(3),* 529-545.

Cicchetti, D., & Curtis, W. J. (2005). An event-related potential study of the processing of affective facial expressions in young children who experienced maltreatment during the first year of life. *Development and Psychopathology, 17(3)*, 641-677.

Clark-Polner, E., Johnson, T., & Barrett, L. F. (2017). Multivoxel pattern analysis does not provide evidence to support the existence of basic emotions. *Cerebral Cortex, 27,* 1944-1948.

Clore, G. L. & Ortony, A. (1991). What more is there to emotion concepts than prototypes? *Journal of Personality and Social Psychology*, *60*, 48-50.

Clore GL, Ortony A. (2013). Psychological Construction in the OCC Model of Emotion. *Emotion Review,* 5(4):335-343. doi:10.1177/1754073913489751.

Clore, G. L., & Ortony, A. (2008). Appraisal theories: How cognition shapes affect into emotion. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of Emotions* (3rd ed., pp. 628-642). New York, NY: Guilford Press.

Cole, P. M. (1986). Children's spontaneous control of facial expression. *Child Development*, 57, 1309-1321.

Cole, P. M., Jenkins, P. A., & Shott, C. T. (1989). Spontaneous expressive control in blind and sighted children. *Child Development*, 60, 683-688.

Cordaro, D. T., Sun, R., Keltner, D., Kamble, S., Huddar, N., & McNeil, G. (2017). Universals and Cultural Variations in 22 Emotional Expressions Across Five Cultures. *Emotion*. doi:10.1037/emo0000302

Corneanu, C. A., Simon, M. O., Cohn, J. F., & Guerrero, S. E. (2016). Survey on rgb, 3d, thermal, and multimodal approaches for facial expression recognition: History, trends, and affect-related applications. *IEEE transactions on pattern analysis and machine intelligence*, 38(8), 1548-1568.

Cowen, A. S. & Keltner, D. (2017). Self-report captures 27 distinct categories of emotion bridged by continuous gradients. *Proceedings of the National Academy of Sciences*, www.pnas.org/cgi/doi/10.1073/pnas.1702247114

Craig, A. D. (2015). *How do you feel?: an interoceptive moment with your neurobiological self*. Princeton University Press.

Crittenden, A. N., & Marlowe, F. W. (2008). Allomaternal care among the Hadza of Tanzania. *Human Nature, 19*(3), 249-262.

Crivelli, C., Carrera, P., & Fernández-Dols, J.-M. (2015). Are smiles a sign of happiness? Spontaneous expressions of judo winners. *Evolution and Human Behavior, 36(1)*, 52-58.

Crivelli, C., Jarillo, S., & Fridlund, A. J. (2016). A multidisciplinary approach to research in small-scale societies: Studying emotions and facial expressions in the field. *Frontiers in psychology*, 7:1073. doi:10.3389/fpsyg.2016.01073

Crivelli, C., Jarillo, S., Russell, J. A., & Fernández-Dols, J. M. (2016). Reading emotions from faces in two indigenous societies. *Journal of Experimental Psychology: General*, *145(7*), 830-843. doi:10.1037/xge0000172

Crivelli, C., Russell, J. A., Jarillo, S., & Fernández-Dols, J. M. (2016). The fear gasping face as a threat display in a Melanesian society. *Proceedings of the National Academy of Sciences, U.S.A.*, *113(44)*, 12403-12407. doi:10.1073/PNAS.1611622113

Crivelli C, Jarillo S and Fridlund AJ (2016) A Multidisciplinary Approach to Research in Small-Scale Societies: Studying Emotions and Facial Expressions in the Field. *Front. Psychol.* 7:1073. doi: 10.3389/fpsyg.2016.01073

Cunningham, D. W., & Wallraven, C. (2009). Dynamic information for the recognition of conversational expressions. *Journal of Vision, 9,* 1-17.

da Silva Ferreira, G. C., Crippa, J. A., & de Lima Osório, F. (2014). Facial emotion processing and recognition among maltreated children: a systematic literature review. *Frontiers in psychology*, *5*, 1460.

Danziger, K.(2006). Universalism and indigenization in the history of modern psychology. In *Internationalizing the History of Psychology* (Brock, A.C.,ed.),pp.208–225, New York University Press

Darwin, C. (1859/1964). *On the Origin of Species*. (A facsimile of the first edition. ed.). Cambridge, MA.: Harvard University Press.

Darwin, C. (1872/1965). *The Expression of the Emotions in Man and Animals* (Chicago: University of Chicago Press).

Dashiell, J. F. (1927). A new method of measuring reactions to facial expression of emotion. *Psychological Bulletin*, *24,* 174-175.

Davidson, R. J. (1992). Anterior cerebral asymmetry and the nature of emotion. Brain Cogn, 20(1), 125-151.

de Gelder, B. (2016). *Emotions and the body*. New York: Oxford University Press.

de Melo, C. Peter Carnevale, Stephen Read and Jonathan Gratch. Reading people's minds from emotion expressions in interdependent decision making. *Journal of Personality and Social Psychology, 106(1)*, 2014, pp. 73-88

de Mendoza, A.H., Fernández-Dols, J. M., Parrott, W. G., & Carrera, P. (2010). Emotion terms, category structure, and the problem of translation: The case of shame and vergüenza. *Cognition & Emotion, 24(4)*, 661-680.

DeCarlo, L. T. (2012). On a signal detection approach to -alternative forced choice with bias, with maximum likelihood and Bayesian approaches to estimation. *Journal of Mathematical Psychology, 56(3)*, 196-207. doi:10.1016/j.jmp.2012.02.004

Denham, S. A. (1998). *Emotional development in young children*. Guilford Press.

Denison, S., Reed, C. & Xu, F. (2013) The emergence of probabilistic reasoning in very young infants: evidence from 4.5- and 6-month-old infants. *Developmental Psychology, 49(2),* 243-249.

DiGirolamo, M. A., & Russell, J. A. (2017). The emotion seen in a face can be a methodological artifact: The process of elimination hypothesis. *Emotion, 17(3)*, 538-546. doi:10.1037/emo0000247

Ding, Y., K. Prepin, J. Huang, C. Pelachaud, and T. Artie`res, (2014) "Laughter animation synthesis," in Proceedings of the 2014 International Conference on Autonomous Agents and Multi-agent Systems, ser. AAMAS '14, 2014, pp. 773–780.

Dondi, M., Gervasi, M. T., Valente, A., Vacca, T., Bogana, G., De Bellis, I., ... & Oster, H. (2012, July). Spontaneous facial expressions of distress in fetuses. In *14th European Conference on facial Expression* (pp. 16-18).

Dondi, M., Messinger, D., Colle, M., Tabasso, A., Simion, F., Barba, B. D., & Fogel, A. (2007). A new perspective on neonatal smiling: differences between the judgments of expert coders and naive observers. *Infancy, 12(3)*, 235-255.

Doyle, C. M., & Lindquist, K. A. (2018). When a word is worth a thousand pictures: Language shapes perceptual memory for emotion. *Journal of Experimental Psychology: General, 147(1*), 62-73. doi:10.1037/xge0000361

Duchenne CB (1990/1862). *The Mechanism of Human Facial Expression* (Renard, Paris); reprinted (1990) (Cambridge Univ Press, London).

Duran, J. I., & Fernandez-Dols J-M. (2018). *Do emotions result in their predicted facial expressions? A meta-analysis of studies on the link between expression and emotion*. PsyArXiv preprint, available at https://psyarxiv.com/65qp7

Duran, J. I., Reisenzein, R., & Fernandez-Dols, J-M. (2017). Coherence between emotions and facial expressions: A research synthesis. In J-M. Fernandez-Dols & J. A. Russell (Eds.), *The science of facial expression* (p. 107-129). NY: Oxford U Press.

Edelman, G. M., & Tononi, G. (2000). *A universe of consciousness: How matter becomes imagination*. New York: Basic books.

Eibl-Eibesfeldt, I.(1989) *Human Ethology*. Aldine deGruyter.

Eibl-Eibesfeldt, I. (1972). Similarities and differences between cultures in expressive movements. In R. A. Hinde, *Non-Verbal Communication*. Oxford, England: Cambridge U. Press.

Ekman, P. (1972). Universals And Cultural Differences In facial Expressions Of Emotions. In J. Cole (Ed.), *Nebraska Symposium on Motivation*, 1971 (pp. 207-283). Lincoln, Nebraska: University of Nebraska Press.

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

Ekman, P. (1994). Strong evidence for universals in facial expressions: A reply to Russell's mistaken critique. *Psychological Bulletin*, *115*, 268-287.

Ekman, P. (2016). What scientists who study emotion agree about. *Perspectives on Psychology Science*, *11(1)*, 31-34. doi:10.1177/1745691615596992

Ekman, P. (2017). Facial expressions. In J-M. Fernandez-Dols & J. A. Russell (Eds.), *The science of facial expression* (p. 39-56). NY: Oxford U Press.

Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review*, *3(4)*, 364-370.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, *17(2)*, 124.

Ekman P, Friesen WV (1976) *Pictures of Facial Affect* (Consulting Psychologists Press, Palo Alto, CA).

Ekman P, Friesen WV (1978) *Facial Action Coding System: A Technique for the Measurement of Facial Movement* (Consulting Psychologists Press, Palo Alto, CA).

Ekman, P., & Oster, H. (1979). Facial expressions of emotion. *Annual Review of Psychology*, 30(1), 527-554.

Ekman, P., Friesen, W., & Ellsworth, P. (1972). *Emotion in the human face: guidelines for research and an integration of findings*. Elmsford, N.Y.: Pergamon Press.

Ekman, P., Friesen, W.V., and Hager, J.C. (2002). *Facial Action Coding System: The Manual On CD ROM* (Salt Lake City: The Human Face)

Ekman, P., Levenson, R. W., &Friesen, W. V. (1983). Autonomic nervous system activity distinguishes among emotions, *Science*, *221(4616)*, 1208-1210.

Ekman, P., Sorenson, E. R., & Friesen, W. V. (1969). Pan-cultural elements in facial displays of emotion. *Science*, *164(3875)*, 86-88.

Ekman, P. (1980). *The Face of Man.* New York: Garland Publishing, Inc.

Elfenbein, H., Beaupre, M., Levesque, M., & Hess, U. (2007). Toward a dialect theory: cultural differences in the expression and recognition of posed facial expressions. Emotion, 7, 131-146.

Elfenbein, H. A. (2013). Nonverbal dialects and accents in facial expressions of emotion. *Emotion Review*, 5, 90-96.

Elfenbein, H. A. (2017). Emotional dialects in the language of emotion. In J-M. Fernandez-Dols & J. A. Russell (Eds.), *The science of facial expression* (p. 479-496). NY: Oxford U Press.

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: a meta-analysis. *Psychological Bulletin, 128(2)*, 203-235.

Emde, R. N., & Koenig, K. L. (1969). Neonatal smiling and rapid eye movement states. *Journal of the American Academy of Child Psychiatry, 8(1)*, 57-67.

Essa, I. A., & Pentland, A. P. (1997). Coding, analysis, interpretation, and recognition of facial expressions. *IEEE transactions on pattern analysis and machine intelligence, 19(7)*, 757-763.

Feldman, R. (2016). The neurobiology of mammalian parenting and the biosocial context of human caregiving. *Hormones and Behavior, 77*, 3-17.

Feng D., Jeong, D., Krämer, N., Miller, L. and Marsella, S. "Is It Just Me?: Evaluating Attribution of Negative Feedback as a Function of Virtual Instructor's Gender and Proxemics", in *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems*, May 2017, pp. 810-818, International Foundation for Autonomous Agents and Multiagent Systems

Fernandez-Dols, J-M. (2017). Natural facial expression: A view from psychological construction and pragmatics. In J-M. Fernandez-Dols & J. A. Russell (Eds.), *The science of facial expression* (p. 457-378). NY: Oxford U Press.

Fernandez-Dols, J-M., Sanchez, F., Carrera, P., & Ruiz-Belda, M-A. (1997). Are spontaneous expressions and emotions linked? An experimental test of coherence. *Journal of Nonverbal Behavior, 21*, 163-177.

Fernández-Dols, J.-M., & Crivelli, C. (2013). Emotion and expression: Naturalistic studies. *Emotion Review, 5(1)*, 24-29. doi:10.1177/1754073912457229

Fernández-Dols, J.-M., & Ruiz-Belda, M.-A. (1995). Are smiles a sign of happiness? Gold medal winners at the Olympic Games. *Journal of Personality and Social Psychology, 69(6)*, 1113-1119.

Fiorentini, C. & Viviani, P. (2011). Is there a dynamic advantage for facial expressions? *Journal of Vision, 11*, 1-15.

Firestein, S. (2012). *Ignorance: How it drives science*: Oxford University Press.

Flom, R., & Bahrick, L. E. (2007). The development of infant discrimination of affect in multimodal and unimodal stimulation: The role of intersensory redundancy. *Developmental psychology, 43(1)*, 238.

Frank, M. G., & Stennett, J. (2001). The forced-choice paradigm and the perception of facial expressions of emotion. *Journal of Personality and Social Psychology, 80*, 75–85.

Fridlund, A. (2017). The behavioral ecology view of facial displays, 25 years later. In J-M. Fernandez-Dols & J. A. Russell (Eds.), *The science of facial expression* (p. 77-92). NY: Oxford U Press.

Fridlund, A. J. (1991). Sociality of solitary smiling: Potentiation by an implicit audience. *Journal of Personality and Social Psychology, 60(2),* 229-240.

Fridlund, A. J. (1994). *Human facial expression: An evolutionary view*. San Diego, CA: Academic Press.

Friedman, B. H. (2010). Feelings and the body: the Jamesian perspective on autonomic specificity of emotion. *Biological Psychology, 84(3)*, 383-393.

Galati. D., Sini, B., Schmidt, S., & Tinti, C. (2003). Spontaneous facial expressions in congenitally blind and sighted children aged 8-11. *Journal of Visual Impairment & Blindness, 97*, 418-428.

Galati, D., Miceli, R., & Sini, B. (2001). Judging and coding facial expression of emotions in congenitally blind children. *International Journal of Behavioral Development, 25(3),* 268-278. doi:10.1080/01650250042000393

Galati, D., Scherer, K. R., & Ricci-Bitti, P. E. (1997). Voluntary facial expression of emotion: comparing congenitally blind with normally sighted encoders. *Journal of Personality and Social Psychology, 73(6),* 1363-1379.

Gandhi, T. K., Singh, A. K., Swami, P., Ganesh, S., & Sinha, P. (2017). Emergence of categorical face perception after extended early-onset blindness. *Proceedings of the National Academy of Sciences*, 114 (23) 6139-6143.

Gendron, M., & Barrett, L. F. (2009). Reconstructing the past: A century of ideas about emotion in psychology. *Emotion Review, 1*, 316-339. PMC2835158

Gendron, M., & Barrett, L. A. (2017). Facing the past: A history of the face in psychological research on emotion perception. In J-M. Fernandez-Dols & J. A. Russell (Eds.), *The science of facial expression* (p. 15-36). NY: Oxford U Press.

Gendron, M., Crivelli, C., & Barrett, L. F. (2018). Universality reconsidered: Diversity in meaning making of facial expressions. Current Directions in Psychological Science.

Gendron, M., Hoemann, K., Crittenden, A. N., Msafiri, S., Ruark, G., & Barrett, L. F. (2018). *Emotion perception in the Hadza hunter-gatherers*. PsyArXiv preprint, available https://psyarxiv.com/pf2q3/

Gendron, M., Lindquist, K., Barsalou, L., & Barrett, L. F. (2012). Emotion words shape emotion percepts. *Emotion, 12*, 314-325.

Gendron, M., Mesquita, B., & Barrett, L. F. (2013). Emotion perception: Putting a face in context. Chapter in D. Reisberg (Ed.), *Oxford Handbook of Cognitive Psychology* (p. 539-556). New York: Oxford University Press.

Gendron, M., Roberson, D., & Barrett, L. F (2015). Cultural variation in emotion perception is real: A response to Sauter et al. *Psychological Science, 26*, 357-359.

Gendron, M., Roberson, D., van der Vyver, J. M., & Barrett, L. F. (2014a). Perceptions of emotion from facial expressions are not culturally universal: evidence from a remote culture. *Emotion, 14(2)*, 251-262. doi:10.1037/a0036052

Gendron, M., Roberson, D., van der Vyver, J. M., & Barrett, L. F. (2014b). Cultural relativity in perceiving emotion from vocalizations. *Psychological Science*, *25*, 911-920.

Gewald, J.-B. (2010). Remote but in contact with history and the world. *Proceedings of the National Academy of Sciences, 107(18),* E75. doi:10.1073/pnas.1001284107

Gilbert, D. T. (1998). Ordinary personology. In S. T. Fiske & L. Gardner (Eds.), *The handbook of social psychology* (Vol. 1 and 2, pp. 89-150). New York, NY: McGraw-Hill.

Gold, J. M., Barker, J. D., Barr, S., Bittner, J. L., Bromfield, W. D., Chu, N., et al. (2013). The efficiency of dynamic and static facial expression recognition. *Journal of Vision*, 13(5): 23, 1–12

Goldstone R. (1994). An efficient method for obtaining similarity data. *Behavior Research Methods, Instruments, & Computers*,1994; 26:381–386.

Goldstone, R. L., Steyvers, M., & Rogosky, B. J. (2003). Conceptual interrelatedness and caricatures. *Memory & Cognition, 31,* 169–180.

Granchrow, J. R., Steiner, J. E., & Daher, M. (1983). Neonatal facial expressions in response to different qualities and intensities of gustatory stimulation. *Infant Behavior & Development, 6,* 153–157.

Gross, C. T., & Canteras, N. S. (2012). The many paths to fear. *Nature Reviews Neuroscience, 13(9)*, 651-658.

Gross, D., & Harris, P. L. (1988). False beliefs about emotion: Children's understanding of misleading emotional displays. *International Journal of Behavioral Development, 11(4),* 475-488.

Gross, J. J., & Barrett, L. F. (2011). Emotion generation and emotion regulation: One or two depends on your point of view. *Emotion Review, 3*, 8-16. PMC3072688

Grossmann, T. (2010). The development of emotion perception in face and voice during infancy. *Restor Neurol Neurosci, 28(2):*219-36

Grossmann, T. (2015). The development of social brain functions in infancy. *Psychological Bulletin, 141(6),* 1266-1287. doi:10.1037/bul0000002

Guillory, S. A. & Bujarski, K. A. (2014). Exploring emotion using invasive methods: Review of 60 years of human intracranial electrophysiology. *Social Cognitive and Affective Neuroscience,* doi:10.1093/scan/nsu002

Gunnery, S. D., & Hall, J. A. (2014). The Duchenne Smile and Persuasion. *Journal of Nonverbal Behavior, 38(2),* 181-194. doi:10.1007/s10919-014-0177-1

Gunnery, S. D., Hall, J. A., & Ruben, M. A. (2013). The deliberate Duchenne smile: Individual differences in expressive control. *Journal of Nonverbal Behavior*, *37(1)*, 29-41. doi:10.1007/s10919-012-0139-4

Gweon, H., & Schulz, L. E. (2011). 16-month-olds rationally infer causes of failed actions. *Science, 332(6037)*, 1524.

Gweon, H., Tenenbaum, J., & Schulz, L. E. (2010). Infants consider both the sample and the sampling process in inductive generalization. *Proceedings of the National Academy of Sciences, 107(20)*, 9066-9071,

Haidt, J., & Keltner, D. (1999). Culture and facial expression: Open-ended methods find more expressions and a gradient of recognition. *Cognition & Emotion*, *13(3)*, 225-266.

Hao, Wang, Peng, & Ji, 2018. Facial action unit recognition augmented by their dependencies. Presented at 2018 13th IEEE International Conference on Automatic Face and Gesture Recognition.

Harris, P. L. (1994). The child's understanding of emotion: Developmental change and the family environment. *Journal of Child Psychology and Psychiatry, 35(1)*, 3-28.

Hata, T., Hanaoka, U., Mashima, M., Ishimura, M., Marumo, G., & Kanenishi, K. (2013). Four-dimensional HDlive rendering image of fetal facial expression: a pictorial essay. *Journal of Medical Ultrasonics, 40(4)*, 437-441.

Haviland, J. M., & Lelwica, M. (1987). The induced affect response: 10-week-old infants' responses to three emotion expressions. *Developmental Psychology, 23(1)*, 97-104..

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behav Brain Sci, 33(2-3)*, 61-83; discussion 83-135. doi:10.1017/S0140525X0999152X

Hertenstein, M.J. & Campos,J.J. (2004). The retention effects of an adult's emotional displays on infant behavior. *Child Development*, 75: 595-613.

Hess, U., & Hareli, S. (2017). The social signal value of emotions: The role of contextual factors in social inferences drawn from emotion displays. In J-M. Fernandez-Dols & J. A. Russell (Eds.), The science of facial expression (p. 375-392). NY: Oxford U Press.

Hicks, R. R. (2012). *Forty studies that changed psychology: Explorations into the history of psychological research*, 6th Ed. Pearson, Prentice Hall.

Hjortsjö, C. H. (1969). *Man's face and mimic language*. Studen litteratur.

Hoehl, S., & Striano, T. (2008). Neural Processing of Eye Gaze and Threat-Related Emotional facial Expressions in Infancy. *Child Development, 79(6)*, 1752-1760.

Hoemann, K., Crittenden, A. N., Ruark, G., Gendron, M., & Barrett, L. F. (in press). Context facilitates the cross-cultural perception of emotion. *Emotion*.

Holodynski, M., & Friedlmeier, W. (2006). *Development of emotions and emotion regulation* (Vol. 8). Springer Science & Business Media.

Hortensius, R., Schutter, D. J., & Harmon-Jones, E. (2012). When anger leads to aggression: induction of relative left frontal cortical activity with transcranial direct current stimulation increases the anger–aggression relationship. *Soc Cogn Affect Neurosci, 7(3)*, 342-347.

Hout, M. C., Goldinger, S. D., & Ferguson, R. W. (2013). The versatility of SpAM: A fast, efficient, spatial method of data collection for multidimensional scaling. *Journal of Experimental Psychology: General, 142*(1), 256-281.

Hoyt, C., Blascovich, J., & Swinth, K. (2003). Social inhibition in immersive virtual environments. *Presence, 12*(2), 183-195.

Hutto, J.R, Vattoth, S (2015), "A practical review of the muscles of facial mimicry with special emphasis on the superficial musculoaponeurotic system." *American Journal of Radiology 204*: W19-W26.

Iwata, J., & LeDoux, J. E. (1988). Dissociation of associative and nonassociative concomitants of classical fear conditioning in the freely behaving rat. *Behavioral Neuroscience, 102(1)*, 66-76.

Izard, C. E. (2007). Basic emotions, natural kinds, emotion schemas, and a new paradigm. *Perspectives on Psychological Science, 2*, 260-280.

Izard, C. E. (1971). *The face of emotion*. Appleton-Century-Crofts.

Izard, C. E. (2010). The many meanings/aspects of emotion: Definitions, functions, activation, and regulation. *Emotion Review, 2(4)*, 363-370.

Izard, C. E., & Malatesta, C. Z. (1987). Perspectives on emotional development: Differential emotions theory of early emotional development. In J.D. Osofsky (Ed.), *Handbook of infant development* (2nd ed., pp. 494–554). New York: Wiley Interscience.

Izard, C. E., Fantauzzo, C. A., Castle, J. M., Haynes, O. M., Rayias, M. F., & Putnam, P. H. (1995). The ontogeny and significance of infants' facial expressions in the first 9 months of life. *Developmental Psychology, 31(6)*, 997.

Izard, C. E., Hembree, E., & Huebner, R. (1987). Infants' emotional expressions to acute pain: Developmental changes and stability of individual differences. *Developmental Psychology, 23,* 105–113.

Izard, C. E., Hembree, E., Dougherty, L., & Spirrizi, C. (1983). Changes in 2-to 19-month-old infants' facial expressions following acute pain. *Developmental Psychology, 19(3)*, 418-426.

Izard, C. E., Woodburn, E. M., & Finlon, K. J. (2010). Extending emotion science to the study of discrete emotions in infants. *Emotion Review, 2(2)*, 134-136.

Jack, R. E., & Schyns, P. G. (2017). Toward a social psychophysics of face communication. *Annual Review of Psychology, 68*, 269-297. doi:10.1146/annurev-psych-010416-044242

Jack, R. E., Crivelli, C., & Wheatley, T. (2018). Data-Driven Methods to Diversify Knowledge of Human Psychology. *Trends in Cognitive Science, 22(1)*, 1-5. doi:10.1016/j.tics.2017.10.002

Jack, R. E., Sun, W., Delis, I., Garrod, O. G., & Schyns, P. G. (2016). Four not six: Revealing culturally common facial expressions of emotion. *Journal of Experimental Psychology: General*, 145(6), 708-730. doi:10.1037/xge0000162

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

James, W. (1890/2007). *The Principles of Psychology* (Vol. 1). New York: Dover.

Jeni, Laszlo, Cohn, Jeffrey F, & De la Torre, Fernando. (2013). Facing imbalanced data recommendations for the use of performance metrics. *Proceedings of the Affective*

Computing and Intelligent Interaction, Geneva, Switzerland. Los Alamitos, CA: IEEE Computer Society.

Jones, N. B. (2016). *Demography and Evolutionary Ecology of Hadza Hunter-Gatherers* (Vol. 71): Cambridge University Press.

Kamachi, M., Bruce, V., Mukaida, S., Gyoba, J., Yoshikawa, S., & Akamatsu, S. (2001). Dynamic properties influence the perception of facial expressions. *Perception*, *30*, 875–887.

Kärtner, J., Holodynski, M., & Wörmann, V. (2013). Parental ethnotheories, social practice and the culture-specific development of social smiling in infants. *Mind, Culture, and Activity, 20(1)*, 79-95.

Kayyal, M. H., & Russell, J. A. (2013). Americans and Palestinians Judge Spontaneous facial Expressions of Emotion. *Emotion, 13(5)*, 891-904. doi:10.1037/a0033244

Keltner, D., & Cordaro, D. T. (2017). Understanding multimodal emotional expressions: Recent advances in Basic Emotion Theory. In J-M. Fernandez-Dols & J. A. Russell (Eds.), *The science of facial expression* (p. 57-75). NY: Oxford U Press.

Keltner, D. (1995). Signs of appeasement: Evidence for the distinct displays of embarrassment, amusement, and shame. Journal of Personality and Social Psychology, 68(3), 441-454.

Keltner, D., Sauter, D., Tracy, J., & Cowen, A. (in press). Emotional expression: Advances in Basic Emotion Theory. *Journal of Nonverbal Behavior*.

Kim, J., R. Hill, P. Durlach, H. Lane, E. Forbell, M. Core, S. Marsella, D. Pynadath, and J. and Hart (2009). "BiLAT: A Game-Based Environment for Practicing Negotiation in a Cultural Context. Special Issue on Ill-Defined Domains", International Journal of AI in Education, vol. 19, no. 3, pp. 289-308

Kim, Y., Baylor, A. L. & Shen, E. (2007). Pedagogical agents as learning companions: The impact of agent affect and gender. *Journal of Computer Assisted Learning* (JCAL), 23(3), 220-532.

Kobiella, A., Grossmann, T., Reid, V. M., & Striano, T. (2008). The discrimination of angry and fearful facial expressions in 7-month-old infants: An event-related potential study. *Cognition and Emotion, 22(1),* 134-146.

Koster-Hale, J., Bedny, M., & Saxe, R. (2014). Thinking about seeing: perceptual sources of knowledge are encoded in the theory of mind brain regions of sighted and blind adults. *Cognition, 133(1),* 65-78. doi:10.1016/j.cognition.2014.04.006

Kouo, J. L., & Egel, A. L. (2016). The Effectiveness of Interventions in Teaching Emotion Recognition to Children with Autism Spectrum Disorder. *Review Journal of Autism and Developmental Disorders, 3(3),* 254-265. doi:10.1007/s40489-016-0081-1

Kragel, P.A. & LaBar, K. S. (2013). Multivariate pattern classification reveals autonomic and experiential representations of discrete emotions. *Emotion, 13,* 681-90.

Kragel, P.A. & LaBar, K. S. (2015). Multivariate neural biomarkers of emotional states are categorically distinct. *Social Cognitive and Affective Neuroscience, 10,* 1437-1448.

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: a review. *Biological Psychology, 84(3),* 394-421. doi:10.1016/j.biopsycho.2010.03.010

Krumhuber, E. G. & Kappas, (2005). Moving smiles: The role of dynamic components for the perception of the genuiness of smiles. *Journal of Nonverbal Behavior, 29,* 3-24.

Krumhuber, E. G., Kappas, A., & Manstead, A. S. R. (2013). Effects of dynamic aspects of facial expressions: A review. *Emotion Review, 5,* 41-46.

Krumhuber, E., Manstead, A. S. R., Cosker, D., Marshall, D., & Rosin, P. L. (2009). Effects of dynamic attributes of smiles in human and synthetic faces: A simulated job interview setting. *Journal of Nonverbal Behavior, 33*, 1–15.

Krumhuber, E., Manstead, A., Cosker, D., Marshall, D., Rosin, P. L., & Kappas, A. (2007). Facial dynamics as indicators of trustworthiness and cooperative behavior. *Emotion, 7(4),* 730-735.

Kuhn, T. (1962). The structure of scientific revolutions. Chicago, IL: University of Chicago Press.

Kuppens, P., Tuerlinckx, F., Russell, J. A and Barrett, L. F. (2013). The relationship between valence and arousal in subjective experience. *Psychological Bulletin, 139,* 917-940.

Kuppens, P., Van Mechelen, I., & Rijmen, F. (2008). Toward disentangling sources of individual differences in appraisal and anger. *Journal of Personality, 76(4)*, 969-1000.

Kuppens, P., Van Mechelen, I., Smits, D. J. M., De Boeck, P., & Ceulemans, E. (2007). Individual differences in patterns of appraisal and anger experience. *Cognition & Emotion, 21(4),* 689-713. doi:Doi 10.1080/02699930600859219

Landau, B., & Shipley, E. (2001). Labelling patterns and object naming. *Developmental Science, 4(1),* 109-118.

Lang, P.J., Greenwald, M.K., Bradley, M.M., & Hamm, A.O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology, 30,* 261–273.

Lazarus, R. (1991). *Emotion and Adaptation*. NY, Oxford University Press.

Lebois, L. A. M., Wilson-Mendenhall, C. D., Simmons, W. K., Barrett, L. F., & Barsalou, L.W. (2018). Learning situated emotions. *Neuropsychologia*. https://doi.org/10.1016/j.neuropsychologia.2018.01.008

Lee, D. H. & Anderson, A. K. (2017). Form and function of facial expressive origins. In J-M. Fernandez-Dols & J. A. Russell (Eds.), *The science of facial expression* (p. 173-194). NY: Oxford U Press.

Lee, D. H., Mirza, R., Flanagan, J. G., & Anderson, A. K. (2014). Optical origins of opposing facial expression actions. *Psychological Science, 25(3)*, 745-752.

Leitzke, B. T., & Pollak, S. D. (2016). Developmental changes in the primacy of facial cues for emotion recognition. *Developmental Psychology*, *52*(4), 572.

Leppänen, J. M., & Nelson, C. A. (2009). Tuning the developing brain to social signals of emotions. *Nature Reviews Neuroscience, 10,* 37-47.

Leppänen, J. M., Moulson, M. C., Vogel-Farley, V. K., & Nelson, C. A. (2007). An ERP study of emotional face processing in the adult and infant brain. *Child Development, 78(1),* 232-245.

Leppänen, J. M., Richmond, J., Vogel-Farley, V. K., Moulson, M. C., & Nelson, C. A. (2009). Categorical representation of facial expressions in the infant brain. *Infancy, 14(3),* 346-362.

Levenson, R. W. (2011). Basic emotion questions. *Emotion Review, 3(4),* 379-386.

Lewis, M., & Sullivan, M. W. (Eds.). (2014). *Emotional development in atypical children*. Psychology Press.

Lewis, M., Ramsay, D. S., & Sullivan, M. W. (2006). The relation of ANS and HPA activation to infant anger and sadness response to goal blockage. *Developmental Psychobiology, 48(5),* 397-405.

Lewis, M., Sullivan, M. W., & Kim, H. M. S. (2015). Infant approach and withdrawal in response to a goal blockage: Its antecedent causes and its effect on toddler persistence. *Developmental Psychology*, *51*(11), 1553.

Lindquist, K., & Barrett, L. F. (2008). Emotional complexity. Chapter in M. Lewis, J. M. Haviland-Jones, and L.F. Barrett (Eds.), *The handbook of emotion*, 3rd Edition (p. 513-530). New York: Guilford.

Lindquist, K. A., Gendron, M., Barrett, L. F., & Dickerson, B.C. (2014). Emotion perception but not affect perception is impaired with semantic memory loss. *Emotion*, 14, 375-87.

Lindquist, K. A., Satpute, A. B., Wager, T. D., Weber, J., & Barrett, L. F. (2016). The brain basis of positive and negative affect: Evidence from a meta-analysis of the human neuroimaging literature. *Cerebral Cortex*. 26, 1910-22.

Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., & Barrett, L. F. (2012). The brain basis of emotion: a meta-analytic review. *Behavioral and Brain Sciences, 35(3)*, 121-143.

Marsella, S. & Gratch, J. (2016). Computational models of emotion as psychological tools. In L. F. Barrett, M. Lewis and J. Haviliand-Jones (Eds.) *Handbook of Emotions*, 4th Edition. New York: Guilford Press.

Marsella, S., Johnson, W.L., and LaBore, C. (2000). Interactive pedagogical drama. In Proceedings of the fourth international conference on Autonomous agents (AGENTS '00). ACM, New York, NY, USA, 301-308.

Martin, N. G., Maza, L., McGrath, S. J., & Phelps, A. E. (2014). An examination of referential and affect specificity with five emotions in infancy. *Infant Behavior and Development*, *37*(3), 286-297.

Martin, J., Rychlowska, M., Wood, A., and Niedenthal, P. (2017). Smiles as multipurpose social signals. *Trends in Cognitive Sciences, 21*, 864-877.

Martinez, A. M. (2017a). Computational models of face perception. *Current Directions in Psychological Science, 26(3)*, 263-269.

Martinez, A. M. (2017b). Visual perception of facial expressions of emotion. *Current Opinion in Psychology*, *17*, 27-33.

Martinez, A., & Du, S. (2012). A model of the perception of facial expressions of emotion by humans: Research overview and perspectives. *Journal of Machine Learning Research, 13*, 1589-1608.

Matias, R., & Cohn, J. F. (1993). Are max-specified infant facial expressions during face-to-face interaction consistent with differential emotions theory? *Developmental Psychology, 29(3)*, 524.

Matsumoto, D. (1990). Cultural similarities and differences in display rules. *Motivation and Emotion, 14*, 195-214.

Matsumoto, D. & Willingham, B. (2006). The thrill of victory and the agony of defeat: Spontaneous expressions of medal winners of the 2004 Athens Olympic Games. *Journal of Personality and Social Psychology, 91*, 568-581.

Matsumoto, D., & Willingham, B. (2009). Spontaneous facial expressions of emotion of congenitally and noncongenitally blind individuals. *Journal of Personality and Social Psychology*, *96(1)*, 1-10.

Matsumoto, D., Keltner, D., Shiota, M., Frank, M., & O'Sullivan, M. (2008). Facial expressions of emotions. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of Emotion,* 3rd Ed. (pp. 211-234). New York: Macmillan.

Mayr, E. (2004). *What makes biology unique? Considerations on the autonomy of a scientific discipline*. New York: Cambridge University Press.

McCall, C., Blascovich, J., Young, A., & Persky, S. (2009). Proxemic behaviors as predictors of aggression towards Black (but not White) males in an immersive virtual environment. *Social Influence, 4(2)*, 138 - 154.

Medin, D. L., Ojalehto, B., Marin, A., & Bang, M. (2017). Systems of (non-) diversity. *Nature Human Behaviour, 1*, 0088.

Mesquita, B. & Frijda, N. H. (1992). Cultural variations in emotions: A review. *Psychological Bulletin, 2*, 179-204.

Mesquita, B., Boiger, M., & De Leersnyder, J. (2016). The cultural construction of emotions. *Current Opinion in Psychology*, *8*, 31-36.

Messinger, D. S. (2002). Positive and negative: Infant facial expressions and emotions. *Current Directions in Psychological Science, 11(1)*, 1-6.

Michel, G. F., Camras, L. A., & Sullivan, J. (1992). Infant interest expressions as coordinative motor structures. *Infant Behavior and Development, 15(3)*, 347-358.

Miles, L. and Johnston, L. (2007). Detecting Happiness: Perceiver Sensitivity to Enjoyment and Non-Enjoyment Smiles. *Journal of Nonverbal Behavior, 31*, 259-275.

Miller, G. A. (1995). WordNet: a lexical database for English. *Communications of the ACM, 38(11)*, 39-41.

Mobbs, D., Weiskopf, N., Lau, H., Featherstone, E., Dolan, R., & Frith, C. (2006). The Kuleshov Effect: The influence of contextual framing on emotional attributions. *Social Cognitive and Affective Neuroscience, 1*, 95–106.

Mollahosseini, A., Hassani, B., Salvador, M. J., Abdollah, H., Chan, D., & Mahoor, M. N. (2016). *Facial expression recognition from world wild web*. arXiv preprint, available at https://arxiv.org/abs/1605.03639

Montagu, J. (1994). *The expression of the passions: The origin and influence of Charles Le Brun's lecture on general and particular expression*. London: Yale University Press.

Montague, D. P., & Walker-Andrews, A. S. (2001). Peekaboo: a new look at infants' perception of emotion expressions. *Developmental Psychology, 37(6)*, 826.

Moses, L. J., Baldwin, D. A., Rosicky, J. G., & Tidball, G. (2001). Evidence for referential understanding in the emotions domain at twelve and eighteen months. *Child Development, 72(3)*, 718-735.

Motley, M. T., & Camden, C. T. (1988). Facial expressions of emotion: A comparison of posed expressions versus spontaneous expressions in an interpersonal communication setting. *Western Journal of Speech Communication, 52*, 1-22.

Mumme, D. L., Fernald, A., & Herrera, C. (1996). Infants' responses to facial and vocal emotional signals in a social referencing paradigm. *Child Development*, *67*(6), 3219-3237.

Müri, R.M. (2015). Cortical control of facial expression. *The Journal of Comparative Neurology 524*, 1578-1585.

Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.

Naab, P. J., & Russell, J. A. (2007). Judgments of emotion from spontaneous facial expressions of New Guineans. *Emotion, 7(4)*, 736-744.

Namba, S., Makihara, S., Kabir, R. S., Miyatani, M., & Nakao, T. (2016). Spontaneous facial expressions are different from posed facial expressions: Morphological properties and dynamic sequences. *Current Psychology, 36*, 593-605.

Neidle, Carol, Judy A. Kegl, Dawn MacLaughlin, Benjamin Bahan & Robert G. Lee. 2000. *The syntax of American Sign Language*. Cambridge: MIT Press.

Nelson, K. E., Welsh, J. A., Trup, E. M. V., & Greenberg, M. T. (2011). Language delays of impoverished preschool children in relation to early academic and emotion recognition skills. *First Language*, *31*, 164–194.

Nelson, N. L. & Russell, J. A. (2011). Preschoolers' use of dynamic facial, bodily, and vocal cues to emotion. *Journal of Experimental Child Psychology, 110*, 52-61.

Nelson, N. L., & Russell, J. A. (2013). Universality revisited. *Emotion Review, 5*, 8-15.

Nelson, N. L., & Russell, J. A. (2016). Building emotion categories: Children use a process of elimination when they encounter novel expressions. *Journal of Experimental Child Psychology, 151*, 120-130.

Nelson, N. L., Hudspeth, K., & Russell, J. A. (2013). A story superiority effect for disgust, fear, embarrassment, and pride. *British Journal of Developmental Psychology, 31(3)*, 334-348.

Neth, D., & Martinez, A. M. (2009). Emotion perception in emotionless face images suggests a norm-based representation. *Journal of Vision, 9(1)*, 5.

Nezlek, J. B., Vansteelandt, K., Van Mechelen, I., & Kuppens, P. (2008). Appraisal-emotion relationships in daily life. *Emotion*, *8(1)*, 145-150.

Ngo, N. & Isaacowitz, D. M. (2015). Use of context in emotion perception: The role of top-down control, cue type, and perceiver's age. *Emotion, 15*, 292-302.

Niewiadomski, R, Ding, Y., Mancini, M., Pelachaud, C., Volpe, G., Camurri, A. (2015). Perception of intensity incongruence in synthesized multimodal expressions of laughter, Affective Computing and Intelligent Interaction (ACII), 2015 International Conference on, Xi'an, pages 684-690.

Norenzayan, A., & Heine, S. J. (2005). Psychological universals: What are they and how can we know? *Psychological Bulletin, 131(5),* 763-784.

Obrist, P. A. (1981). *Cardiovascular psychophysiology: A perspective.* New York: Plenum.

Obrist, P. A., Webb, R. A., Sutterer, J. R., & Howard, J. L. (1970). The cardiac-somatic relationship: some reformulations. *Psychophysiology, 6(5)*, 569-587.

Ochs, M., Niewiadomski, R. and Pelachaud, C. (2010). How a virtual agent should smile?: morphological and dynamic characteristics of virtual agent's smiles. In Proceedings of the 10th international conference on Intelligent virtual agents (IVA'10), Springer-Verlag, Berlin, Heidelberg, 427-440.

Ortony, A., Clore, G.L., Collins, A. (1990). *The cognitive structure of emotions*. Cambridge University Press .

Osgood, C. E., May, W H., & Miron, M. S. (1975). *Cross-cultural universals of affective meaning.* Urbana: University of Illinois Press.

Oster, H., Hegley, D., & Nagel, L. (1992). Adult judgments and fine-grained analysis of infant facial expressions: Testing the validity of a priori coding formulas. *Developmental Psychology, 28*, 1115.

Oster, H. (2005). The repertoire of infant facial expressions: an ontogenetic perspective. In J. Nadel & D. Muir (Eds.), *Emotional development: Recent research advances* (pp. 261-292). New York, NY, US: Oxford University Press.

Oster, H. (2003). *BabyFACS: Facial Action Coding System for Infants and Young Children*. Unpublished manual. New York University.

Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. New York: Oxford University Press.

Parkinson, B. (1997). Untangling the appraisal- emotion connection. *Personality and Social Psychology Review, 1(1)*, 62-79.

Parr, T. (2005). *The feelings book.* New York: Little, Brown books for young readers.

Parr, L.A., Waller, B.M., Burrows, A.M., Gothard, K.M., and Vick, S.-J. (2010). MaqFACS: A muscle-based facial movement coding system for the rhesus Macaque. *American Journal of Physical Anthropology 143*, 625-630.

Pavlenko, A. (2014). *The bilingual mind: And what it tells us about language and thought*. New York: Cambridge University Press.

Peltola MJ, Leppänen JM, Palokangas T, Hietanen JK. (2008). Fearful faces modulate looking duration and attention disengagement in 7-month-old infants. *Developmental Science, 11*, 60–68.

Perlman, S.B., Kalish, C.W., and Pollak, S.D. (2008). The role of maltreatment experience in children's understanding of the antecedents of emotion. *Cognition and Emotion, 22 (4),* 651-670.

Pinker, S. (1997). *How the mind works*. New York: Norton.

Plate, RC. Fulvio, JM.; Shutts, K; Green, CS; Pollak, SD. (2018). Probability learning: Changes in behavior across time and development. *Child Development, 89*, 205-218.

Pollak, S. D. (2009) The emergence of emotion: Experience, development and biology. In D. Cicchetti and M. R. Gunnar (Eds.), *Minnesota Symposia on Child Psychology: Meeting* *the challenge of translational research in child psychology* (Vol 35, p. 189-238). Hoboken, NJ: Wiley.

Pollak SD. (2015) Multilevel developmental approaches to understanding the effects of child maltreatment: Recent advances and future challenges. *Development and Psychopathology*, 2Nov27(4 Pt 2):1387-97.

Pollak, S. D., & Kistler, D. J. (2002). Early experience is associated with the development of categorical representations for facial expressions of emotion. *Proceedings of the National Academy of Sciences, 99(13)*, 9072-9076.

Pollak, S. D., Messner, M., Kistler, D. J., & Cohn, J. F. (2009). Development of perceptual expertise in emotion recognition. *Cognition, 110(2)*, 242-247.

Pollak, S.D. and Tolley-Schell, S.A.(2003). Selective attention to facial emotion in physically abused children. *Journal of Abnormal Psychology, 112*, 323-338.

Pollak, S.D., & Sinha, P. (2002). Effects of early experience on children's recognition of facial displays of emotion. *Developmental Psychology, 38*, 784-791.

Pollak, S.D., Cicchetti, D., Hornung, K., & Reed, A. (2000). Recognizing emotion in faces: Developmental effects of child abuse and neglect. *Developmental Psychology, 36*, 679-688. PMID: 9886228

Pollak, S.D., Vardi, S., Bechner, A.M.P., and Curtin, J.J. (2005). Physically abused children's regulation of attention in response to hostility. *Child Development, 76(5)*, 968-977.

Pruitt, D. G. & M. J. Kimmel (1977). Twenty years of experimental gaming: Critique, synthesis, and suggestions for the future. *Annual Review Psychology 28*, 363-392.

Quigley, K. S., & Barrett, L. F. (2014). Is there reliability and specificity of autonomic changes during emotional episodes? Guidance from the conceptual act theory and psychophysiology. *Biological Psychology, 98,*82-94.

Rehm, M. & Andre', E. (2005). Catch me if you can: exploring lying agents in social settings. 4th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS), ACM, New York, NY, USA, 937-944.

Reissland, N., Francis, B., & Mason, J. (2013). Can healthy fetuses show facial expressions of "pain" or "distress"? *PLoS One, 8(6),* e65530.

Reissland, N., Francis, B., Mason, J., & Lincoln, K. (2011). Do facial expressions develop before birth? *PLoS One, 6(8),* e24081.

Rickel, R., Marsella, S., Gratch, J., Hill, R., Traum, D. and Swartout, B. (2002). Towards a New Generation of Virtual Humans for Interactive Experiences, in IEEE Intelligent Systems July/August 2002, pp. 32-38.

Riggins v. Nevada (1992). 504 U.S. 127 (Kennedy, J., concurring), p. 142.

Rinn, W. E. (1984). The neuropsychology of facial expression: a review of the neurological and psychological mechanisms for producing facial expressions. Psychological bulletin, 95(1), 52.

Robbins, J. & Rumsey, A. (2008). Introduction: Cultural and linguistic anthropology and the opacity of other minds. *Anthropological Quarterly, 81,* 407-420.

Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: evidence for an accessibility model of emotional self-report. *Psychological Bulletin, 128(6)*, 934.

Roch-Levecq, A.-C. (2006). Production of basic emotions by children with congenital blindness: Evidence for the embodiment of theory of mind. *British Journal of Developmental Psychology, 24(3)*, 507-528.

Roseman, I. J. (2001). A model of appraisal in the emotion system: Integrating theory, research, and applications. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, methods, research* (pp. 68-91). New York: Oxford University Press.

Rosenberg, E. L., & Ekman, P. (1994). Coherence between expressive and experiential systems in emotion. *Cognition & Emotion, 8*, 201-229.

Rosenberg, E. L., & Ekman, P. (1995). Conceptual and methodological issues in the judgment of facial expressions of emotion. *Motivation and Emotion, 19(2)*, 111-138.

Rosenstein, D., & Oster, H. (1988). Differential facial responses to four basic tastes in newborns. *Child Development, 59*, 1555-1568.

Röttger-Rössler, B., Scheidecker, G., Funk, L., & Holodynski, M. (2015). Learning (by) feeling: A cross-cultural comparison of the socialization and development of emotions. *Ethos, 43(2)*, 187-220.

Rozin, P., Hammer, L., Oster, H., Horowitz, T., & Marmora, V. (1986). The child’s conception of food: Differentiation of categories of rejected substances in the 16 months to 5 year age range. *Appetite, 7*, 141–151.

Ruba, A. L., Johnson, K. M., Harris, L. T., & Wilbourn, M. P. (2017). Developmental changes in infants’ categorization of anger and disgust facial expressions. *Developmental Psychology*, *53*(10), 1826.

Rudoivic, O., Lee, J., Dai, M., Schuller, B., & Picard, R. W. (2018). Personalized machine learning for robot perception of affect and engagement in autism therapy. *Science Robotics, 3*, eaao6760. June 27, 2018.

Ruiz-Belda, M. A., Fernandez-Dols, J-M., Carrera, P., & Barchard, K. (2003). Spontaneous facial expressions of happy bowlers and soccer fans. *Cognition and Emotion, 17*, 315-326.

Russell, J. A. (1991). Culture and the categorization of emotions. *Psychological Bulletin*, 110(3), 426-450.

Russell, J. A. (1993). Forced-choice response format in the study of facial expressions. *Motivation and Emotion, 17*, 41-51.

Russell, J. A. (1994). Is there universal recognition of emotion from facial expressions? A review of the cross-cultural studies. *Psychological Bulletin, 115(1)*, 102-141.

Russell, J. A. (1995). Facial expressions of emotion: what lies beyond minimal universality? *Psychological Bulletin, 118(3),* 379-391.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110(1)*, 145-172.

Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology, 76*, 805-819.

Russell, J. A., & Widen, S. C. (2002a). A label superiority effect in children's categorization of facial expressions. *Social Development, 11(1),* 30-52.

Russell, J. A., & Widen, S. C. (2002b). Words versus faces in evoking preschool children's knowledge of the causes of emotions. *International Journal of Behavioral Development, 26(2)*, 97-103.

Russell, J. A., Hosie, C. D., Gray, C., Scott, C., Hunter, N., Banks, J. S., & Macaualy, M. C. (1998). The development of theory of mind in deaf children. *Journal of Child Psychology and Psychiatry, 39*, 903–910.

Rychlowska, M., Jack, R. E., Garrod, O. G. B., Schyns, P. G., Martin, J. D., & Niedenthal, P. M. (2017). Functional smiles: tools for love, sympathy, and war. *Psychological Science*, 28, 1259-1270.

Rychlowska, M., Miyamoto, Y., Matsumoto, D., Hess, U., Gilboa-Schechtman, E., Kamble, S., Muluk, H, Masuda, T., Niedenthal, P.M. (2015). Heterogeneity of long-history migration explains cultural differences in reports of emotional expressivity and the functions of smiles. *Proceedings of the National Academy of Sciences of the United States of America 112(19)*, E2429-36.

Saarimäki, H., Gotsopoulos, A., Jaaskelainen, I. P., Lampinen, J., Vuilleumier, P., Hari, R., Sams, M., & Nummenmaa, L. (2016). Discrete neural signatures of basic emotions. *Cerebral Cortex, 26*, 2563-2573.

Saarni, C., Campos, J. J., Camras, L. A., & Witherington, D. (2006). Emotional development: Action, communication, and understanding. In Damon, W., Lerner, R. M., & Eisenberg, N. (Eds.), *Handbook of child psychology, social, emotional, and personality development* (Vol. 3, p. 226-299). New York: John Wiley & Sons.

Sato, W., & Yoshikawa, S. (2004). The dynamic aspects of emotional facial expressions. *Cognition and Emotion, 18,* 701-710.

Scarantino, A. (2016). The philosophy of emotions and its impact on affective science. In L.F. Barrett, M. Lewis and J.M. Haviland-Jones (Eds.), *Handbook of Emotions*, 4th Ed. (p. 3-48). New York: Guilford Press.

Scarantino, A., & Griffiths, P. (2011). Don't give up on basic emotions. *Emotion Review, 3(4),* 444-454.

Scherer, K. R., Mortillaro, M., & Mehu, M. (2017). Facial expression is driven by appraisal and generates appraisal inferences. In J-M. Fernandez-Dols & J. A. Russell (Eds.), *The Science of Facial Expression* (p. 353-374). New York: Oxford U Press.

Scherer, K. R., Mortillaro, M., Rotondi, I., Sergi, I., & Trznadel, S. (2018). Appraisal-driven facial actions as building blocks for emotion inference. *Journal of Personality and Social Psychology*, *114*(3), 358-379.

Schwartz, G. M., Izard, C. E., & Ansul, S. E. (1985). The 5-month-old's ability to discriminate facial expressions of emotion. Infant Behavior and Development, 8(1), 65-77.

Searle, J. R. (1992). The rediscovery of the mind: MIT press.

Searle, J. R. (2004). Mind: a brief introduction: Oxford University Press.

Sears, M. S., Repetti, R. L., Reynolds, B. M., & Sperling, J. B. (2014). A naturalistic observational study of children's expressions of anger in the family context. *Emotion, 14(2)*, 272.

Serrano, J. M., Iglesias, J., & Loeches, A. (1992). Visual discrimination and recognition of facial expressions of anger, fear, and surprise in 4-to 6-month-old infants. *Developmental Psychobiology, 25(6),* 411-425.

Shackman, J. E., & Pollak, S. D. (2014). Impact of physical maltreatment on the regulation of negative affect and aggression. *Development and Psychopathology*, *26*, 1021-1033.

Shackman, J. E., Fatani, S., Camras, L. A., Berkowitz, M. J., Bachorowski, J. A., & Pollak, S. D. (2010). Emotion expression among abusive mothers is associated with their children's emotion processing and problem behaviours. *Cognition and Emotion, 24*(8), 1421-1430.

Shackman, J.E., Shackman, A.J., and Pollak, S.D. (2007). Physical abuse amplifies attention to threat and increases anxiety in children. *Emotion, 7*, 838-852.

Shariff, A. F., & Tracy, J. (2011). What are emotion expressions for? *Current Directions in Psychological Science, 20(6)*, 395-399.

Shaver, P, Schwartz, J., Kirson, D, & O'Connor, C. (1987). Emotion knowledge: Further exploration of a prototype approach. *Journal of Personality and Social Psychology, 52,*1061-1086.

Shepard, R. N., & Cooper, L. A. (1992). Representation of colors in the blind, color-blind, and normally sighted. *Psychological Science, 3(2),* 97-104.

Shiota, M.N., Campos, B., & Keltner, D. (2003). The faces of positive emotion: Prototype displays of awe, amusement, and pride. *Annals of the New York Academy of Sciences, 1000,* 296-299.

Siegel, E. H., Sands, M. K., Van den Noortgate, W., Condon, P., Chang, Y., Dy, J., Quigley, K. S., & Barrett, L. F. (2018). Emotion fingerprints or emotion populations? A meta-analytic investigation of autonomic features of emotion categories. *Psychological Bulletin, 144(4),* 343-393.

Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review, 4,* 108–131.

Smith, L. B., Jayaraman, S., Clerkin, E., & Yu, C. (2018). The developing infant creates a curriculum for statistical learning. *Trends in Cognitive Sciences*, /doi.org/10.1016/j.tics.2018.02.004

Smith, T.W. (2016). The book of human emotion. Little, Brown & Company: New York, NY.

Snyder, P. J., Kaufman, R., Harrison, J., & Maruff, P. (2010). Charles Darwin's emotional expression "experiment" and his contribution to modern neuropharmacology. *Journal of the History of the Neurosciences*, 19:2, 158-170, doi: 10.1080/09647040903506679

Soken, N. H., & Pick, A. D. (1999). Infants' perception of dynamic affective expressions: do infants distinguish specific expressions? *Child Development, 70(6)*, 1275-1282.

Sorenson, E. R. (1975). Culture and the expression of emotion. In T. R. Williams (Ed.), *Psychological Anthropology* (pp. 361-372). Chicago: Aldine.

Srinivasan, R., and Martinez, A.M. (2018). Cross-cultural and cultural-specific production and perception of facial expressions of emotion in the wild. IEEE Transactions on Affective Computing. ArXiv preprint, available at https://arxiv.org/pdf/1808.04399.pdf

Sroufe, A. L. (1996). *Emotional development: The organization of emotional life in the early years*. New York, NY: Cambridge University Press.

Stebbins, G. & Delsarte, F. (1887/2013). Delsarte system of expression. E. S. Werner.

Stein, M., Ottenberg, P., & Roulet, N. (1958). A study of the development of olfactory preferences. *Archives of Neurology & Psychiatry, 80,* 264–266.

Stemmler, G. (2004). Physiological processes during emotion. In P. Philippot & R. S. Feldman (Eds.), *The regulation of emotion* (pp. 33–70). Mahwah, NJ: Erlbaum & Associates.

Stemmler, G., Aue, T., & Wacker, J. (2007). Anger and fear: Separable effects of emotion and motivational direction on somatovisceral responses. *International Journal of Psychophysiology, 66(2),* 141-153.

Stenberg, C. R., Campos, J. J., & Emde, R. N. (1983). The facial expression of anger in seven-month-old infants. *Child Development*, 178-184.

Sterling, P. (2012). Allostasis: a model of predictive regulation. *Physiology & Behavior*, *106(1),* 5-15.

Stephens, C. L., Christie, I.C., & Friedman, B. H. (2010). Autonomic specificity of basic emotions: evidence from pattern classification and cluster analysis. *Biological Psychology, 84,* 463-73.

Stevenson, R. J., Oaten, M. J., Case, T. I., Repacholi, B. M., & Wagland, P. (2010). Children's response to adult disgust elicitors: Development and acquisition. *Developmental Psychology, 46(1),* 165.

Sullivan, M. W., & Lewis, M. (2003). Contextual determinants of anger and other negative expressions in young infants. *Developmental Psychology, 39(4),* 693.

Susskind, J. M., Lee, D. H., Cusi, A., Feiman, R., Grabski, W., & Anderson, A. K. (2008). Expressing fear enhances sensory acquisition. *Nature Neuroscience, 11(7),* 843-850.

Tassinary, L. G., & Cacioppo, J. T. (1992). Unobservable facial actions and emotion. *Psychological Science, 3(1),* 28-33.

Tinbergen, Niko. (1953). *The Herring Gull's World*. London: Collins.

Todorov, A. (2017). *Face value: The irresistible influence of first impressions*. Princeton, NJ: Princeton University Press.

Tooby, J., & Cosmides, L. (2008). The evolutionary psychology of the emotions and their relationship to internal regulatory variables. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), The Handbook of Emotions (3rd ed.). New York, NY: The Guilford Press.

Tovote, P., Fadok, J. P., & Lüthi, A. (2015). Neuronal circuits for fear and anxiety. *Nature Reviews Neuroscience, 16(6)*, 317-331.

Tracy, J. L., & Matsumoto, D. (2008). The spontaneous expression of pride and shame: Evidence for biologically innate nonverbal displays. *Proceedings of the National Academy of Sciences of the United States of America, 105(33)*, 11655-11660.

Tracy, J. L., & Robins, R. W. (2008). The nonverbal expression of pride: evidence for cross21cultural recognition. *Journal of Personality and Social Psychology, 94(3)*, 516.

Tracy, J., & Randles, D. (2011). Four models of basic emotions: A review of Ekman and Cordaro, Izard, Levenson, and Panksepp and Watt. *Emotion Review, 3(4)*, 397-405.

Turati, C. (2004). Why Faces Are Not Special to Newborns An Alternative Account of the Face Preference. *Current Directions in Psychological Science, 13(1)*, 5-8.

Turati, C., Di Giorgio, E., Bardi, L., & Simion, F. (2010). Holistic face processing in newborns, 3-month-old infants and adults: Evidence from the composite face effect. *Child Development, 81*, 1894-1905.

Turner, J. R., & Carroll, D. (1985). Heart rate and oxygen consumption during mental arithmetic, a video game, and graded exercise: Further evidence of metabolically exaggerated cardiac adjustments? *Psychophysiology, 22*, 261–267.

Tversky, A. (1977). Features of similarity. *Psychological Review, 84*, 327-352.

Tversky, A.; Kahneman, D. (1974). "Judgment under Uncertainty: Heuristics and Biases" *Science, 185 (4157)*, 1124–1131.

Vaillant-Molina, M., & Bahrick, L. E. (2012). The role of intersensory redundancy in the emergence of social referencing in 5½-month-old infants. *Developmental Psychology*, *48*(1), 1.

Vaish, A., & Striano, T. (2004). Is visual reference necessary? Contributions of facial versus vocal cues in 12-month-olds' social referencing behavior. *Developmental Science, 7*(3), 261-269.

Valente, D., Theurel, A., & Gentaz, E. (2017). The role of visual experience in the production of emotional facial expressions by blind people: a review. *Psychon Bull Rev*. doi:10.3758/s13423-017-1338-0

Vallacher, R. R., & Wegner, D. M. (1987). What do people think their doing? Action identification and human behavior. *Psychological Review, 94*, 3-15.

Valliant-Molina, M. & Bahrick, L.E. (2012) The role of intersensory redundancy in the emergence of social referencing in 51⁄2-month-old infants . *Developmental Psychology*, 48, 1-9.

Valstar, M., Zafeiriou, S., & Pantic, M. (2017). Facial actions as social signals. In J. K. Burgoon, N. Magnenat-Thalmann, M. Pantic & A. Vinciarelli (Eds.), Social Signal Processing (p. 123-154).

Van den Stock, J., Righart, R., & De Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion, 7(3),* 487.

Vick, S.-J., Waller, B.M., Parr, L.A., Smith Pasqualini, M., and Bard, K.A. (2007). A cross-species comparison of facial morphology and movement in humans and chimpanzees using FACS. *Journal of Nonverbal Behavior, 31*, 1-20.

Voorspoels, W., Vanpaemel, W., & Storms, G. (2011). A formal idea-based account of typicality. *Psychon Bull Rev, 18*, 1006-1014.

Wade, N. J. (2016). Faces and photography in 19th-century visual science. *Perception, 45*, 1008-1035.

Wager, T. D., Kang, J., Johnson, T. D., Nichols, T. E., Satpute, A. B., & Barrett, L. F. (2015). A Bayesian model of category-specific emotional brain responses. *PLOS Computational Biology*. 11(4): e1004066. doi:10.1371/ journal.pcbi.1004066

Walker-Andrews, A. S. (2005). Perceiving social affordances: The development of emotion understanding. The development of social cognition and communication, 93-116.

Walker-Andrews, A. S. (1997). Infants' perception of expressive behaviors: differentiation of multimodal information. Psychological bulletin, 121(3), 437.

Walker-Andrews, A. S., & Dickson, L. R. (1997). Infants’ understanding of affect. The development of social cognition, 161-186.

Walker-Andrews, A. S., & Lennon, E. (1991). Infants' discrimination of vocal expressions: Contributions of auditory and visual information. Infant Behavior and Development, 14(2), 131-142.

Wallbott, H. G. (1988). In and out of context: Influences of facial expression and context information on emotion attributions. British Journal of Social Psychology, 27, 357-369.

Walle, E. A., & Campos, J. J. (2014). The development of infant detection of inauthentic emotion. Emotion, 14(3), 488.

Wang, N., Marsella, S. and Hawkins, T. (2008) Individual Differences in Expressive Response: A Challenge for ECA Design, in Proceedings of Autonomous Agents and Multi-Agent Systems (AAMAS).

Wang, X., Peelen, M. V., Han, Z., He, C., Caramazza, A., & Bi, Y. (2015). How Visual Is the Visual Cortex? Comparing Connectional and Functional Fingerprints between Congenitally Blind and Sighted Individuals. J Neurosci, 35(36), 12545-12559. doi:10.1523/JNEUROSCI.3914-14.2015

Wehrle, T., Kaiser, S., Schmidt, S., & Scherer, K. R. (2000). Studying the dynamics of emotional expression using synthesized facial muscle movements. *Journal of Personality and Social Psychology, 78,* 105-119.

Weiss, B., & Nurcombe, B. (1992). Age, clinical severity, and the differentiation of depressive psychopathology: A test of the orthogenetic hypothesis. *Development and Psychopathology, 4(1),* 113-124.

Widen, S. C. & Russell, J. A. (2013). Children’s recognition of disgust in others. *Psychological Bulletin, 139*, 271-299.

Widen, S. C. (2016). The development of children's concepts of emotion. In L. F. Barrett, M. Lewis, J. M., & Haviland-Jones (Eds.), *Handbook of emotions* (4th ed., pp. 307-318). New York: Guilford Press.

Wierzbicka, A. (1986). Human emotions: universal or culture-specific? American Anthropologist, 88(3), 584-594.

Wierzbicka, A. (2014). *Imprisoned in English. The hazards of English as a default language*. Oxford University Press.

Wieser, M. J., & Brosch, T. (2012). Faces in context: a review and systematization of contextual influences on affective face processing. *Front Psychol, 3*, 471. doi:10.3389/fpsyg.2012.00471

Wilson, J. P. & Rule, N. O. (2015). Facial trustworthiness predicts extreme criminal-sentencing outcomes. *Psychological Science, 26*, 1325-1331.

Wilson, J. P. & Rule, N. O. (2016). Hypothetical sentencing decisions are associated with actual capital punishment outcomes: The role of facial trustworthiness. *Social Psychological and Personality Science*, https://doi.org/10.1177/1948550615624142

Wilson-Mendenhall, C., Barrett, L. F., & Barsalou, L. W. (2013). Neural evidence that human emotions share core affective properties. *Psychological Science, 24*, 947-956. PMC4015729

Wilson-Mendenhall, C.D., Barrett, L.F., & Barsalou, L.W. (2015). Variety in emotional life: Within-category typicality of emotional experiences is associated with neural activity in large-scale brain networks. *Social Cognitive and Affective Neuroscience, 10*, 62-71.

Wilson-Mendenhall, C. D., Barrett, L. F., Simmons, W. K., Barsalou, L. W. (2011). Grounding emotion in situated conceptualization. *Neuropsychologia, 49*, 1105-1127.

Wilson, T. D., Houston, C. E., Etling, K. M., Brekke, N. (1996). A new look at anchoring effects: Basic anchoring and its antecedents. *Journal of Experimental Psychology: General*, 125(4): 387–402.

Witherington, D. C., Campos, J. J., & Hertenstein, M. J. (2008). Chapter Sixteen Principles of Emotion and its Development in Infancy. *Blackwell handbook of infant development,* 427.

Witherington, D. C., Campos, J. J., Harriger, J. A., Bryan, C., & Margett, T. E. (2010). Emotion and its development in infancy. *Wiley-Blackwell Handbook of Infant Development, The, Volume 1, Second Edition*, 568-591.

Wolff, P. H. (1987). *The development of behavioral states and the expression of emotions in early infancy: New proposals for investigation*. University of Chicago Press.

Wood, A., Martin, J. D., Alibali, M. W., & Niedenthal, P. M. (2018). A sad thumbs up: incongruent gestures and disrupted sensorimotor activity both slow processing of facial expressions. *Cognition and Emotion*, 1-14.

Wood, A., Rychlowska, M., & Niedenthal, P. (2016). Hetergeneity of long-history migration predictions emotion recognition accuracy. *Emotion, 16*, 413-420.

Wörmann, V., Holodynski, M., Kärtner, J., & Keller, H. (2012). A cross-cultural comparison of the development of the social smile: A longitudinal study of maternal and infant imitation in 6-and 12-week-old infants. *Infant Behavior and Development, 35(3),* 335-347.

Wundt, W. (1998/1897). *Outlines of Psychology* In: Judd, C.H. (Trans.). Leipzig: Wilhelm Engelmann.

Yan, F., Dai, S. Y., Akther, N., Kuno, A., Yanagihara, T., & Hata, T. (2006). Four-dimensional sonographic assessment of fetal facial expression early in the third trimester. *International Journal of Gynecology & Obstetrics*, *94*(2), 108-113.

Yik, M., & Russell, J. A. (1999). Interpretation of faces: A cross-cultural study of a prediction from Fridlund's theory. *Cognition & Emotion, 13*(1), 93-104.

Yin, Y., Nabian, M., Ostadabbas, S., Fan, M., Chou, C., & Gendron, M. (2018). Facial expression and peripheral physiology fusion to decode individualized affective experiences. 2nd IJCAI Workshop on Artificial Intelligence in Affective Computing: *Proceedings of Marching Learning Research, 1*, 1-16.

Yitzhak, N., Giladi, N., Gurevich, T., Messinger, D. S., Prince, E. B., Martin, K., & Aviezer, H. (2017). Gently does it: Humans outperform a software classifier in recognizing subtle, nonstereotypical facial expressions. *Emotion, 8,* 1187-1198.

Young, A. W., & Burton, A. M. (2018). Are we face experts? Trends in Cognitive Sciences, 22(2), 100-110.

Yu H, Garrod OGB, Schyns PG. (2012). Perception-driven facial expression synthesis. *Comput. Graph,* 36,152–62

Zafeiriou, S., Papaioannou, A., Kotsia, I., Nicolaou, M., & Zhao, G. (2016). Facial Affect``In-The-Wild. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (pp. 36-47).

Zebrowitz, L. A. (1997). *Reading faces: Window to the soul?* Westview Press.

Zebrowitz, L. A. (2017). First impressions from faces. *Current Directions in Psychological Science, 26(3)*, 237-242.

Zoll, C., Enz, S., H., S., Aylett, R., & Paiva, A. (2006). *Fighting Bullying with the Help of Autonomous Agents in a Virtual School Environment.* Paper presented at the 7th International Conference on Cognitive Modelling (ICCM-06), Trieste, Italy.

List of Tables

Table 1. A comparison of the facial configurations listed as the expressions of selected emotion categories

Emotion Category	Proposed Expressive Configurations Described as Facial Action Units					
	Matsumoto, Keltner, Shiota, O'Sullivan & Frank (2008)		Cordaro, Sun, Keltner, Kamble, Huddar & McNeil (2017)		Keltner et al. (in press)	Physical Description
	Darwin's (1872) Description	Observed in reseearch	Reference Configuration Used	International Core Pattern		
Amusement	Not listed	Not listed	6, 12, 26 or 27, 55 or 56, a "head bounce" (Shiota, Campos & Keltner, 2003)	6, 7, 12, 16, 25, 26 or 27, 53	6+7+12+25+26+53	Head back, Duchenne smile (6, 7, 12), lips separated, jaw dropped
Anger	4+ 5+ 24+ 38	4 + 5 or 7 + 22+23+24	4 +5 + 7 + 23 (Ekman, Levenson & Friesen, 1983)	4, 7	4+5+17+23+24	Brows furrowed, eyes wide, lips tightened and pressed together
Awe	Not listed	Not listed	1, 5, 26 or 27, 57 and visible inhalation (Shiota et al., 2003)	1, 2, 5, 12, 25, 26 or 27, 53	Not listed	
Contempt	9+ 10+ 22+ 41+ 61 or 62	12 (unilateral) + 14 (unilateral)	12 + 14 (Ekman et al., 1983)	4, 14, 25	Not listed	
Disgust	10+ 16+ 22+ 25 or 26	9 or 10, 25 or 26	9+15+16 (Ekman et al., 1983)	4, 6, 7, 9, 10, 25, 26 or 27	7+9+19+25+26	Eyes narrowed, nose wrinkled, lips parted, jaw dropped, tongue show
Embarrassment	Not listed	Not listed	12, 24, 51, 54, 64 (Keltner & Buswell, 1997)	6, 7, 12, 25, 54, participant dampens smile with 23, 24, frown, etc.)	7+12+15+52+54+64	Eyelids narrowed, controlled smile, head turned and down, (not scored with FACS: hand touches face)
Fear	1+ 2+ 5+ 20	1+2+4+5+20, 25 or 26	1+2+4+5+7+20+26 (Ekman et al., 1983)	1, 2, 5, 7, 25, 26 or 27, participant suddenly shifts entire body backwards in chair	1+2+4+5+7+20+25	Eyebrows raised and pulled together, upper eyelid raised, lower eyelid tense, lips parted and stretched
Happiness	6 + 12	6 + 12	6+12 (Ekman et al., 1983)	6, 7, 12, 16, 25, 26 or 27	6+7+12+25+26	Duchenne smile (6, 7, 12)
Pride	Not listed	Not listed	6, 12, 24, 53, a straightening of the back and pulling back of the shoulders to expose the chest (Shiota et al., 2003)	7, 12, 53, participant sits up straight	53+64	Head up, eyes down
Sadness	1 + 15	1+15, 4,17	1+4+5 (Ekman et al., 1983)	4, 43, 54	1+4+6+15+17	Brows knitted, eyes slightly tightened, lip corners depressed, lower lip raised
Shame	Not listed	Not listed	54, 64 (Keltner & Buswell, 1997)	4, 17, 54	54+64	Head down, eyes down
Surprise	1+ 2 + 5+ 25 or 26	1+2+5+25 or 26	1+2+5+26 (Ekman et al., 1983)	1, 2, 5, 25, 26 or 27	1+2+5+25+26	Eyebrows raised, upper eyelid raised, lips parted, jaw dropped

Note. Darwin's description taken from Matsumoto et al. (2008), Table 13.1. International core patterns (ICPs) refer to expressions of 22 emotion categories that are thought to be conserved across cultures, taken from Cordaro et al. (2017), Tables 4, 5 and 6. A plus sign means "with"; these action units would appear simultaneously. A comma means "sometimes with"; these action units are statistically the most probable to appear, but do not necessarily need to happen simultaneously (David Cordaro, personal communication, 11/11/2018).

Table 2: Criteria used to evaluate the empirical evidence

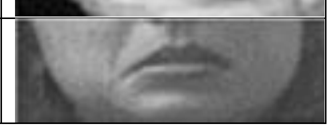
	Expression Production	Emotion Perception
Reliability	When a person is sad, the proposed expression (a frowning facial configuration) should be observed more frequently than would be expected by **chance**. Likewise, for every other emotion category that is subject to a commonsense belief. Reliability is related to a **forward inference**: given that someone is happy, what is the likelihood of observing a smile, *p[set of facial muscle movements | emotion category]*. Chance means that facial configurations occur randomly with no predictable relationship to a given emotional state. This would mean that the facial configuration in question carries no information about the presence or absence of an emotion category. For example, in an experiment that observes the facial configurations associated with instances of happiness and anger, chance levels of scowling or smiling would be 50%. Reliability also depends on the base rate: how frequently people make a particularly facial configuration overall. For example, if a person frequently makes a scowling facial configuration during an experiment examining the expressions of anger, sadness and fear, he will seem to be consistently scowling in anger when in fact he is scowling indiscriminately. Reliability rates between 70% and 90% provide strong evidence for the commonsense view, between 40% and 69% provide moderate support for the commonsense view, and between 20% and 39% provide weak support (Ekman, 1994; Haidt & Keltner, 1999; Russell, 1994).	When a person makes a scowling facial configuration, perceivers should consistently infer that the person is angry. Likewise, for every facial configuration that has been proposed as *the* expression of a specific emotion category. That is, perceivers must consistently make a **reverse inference**: given that someone is scowling, what is the likelihood that he is angry, *p[emotion category | set of facial muscle movements]*. Chance means that emotional states occur randomly with no predictable relationship to a given facial configuration. This would mean that the presence or absence of an emotion category cannot be inferred from the presence or absence of the facial configuration. For example, in an experiment that observes how people perceive 51 different facial configurations, chance levels for correctly labeling a scowling face as anger would be 2%. Reliability also depends on the base rate: how frequently people use a particular emotion label or make a particular emotional inference. For example, if a person frequently labels facial configurations as “angry” during an experiment examining scowling, smiling and frowning faces, she will seem to be consistently perceiving anger when in fact she is labeling indiscriminately. Reliability rates between 70% and 90% provide strong evidence for the commonsense view, between 40% and 69% provide moderate support for the commonsense view, and between 20% and 39% provide weak support (Ekman, 1994; Haidt & Keltner, 1999; Russell, 1994).
Specificity		

	If a facial configuration is diagnostic of a specific emotion category, then the facial configuration should express instances of one and only one emotion category better than chance; it should not consistently express instances of any other mental event (emotion or otherwise) at better than chance levels. For example, to be considered *the* expression of anger, a scowling facial configuration must not express sadness, confusion, indigestion, an attempt to socially influence, etc. at better than chance levels. Estimates of specificity, like reliability, depend on base-rates and on how chance levels are defined.	If a frowning facial configuration is perceived as *the* diagnostic expression of sadness, then a frowning facial configuration should only be labeled as sadness (or sadness should only be inferred from a frowning facial configuration) at above chance levels. And it should not be consistently perceived as expressions of any mental states other than sadness at better than chance levels. Estimates of specificity, like reliability, depend on base-rates and on how chance levels are defined.
Generalizability	Patterns of reliability and specificity should replicate across studies, particularly when different populations are sampled, such as infants, congenitally blind individuals and individuals sampled from diverse cultural contexts, including small-scale, remote cultures. High generalizability across different circumstances ensures that scientific findings are generalizable.	Patterns of reliability and specificity should replicate across studies, particularly when different populations are sampled, such as infants, congenitally blind individuals and individuals sampled from diverse cultural contexts, including small-scale, remote cultures. High generalizability across different circumstances ensures that scientific findings are generalizable.
Validity	Even if a facial configuration is consistently and uniquely observed in relation to a specific emotion category across many studies (strong generalizability), it is necessary to demonstrate that the person in question is really in the expected emotional state. This is the only way that a given facial configuration leads to **accurate** inferences about a person's emotional state. A facial configuration is valid as a **display** or a signal for emotion if and only if it is strongly associated with other measures of emotion, preferably those that are objective and do not rely on anyone's subjective report (i.e., a facial configuration should be strongly and consistently related to perceiver-independent evidence about the emotional state of the expresser).	Even if a facial configuration is consistently and uniquely labeled with a specific emotion word across many studies (strong generalizability), it is necessary to demonstrate that the person making the facial configuration is really in the expected emotional state. This is the only way that a given perception or inference of emotion is **accurate**. A perceiver can only be said to be recognizing an emotional expression if and only if the person being perceived is verifiably in the expected emotional state.

Note: Reliability is also related to **sensitivity**, **consistency, informational value**, and the **true positive rate** (for further description, see Figure 3). Specificity is related to uniqueness, **discreteness**, the **true negative rate** and **referential specificity**. In principle, we can also ask more parametrically whether there is a link between the intensity of an emotional instance and the intensity of facial muscle contractions, but scientists rarely do.

.

Table 3: The Facial Action Coding System (FACS; Ekman & Friesen, 1978) codes for adults

AU	*Description*	*Facial muscles (type of activation)*	
1	Inner brow raiser	Frontalis (pars medialis)	
2	Outer brow raiser	Frontalis (pars lateralis)	
4	Brow lowerer	Corrugator supercilii, depressor supercilii	
5	Upper lid raiser	Levator palpebrae superioris	
6	Cheek raiser	Orbicularis oculi (pars orbitalis)	
7	Lid tightener	Orbicularis oculi (pars palpebralis)	
9	Nose wrinkle	Levator labii superioris alaquae nasi	
10	Upper lip raiser	Levator labii superioris	
11	Nasolabial deepener	Zygomaticus minor	
12	Lip corner puller	Zygomaticus major	
13	Cheeks puffer	Levator anguli oris	
14	Dimpler	Buccinator	
15	Lip corner depressor	Depressor anguli oris	
16	Lower lip depressor	Depressor labii inferioris	

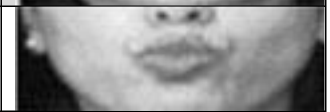
17	Chin raiser	Mentalis	
18	Lip puckerer	Incisivii labii superioris and incisivii labii inferioris	
20	Lip stretcher	Risorius w/ platysma	
22	Lip funneler	Orbicularis oris	
23	Lip tightener	Orbicularis oris	
24	Lip pressor	Orbicularis oris	
25	Lips part	Depressor labii inferioris or relaxation of mentalis, or orbicularis oris	
26	Jaw drop	Masseter, relaxed temporalis and internal terygoid	
27	Mouth stretch	Pterygoids, digastric	
28	Lip suck	Orbicularis oris	
41	Lid Droop		
42	Slit		
43	Eyes Closed		
44	Squint		
45	Blink		
46	Wink		

Table 4: Reliability and specificity: A summary of the evidence

	Reliability	Specificity
Expression Production		
Adults, Developed, Spontaneous, Lab	weak	unknown
Adults, Developed, Spontaneous, Naturalistic	weak	unknown
Adults, Developed, Posed	weak to strong	unknown
Adults, Remote, Spontaneous	unclear	unknown
Adults, Remote, Posed	weak to strong	unknown
Newborns, Infants, Toddlers	unsupported	unsupported
Congenitally Blind	unsupported to weak	unsupported
Emotion Perception		
Adults, Developed, Choice-From-Array	moderate to strong	unknown
Adults, Developed, Reverse Inference (with Choice-From-Array)	moderate	moderate
Adults, Developed, Free-Labeling	weak to moderate	weak
Adults, Developed, Virtual Humans	unknown	unknown
Adults, Remote, Choice-From-Array (before 2008)	moderate to strong	unknown
Adults, Remote, Choice-From-Array (after 2008)	weak to moderate	unsupported
Adults, Remote, Free-Labeling (before 2008)	unsupported to strong	variable
Adults, Remote, Free-Labeling (after 2008)	unsupported	unsupported
Infants, Young Children	unsupported	unsupported

Note. Criteria were adopted from Haidt & Keltner (1999), who suggest that reliability rates of 70±90% are considered strong evidence for universal emotion perception (following Ekman, 1994a); presumably, this would also hold for studies of expression production. Weak evidence is in the range of 20±40% (following Russell, 1994). By interpolation, reliability between 41% and 69% would be considered moderate evidence for reliability. Reliability estimates below 20% are interpreted as findings that clearly do not support the reliability hypothesis. We also adopted these criteria for specificity findings. Developed = studies of participants from the U.S. and other more urban countries. Spontaneous = spontaneous facial movements. Posed = posed facial configurations. Remote = studies of participants from small-scale, remote samples.

Table 5: Common tasks for measuring explicit emotion perception

Concerns	Additional Observations
General Considerations	
Test-retest reliability is rarely evaluated but is critical. A number of contextual factors are known to influence judgments, including a perceiver's internal state.	Test-retest assessments are rarely done for practical reasons.
Participants are typically asked to infer emotional meaning in exaggerated facial configurations. This reduces the ecological validity of the findings for how people infer emotional meaning in faces in the real world. The facial configurations used in most experiments (see Figure 4) are caricatures – they are exaggerated to maximally distinguish one from the another. Caricatures are easier to label (categorize) than are typical stimuli, particularly when the categories in question are highly interrelated (Goldstone, Steyvers, & Rogosky, 2003).	Exaggerated facial configurations have greater "source clarity" (Ekman, Friesen & Ellsworth, 1972)
Participants are typically asked to infer emotional meaning in highly selected facial configurations.	In early studies, a smaller set of exaggerated facial configurations were culled from much larger sets of posed faces (involving several thousand faces; for a discussion, see Gendron & Barrett, 2017; Russell, 1994).
Participants are typically asked to infer emotional meaning in static, non-moving facial configurations (i.e., in photographs rather than movies). This reduces the ecological validity of the findings for how people infer emotional meaning in faces in the real world. In the real world, people have to infer when a set of	There is information in the dynamics of facial movements (Jack & Schyns, 2017; Krumhuber, Kappas & Manstead, 2013), but dynamic facial movements, particularly when they are spontaneous, do not always produce higher levels of agreement in emotion perception studies. Dynamic movements add realism,

movements begin and end; this is called discrimination or detection).	intensity and improve levels of agreement primarily when movements are degraded or are artificial
Participants are typically asked to infer emotional meaning in posed, rather than spontaneous, facial configurations.	Spontaneous or candid facial configurations typically produce much lower levels of agreement in emotion perception studies (e.g., Kayyal & Russell, 2013; Naab & Russell, 2007).
Only a single task used in most experiments (i.e., participants are asked to infer emotion in facial configurations via one method of responding). Ideally, multiple tasks should be used with the same population of participants to see if convergent results are obtained.	This approach is rarely taken, but for an example, see Crivelli et al., 2016; Gendron et al., 2014; Gendron et al., 2018).
Most experiments ask participants to infer emotion in a disembodied face, alone, without context. This reduces the ecological validity of the findings for how people infer emotional meaning in faces in the real world.	A growing number of experiments now show that context is an important, and sometimes dominant, source of information when people infer emotional meaning in a facial configuration. See Box 3 in SOM. For example, Situational information tends to dominate perception of emotion in faces both when situations are common, everyday (Carrera-Levillain & Fernández-Dols, 1994) and even when situations are more ambiguous than the exaggerated facial configurations being judged (Carroll & Russell, 1996, Study 3).
Many studies do not report evidence about the specificity of emotion perceptions, or the frequency with which people infer the non-intended emotional meaning to a facial configuration.	
Until recently, the large majority of experiments included only one pleasant emotion category (happiness) among several unpleasant emotion categories (anger, fear, sadness, etc.). This may be one reason that agreement rates are so high for smiles.	In the last few years, experiments are now including a larger variety of pleasant emotion categories (pride, awe, gratitude, etc.), but there continues to be debate over whether or not they expect these emotion categories are expressed with consistent, specific facial configurations.

Choice-From Array: matching photos of facial configurations and emotion words (with or without brief stories) Response options are limited to those provided in the task	
Words influence how the brain processes visual inputs from faces (e.g., Gendron et al., 2012; Doyle & Lindquist, 2018). Stories can prime action perceptions, as well (Gendron et al., in press). More generally, choice-from-array tasks have been shown to encourage biased perceptual responding using a signal detection analysis (e.g., DeCarlo 2012).	Choice-from-array tasks are easy and efficient.
The fact that participants are exposed to the same facial configurations and emotion words over and over allows them to learn the intended pairings even if they don't know them to begin with (Nelson & Russell, 2016).	
An emotion word does not necessarily have a one-to-one correspondence to a single emotion category for all people in a given culture (i.e., they may differ in emotional granularity; Barrett, 2004, 2017; Lindquist & Barrett, 2008) or people from different cultures.	Concerns about individual word meaning is why choice-from-array using stories is preferable. Also, choice-from-array tasks are usually straightforward for participants to understand.
A small range of answers are pre-determined by the experimenter, making it easier for participants to provide the answers scientists expect. For example, by constraining which words participants were allowed to choose from, frowns were consensually labeled as fear, wide-eyed gasping faces were labeled as surprise (Russell, 1993). Scowling faces are more likely to be perceived as fearful when paired with the description of danger (Carroll & Russell, 1996, Study 1) and appear determined or puzzled depending on the story they are presented with (Carroll & Russell, 1996, Study 2).	Choice-from-array responses are easy for scientists to score. Most studies using continuous judgments (rather than forced choice) find that participants do not infer emotional meaning in facial configurations in a yes/no or on/off sort of way (Russell, 1994).

People are asked to make yes/no decisions about assigning a facial configuration to an emotion category. Multiple emotion words may apply to a single configuration (i.e., people might infer more than one emotional meaning in a face), but the option to infer multiple emotional meanings rarely given to participants.	Continuous judgments, such as on a Likert-type scale ranging from one to seven, would solve both of these problems, and also allow analysis of the similarity among facial configurations (which evidence shows is important, e.g., Jack et al., 2016; Kayal & Russell, 2013). Similarity allows scientists to discover the emotional meanings that people implicitly assign to a facial configuration, rather than having people explicitly state them (see further discussion of similarity below).
A participant might decide that no emotion word provided applies to a facial configuration, but the option to respond this way is rarely given to participants (they are usually forced to choose an emotion word; for discussion, see Frank & Stennett, 2001).	See Cordaro et al. (2016) for an example of this design feature.
If a participant hears a story and is choose between two faces (e.g., a scowl and smile), she can give the expected answer (e.g., scowl) simply by figuring out that smile is NOT correct. For example, after hearing a story about anger, a participant is shown a scowl and a smile and can choose the scowl merely by realizing the smile is not correct (on the basis of valence). This is similar to getting an answer right on a multiple-choice test by eliminating all the alternatives—you don't actually know the right answer, but you figured it out because of the structure of the task. A similar point can be made about showing a single face and asking participants to label it with a word by selecting from among a small set of options. Participants use a process of elimination strategy: words that are not chosen on prior trials are selected more frequently, inflating agreement levels (DiGirolamo & Russell, 2017).	If a participant hears a story about anger and must choose between a scowl and a smile, she can figure out that the scowl is correct merely because she is distinguishing between negative (scowl) and positive (smile). If a participant hears a story about anger and must choose between a scowl and a frown, he can figure out that the scowl is correct merely because he is distinguishing between high arousal (scowl) and low arousal (frown).

In tasks that involve brief stories or vignettes about emotion, only one typical story is offered for each emotion category, making it more difficult to observe any variation within a category.	
Free Sorting: photos of facial configurations are sorted into groupings, such that each grouping represents a category **Face-to-Cue Matching: matching photos of facial configurations to a recording of posed vocalization**	
Most participants still spontaneously use words to guide their sorting and organize their groupings.	Ideal for preverbal participants or those with semantic deficits (e.g., Lindquist et al., 2014).
Similarity Judgments Between Pairs of facial configurations **Perceptual Matching: Indicating whether or not two photos of facial configurations belong to the same emotion category**	
It is inefficient and time consuming to judge the similarity of all pairs of facial configurations. For a set of 100 faces, this requires (100*100)/2 = 5,000 different similarity judgments.	Participants can arrange face stimuli on a computer screen and all pairwise similarity judgments can be computed (the SPAM method proposed by Goldstone, 1994; e.g., see Hout et al., 2013). This procedure also solves the problem that the same pair of stimuli will have a different judged similarity depending on which item is presented first if face pairs are presented sequentially presented faces (the judged similarity of two objects, A and B, can depend on the order in which they are presented; the similarity of A vs B is not always judged to be the same as B vs A; Tversky, 1977). Other advantages are that categories can be discovered, rather than prescribed, and verbal associations are minimized. Analyses of similarity judgments

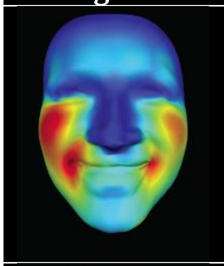
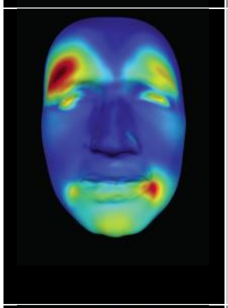
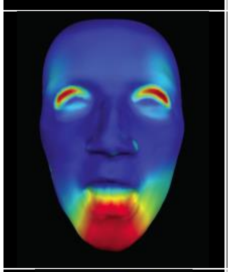
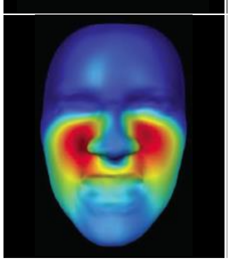
	typically yield more continuous similarity relations between emotion categories along affective dimensions (see Russell & Barrett, 1999).
Free-Labeling: photos of facial configurations are labeled with words offered by participants (unconstrained by experimenter)	
Forcing people to translate faces into words is not a good match, since much of the information from faces cannot be easily captured in words (Ekman, 1994).	This is not a special criticism of free labeling studies -- it applies to all studies that ask people to label a face with words, including the choice-from-array tasks.
Facial expressions did not evolve to represent specific verbal labels (Ekman, 1994, p. 270).	"Regardless of the language, of whether the culture is Western or Eastern, industrialized or preliterate, these facial expressions are labeled with the same emotion terms: happiness, sadness, anger, fear, disgust, and surprise" (Ekman, 1972, p. 278).
There is no widely accepted method for categorizing freely provided responses. (Ekman, 1994, p. 274).	Most scientists group together similar words (synonyms), so that a variety of words can be used to show evidence of a correct response (e.g., a frowning face, which is the proposed expression for sadness, could be labeled as "sad," "grieving," "disappointed," "blue," "despairing," and so on. Scientists routinely use databases that indicate synonyms, like WORDNET (used in Srinivasan & Martinez (2018). Also, it is possible to do data-driven groupings of emotion words into semantic categories (e.g., Jack et al., 2016; Shaver et al., 1987). The more serious problem is that early studies using free-labeling (e.g., Boucher & Carlson, 1980; Izard, 1971) did not provide enough information in the method sections about how freely provided labels were grouped.
Using freely chosen labels in a study of different cultures is difficult because it may be hard to find adequate translations (Ekman, 1994, p. 274). A given emotion word, like sadness, can correspond to different emotion concepts (with different	This is not a special criticism of free labeling studies – it holds for any experiment that uses emotion words requiring translation, including choice-from-array tasks. A standard solution to this problem is to use both forward and backward

features) in different languages (e.g., Wierzbicka, 1986, 2014). A single emotion word in one language can refer to more than one concept in another language (e.g., Pavlenko, 2014). Some languages have no one-to-one translation for English emotion words and some emotion concepts in other languages are not directly translatable into English emotion words (see Barrett, 2017; Russell, 1991; Jack et al., 2016).	translation (e.g., a word spoken in Hadzane is translated into English and then back translated into Hadzane; if there is no broken telephone, then the translation has fidelity). An even better method is to elicit features for the emotion words in question, including typicality of those features, to determine the fidelity of translation (e.g., de Mendoza et al., 2010) Scientifically, issues with translation are manageable if scientists allow phrases to stand in for specific words.
Using only single words will always fail to capture much of the rich information in faces.	Participants often provide multiple words or even longer descriptions of situations, behaviors, or behaviors in situations (e.g., see Gendron et al., 2014; Russell, 1994). Such data are time consuming to code and analyze.
Even when participants are told that photographs are of people trying to express an emotion, they often offer non-emotion labels. For example, Izard (1971) found that people offered labels such as deliberating, clowning, skepticism, pain, and so on (as reported in Russell, 1994).	This is not necessarily evidence that participants did not understand the task asked of them. It might be evidence that these facial configurations are not specific for expressing emotions.

Note. Response tasks are arrayed in order from those that constrain participants' responses most, making it difficult to observe evidence that can disconfirm commonsense beliefs about emotion to those that are least constrained, making it easier to observe variation and disconfirm commonsense beliefs. Choice-from-array = participants are shown a facial configuration and asked to infer its emotional meaning by choosing an emotion word from a small set of words; or, participants presented with an emotion word that labels an emotion category (e.g., sadness) or a brief story about a typical instance of an emotion category (e.g. "the boy's much loved dog just died and he is sad") along with two or three photographs of faces (typically posed into one of the configurations presented in Figure 4) and then asked to choose the facial configuration that they judge best matches the emotional episode described in the word or vignette. Typically, each emotion category is represented by a single scenario. Free sorting = Participants are given photographs of facial configurations and asked to sort them into emotion categories by piles on a big table or on a computer screen. Pairwise similarity judgments = participants rate the similarity of all possible pairs of face stimuli (e.g., on a scale of 0-6). For detailed design concerns about choice-from-array tasks, see Russell (1994, 1995).

.

Table 6: Culturally common facial configurations discovered using the reverse correlation method

Facial Configuration	AU Description	Associated Emotion Words – U.K.	Associated Emotion Words - China
	6+12+13+14	delighted, joy, happy, cheerful, contempt, pride	joyful, delighted, happy, glad, feel well, pleasantly surprised, embarrassed, pride
	4+20+24+43	fear, scared, anxious, upset, miserable, sad, depressed, shame, embarrassed	afraid, anxious, distressed, broken-hearted, sorrow and sadness, having a hard time, grief, dismay, anguish, worry, vexed, unhappy, shame, despise
	2+5+26+27	ecstatic, excited, surprised, frightened, terrified	amazed, greatly surprised, alarmed and panicky, scared, fear
	7+9+16+22	hate, disgust, fury, rage, anger	disgusted, bristle with anger, furious, wild wrath, storm of fury, storm of anger, indignant, rage

Note. Facial configurations extracted using reverse correlation from 62 models of facial configurations. Red coloring indicates stronger AU presence and blue indicates weakest AU presence. Some words and phrases that refer to emotion categories in Chinese are not considered emotion categories in English. Modified from Jack et al. (2016) and reproduced with permission.

Table 7: Summary of cross-cultural emotion perception in small-scale societies

	Unsupported			Weak Support			Moderate Support			Strong Support		
	Culture	N	Citation	Culture	N	Citation	Culture	N	Citation	Culture	N	Citation
Free-labeling	Fore, PNG[a]	100	Sorenson (1975), Sample 2 [c]							Sadong, Borneo [a1]	15	Sorenson (1975), Sample 4 [b]
	Bahinemo, PNG	71	Sorenson (1975) , Sample 3									
	Hadza, Tanzania	43	Gendron et al. (2018), Study 1									
	Trobrianders, PNG	32[f]	Crivelli et al. (2017), Study 1									
Cue-to-cue matching				Shuar, Ecuador	23	Bryant & Barrett (2008), Study 2						
Choice-from array: Matching face and words	Fore, PNG[a]	32	Ekman et al. (1969)[c]									
	Mwani, Mozambique	36[ef]	Crivelli, Jarillo et al. (2016), Study 2							Sadong, Borneo [a1]	15	Ekman et al. (1969) [b]
	Trobrianders, PNG	24[f]	Crivelli et al. (2017), Study 2	Dioula, Burkina Faso[a]	39	Tracy & Robbins (2008), Study 2						
	Trobrianders, PNG	68 [ef]	Crivelli, Jarillo et al. (2016), Study 1									
	Trobrianders, PNG	36 [f]	Crivelli, Russell et al. (2016), Study 1a									
Choice-from array: Matching face and scenario	Hadza, Tanzania	54	Gendron et al. (2018), Study 2				Dani, New Guinea[a]	34	Described in Ekman (1972) [g]			
										Fore, PNG [a1]	189, 130 [e]	Ekman & Friesen (1971) [d]
							Fore, PNG [a1]	189, 130 [e]	Sorenson (1975), Sample 1 [d]			

Notes. Findings summarized for anger, disgust, fear, sadness and surprise; happiness is the only pleasant category tested in all studies but Tracy & Robins (2008) and therefore perception can be (and likely is) guided by distinguishing valence in those studies. All studies used photographs of posed facial configurations that are similar to those in Figure 4, except Crivelli, Jarillo et al. (2016), Study 2 and Crivelli et al. (2017), Study 1. The Bryant & HC Barrett (2008) study was designed to examine emotion perception from vocalizations but is included because perceivers matched them to faces; in addition, participants were tested in a second language (Spanish) in which they received training. All choice-from-array studies did not carefully control whether foils and target facial configurations could be distinguished by valence and/or arousal except Gendron et al. 2018, Study 2. N = sample size. All participants were adults unless otherwise specified as adol=adolescents, ch=children. PNG = Papua New Guinea. Unsupported = reliability and specificity at chance, or any level of reliability above chance combined with evidence of no specificity. Weak support = reliability between 20% and 40% (weak) for at least a single emotion category other than happiness combined above chance specificity for that category or reliability between 41% and 70% (moderate) for at least a single category other than happiness with unknown specificity. Moderate support = reliability between 41% and 70% (moderate) combined with any evidence of above chance specificity those categories or reliability above 70% (strong) for at least a single category other than happiness with unknown specificity. Strong support = strong evidence of reliability (above 70%) and strong evidence of specificity for at least a single emotion category other than happiness. Superscript a: Specificity levels were not reported. Superscript a1: Specificity inferred from reported results. Superscript b: The sample size, marginal means and exact pattern of errors reported for the Sadong samples is identical in Sorenson (1975), Sample 3 and Ekman et al. (1969); Sorenson described using a free-labeling method and Ekman et al. (1969) described using a choice-from-array method in which participants were shown photographs and asked to choose a label from a small list of emotion words; Ekman (1994) indicated, however, that he did not use a free-labeling method, implying that the samples are distinct. Superscript c: Sorenson (1975), Sample 2 included three groups of Fore participants (those with little, moderate and most other group contact). The pattern of findings is nearly identical for the subgroup with the most contact and the data reported for the Fore in Ekman et al. (1969); again, Sorenson described using a free-labeling method and Ekman et al. (1969) described using a choice-from-array method. It is questionable whether the Sadong and the Fore subgroup should be considered isolated (see Sorenson, 1975, p. 362 and 363), but we include them here to avoid falsely dichotomizing cultures as “isolated from” versus “exposed to” one another (Fridlund, 1994; Gewald, 2010). Superscript d: these are likely the same sample because the sample sizes and pattern of data are identical for all emotion categories except for the fear category, which is extremely similar, and for the disgust category which includes responses for contempt in Ekman & Friesen (1971) but was kept separate in Sorenson (1975). Superscript e: participants were children. Superscript f: participants were adolescents. Superscript g:The Dani sample reported in Ekman (1972) is likely a subset of the data from Ekman, Heider, Friesen, and Heider (unpublished manuscript).

Table 8: Recommendations for reading scientific studies about emotion

1. Take note of whether an experiment is studying expressive stereotypes or more variable facial movements.
2. Take note of data on specificity and generalizability; do not focus solely on reliability at above chance levels.
3. Make a distinction between the data in an experiment (what was measured) and how those data are interpreted.
4. Translate “emotional expressions” or “emotional displays” into “facial movements.”
5. Translate “emotion recognition” into “emotion perception” or “emotion inference.”
6. Translate “accuracy” to “agreement,” “consensus” or “reliability.”
7. Give more weight to studies that measure facial movements or study the perception of facial movements made in more naturalistic settings.
8. Take note of studies that measure or manipulate context.
9. Field studies of people from small-scale, remote cultures are often less well-controlled than studies conducted in the laboratory, but they are invaluable in the information that they provide and should be valued.
10. Remember that emotions are not understood as internal states in all cultures. In some cultures they are understood as situated actions.
11. Do not skip the method and results sections and skip to the discussion to learn the results of an experiment. It is important to know what was measured and observed, not just how scientists interpret their measurements.

Table 9: Recommendations for future research

General Recommendations
▪ Take chances on studies that attempt to go beyond merely supporting traditional views of emotion.
▪ Support papers that attempt to study facial movements in real life, measuring context, sampling across cultures even though these studies are often less well controlled than studies in the laboratory, or may use facial stimuli that are less familiar to reviewers than canonical stimulus sets.
▪ Prioritize multidisciplinary studies that combine classical psychology methods with cognitive neuroscience, machine learning, etc.
▪ Support larger scale studies that bridge the lab and the world, that study individual people across many contexts, and measure emotional episodes in high dimensional detail, including physical, psychological and social features; encourage multiple investigators with different areas of expertise to work together.
▪ Support the development of computational approaches.
▪ Create R&D teams that pair psychologists and cognitive scientists trained in the psychology of emotion with engineers and computer scientists.
▪ Increase opportunities to test innovative methods and novel hypotheses, with the acknowledgement that such approaches are likely to elicit resistance from established scientists in the field of emotion.
▪ Generate more studies to identify the underlying neural mechanisms of the production and perception of facial movements.
▪ Direct funding to thornier but necessary new questions and be critical of projects that perpetuate past errors in emotion research.
▪ Direct healthy skepticism to tests, measures, and interventions that rely upon assumptions about "reading facial expressions of emotion" that seem to ignore published evidence and/or ignore integration of contextual information along with facial cues.
▪ Develop systematic, precise ways to describe and/or manipulate the dynamics of specific facial actions.
Stimulus Selection Recommendations

Limitations in stimulus selection can bias results.	▪ For perception studies, incorporate images from the wild (e.g., from multiple internet sources) to capture the full range of facial movements that humans produce in their everyday lives. ▪ For both production studies (where stimuli are designated to evoke emotion) and perception studies, build variation into stimulus sets so conclusions about emotion categories are not inferred (or evoked) from limited stimuli. Consider randomly sampling a variety of stimuli for a given category and treating stimuli as a random variable. ▪ For production studies, ensure that multiple stimuli per emotion category are used to evoke an emotion.
Little is known about the dynamics of the production and perception of emotion signaling.	▪ For perception studies, use dynamic images rather than rely on still images. For production studies, code the temporal dynamics of facial movements. ▪ Attempt to determine full dynamics and the apex of an emotion signal, changes to AUs as signals emerge and recede, and whether the kinematics of distinct AUs are similar or different across sequences or phases of emotion signaling. ▪ Ensure sufficient temporal resolution to allow for event segmentation to be assessed in perception studies.
The role of context is hotly debated, but rarely measured.	▪ Manipulate (or at least measure) the context in which target stimuli are perceived to evaluate whether data are truly stimulus-specific or influenced by context features. ▪ Describe in a systematic way the differences in context, whether for production or perception studies. Theories about the effects of context cannot be resolved until we address how to measure and quantify context.
Sample Selection Recommendations	
Cross-cultural studies can provide powerful insights, but are limited in number and scope.	▪ Quantify, as best as possible, participants' degree of exposure to the west, as well as the amount and type of formal schooling made available to participants. ▪ Harness technology to collect larger numbers of images and video sequences of facial movements across cultures. Use unlabeled classification approaches to discover emotion categories and their expressive forms, rather than continuing to ask whether other cultures are similar to the US. Remember that emotions and mental inferences may be understood differently in different cultures.
Task and Method Design Recommendations	

Measurement versus interpretation of emotion is often blurred in research studies.	▪ Contrast more than one "emotion" category with a baseline, so that conclusions about a specific emotion category are not drawn from a comparison of an emotion versus a no emotion condition. ▪ Compare multiple emotion categories to non-emotion categories in a given study.
New insights about emotion are constrained by reliance on, and assumptions about, traditional categories.	▪ Measure emotional episodes in a multimodal way and attempt to discover explicit criteria for when an emotion is present or absent. Such discovery may require within-person approaches. ▪ Sample broader categories of possible emotion states than the limited categories used in prior research (move beyond categories such happiness, anger, sadness, fear, etc.). Test for variations in intensity within these categories and similarity across categories. ▪ Unless a study design is completely data-driven, explicitly state the theoretical priors of the research team. The distinction is between whether you are seeking to discover versus verify emotion categories. Both approaches are valid, but should be clearly articulated.
Data Analysis Recommendations	
Findings are limited by a failure to consider issues related to forward and reverse inference.	▪ Address issues of reliability and specificity when presenting data on emotion expression and emotion perception. ▪ Use formal signal detection analytics and information theoretic measures rather relying on frequency or levels of agreement. Consider using Bayesian methods so that the null hypothesis can be tested directly.

Figure Captions

Figure 1. Explanatory frameworks guiding the science of emotion: The nature of emotion categories and their concepts. Figure is plotted along two dimensions. Horizontal: represents hypotheses about the surface similarities shared by instances of the same emotion category (e.g., the facial movements that express instances of the same emotion category). Vertical: represents hypotheses about the deep similarities in the mechanisms that cause instances of the same emotion category (e.g., to what extent do instances in the same category share deep, causal features?). Colors represent the type of emotion categories that are proposed in each theoretical framework (green = ad hoc, abstract categories; yellow = prototype or theory-based categories; red = natural kind categories).

Figure 2. Example figures from recently published papers that reinforce the common belief of a one-to-one mapping between a single emotion category and a single facial configuration. **A**. Keltner et al., in press, Table 2, formatting modified. **B**. Shariff & Tracy, 2011, Figure 2. Permissions pending.

Figure 3.Evaluation criteria: Reliability and specificity in relation to forward and reverse inference. Anger and fear are used as the example categories.

Figure 4. Facial action ensembles for commonsense facial configurations. Facial action coding system (FACS) codes that correspond to the commonsense expressive configuration in adults. **A** is proposed expression for anger and corresponds to prescribed EMFACS code for anger (AUs 4, 5, 7, and 23). **B** is proposed expression for disgust and corresponds to prescribed EMFACS code for disgust (AU 10). **C** is proposed expression for fear and corresponds to prescribed EMFACS code for fear (AUs 1, 2, and 5 or 5 and 20). **D** is proposed expression for happiness and corresponds to prescribed EMFACS code for the so-called Duchenne smile (AUs 6 and 12). **E** is proposed expression for sadness and corresponds to prescribed EMFACS code for sadness (AUs 1, 4, 11 and 15 or 1, 4, 15 and 17). **F** is proposed expression for surprise and corresponds to prescribed EMFACS code for surprise (AUs 1, 2, 5, and 26). It was originally proposed that infants express emotions with the same facial configurations as adults. Later research revealed morphological differences between the proposed expressive configurations for adults and infants. Only three out of a possible nineteen proposed configurations for negative emotions from the infant coding scheme were the same as the configurations proposed for adults (Oster et al., 1992). **G**, adapted from Keltner et al., in press, Table 2. **H**. adapted from Shariff & Tracy, 2011, Figure 2. Permissions pending.

Figure 5. Meta-analysis of facial movements during emotional episodes: A summary of effect sizes across studies (Duran et al., 2017). Effect sizes are computed as correlations or proportions (as reported in the original experiments). Results include experiments that reported a correspondence between a facial configuration and its hypothesized emotion category and those that reported a correspondence between individual AUs of that facial configuration and the relevant emotion category; meta-analytic summaries for entire ensembles of AUs only (the facial configurations specified in Figure 2) were even lower than those that appear here.

Figure 6: Comparing posed and spontaneous facial movements. Results from Table 6, Cordaro et al. (2017), degree of overlap between the hypothesized configuration of facial movements for

each emotion category and the “International Core Patterns” derived from participants’ expressive poses; Gabonese participants in Elfenbein et al. (2007), reliability for the anger category is for AU4 + AU5 only; proportion data only from Duran et al., (2017).

Figure 7. Examples of virtual humans. Virtual humans are software-based artifacts that look like and act like people. (A) Feng et al, 2017; (B) Zoll et al., 2006; (C) Hoyt et al., 2003; (D) Marsella et al. 2000.

Figure 8 Emotion perception findings. (**A**) Average effect sizes for perceptions of facial configurations from Elfenbein & Ambady (2002), in which 95% of the articles summarized used choice-from-array to measure participants’ emotion inferences. (**B**) Free-labeling of facial configurations across five language groups from Srinivasan & Martinez (2018). IDs chosen represent the best match to the commonsense facial configurations in Figure 4 based on AUs present. No configuration discovered in this study exactly match the AU configurations proposed by Darwin or documented in prior research. Proportion of times participants offered emotion category labels (or their synonyms) are reported. According to standard scientific criteria, universal expressions of emotion should elicit agreement rates that are considerably higher than those reported here, generally in the 70 ± 90% range, even when methodological constraints are relaxed (Haidt & Keltner, 1999). Specificity data were not available for the Elfenbein & Ambady (2002) meta-analysis.

Figure 9. Map of cross-cultural studies of emotion perception in small-scale societies. People in small scale societies typically live in groupings of several hundred to several thousand that maintain autonomy in social, political and economic spheres. **(A).** Epoch 1 studies, published between 1969 and 1975, were geographically constrained to societies in the South Pacific. (**B).** Epoch 2 studies, published between 2008 and 2017, sample from a broader geographic range including Africa and South America, and are more diverse in the ecological and social contexts of the societies tested. This type of diversity is a necessary condition for discovering the extent of cultural variation in psychological phenomena (Medin et al., 2017). Reproduced with permission from Gendron et al. (2018).

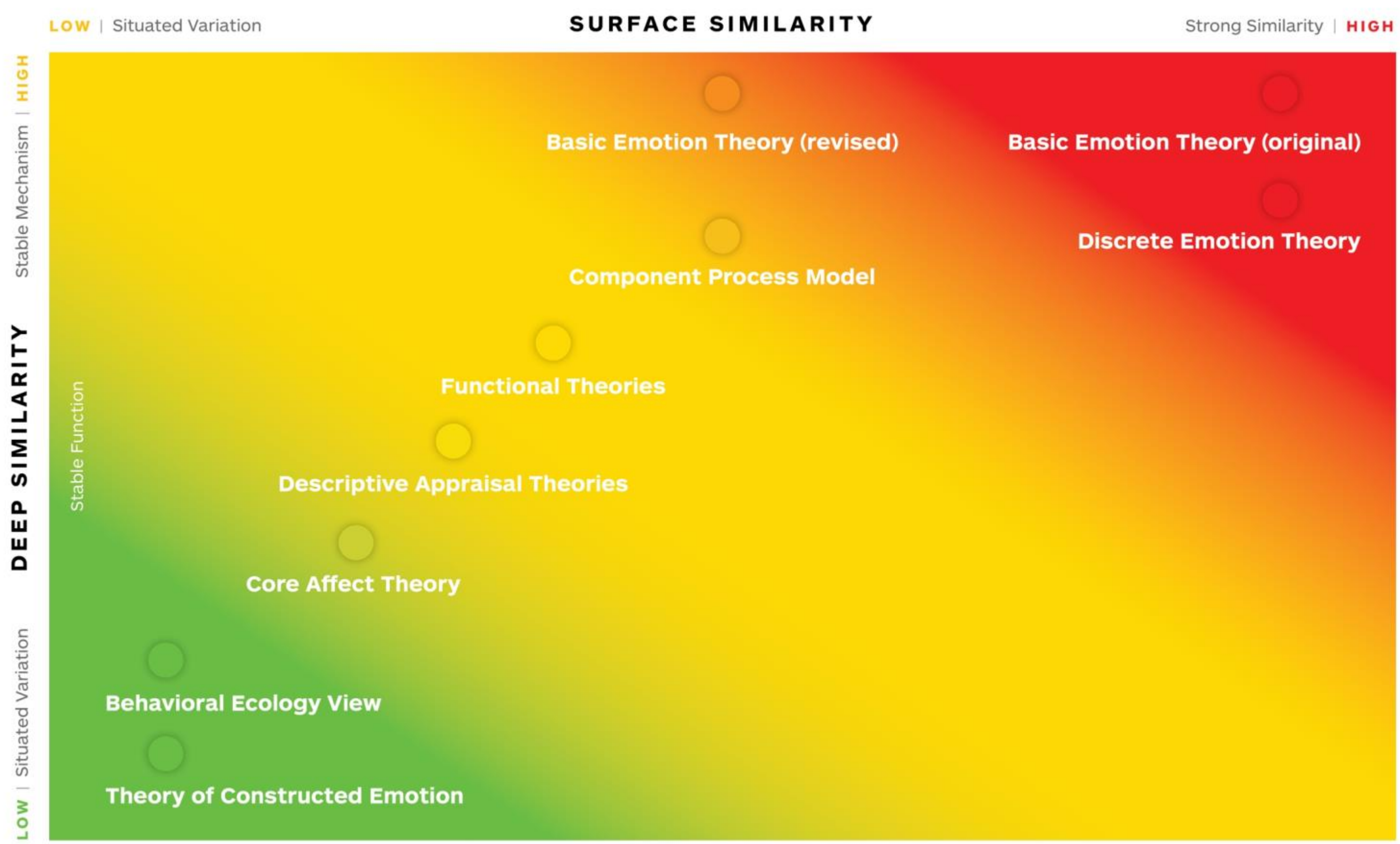
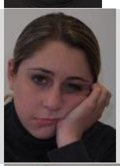


Figure 1. Explanatory frameworks guiding the science of emotion: The nature of emotion categories and their concepts.

A

Emotion	Example Photo	Action Units	Emotion	Example Photo	Action Units	Emotion	Example Photo	Action Units
Amusement		6+7+12+25+26+53	Desire		19+25+26+43	Pain		4+6+7+9+17+18+23+24
Anger		4+5+17+23+24	Disgust		7+9+19+25+26	Pride		53+64
Boredom		43+55	Embarrassment		7+12+15+52+54+64	Sadness		1+4+6+15+17
Confusion		4+7+56	Fear		1+2+4+5+7+20+25	Shame		54+64
Contentment		12+43	Happiness		6+7+12+25+26	Surprise		1+2+5+25+26
Coyness		6+7+12+25+26+52+54+61	Interest		1+2+12	Sympathy		1+17+24+57

B

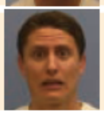
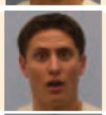
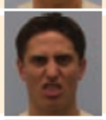
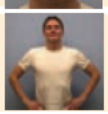
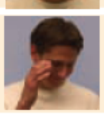
EMOTION EXPRESSION	HYPOTHESIZED PHYSIOLOGICAL FUNCTION	HYPOTHESIZED COMMUNICATIVE FUNCTION	RELEVANT RESEARCH
Happiness 	Research Needed	Communicates a Lack of Threat	Preuschoft & Van Hoof, 1997 Ramachandran, 1998
Sadness 	Research Needed	Tears Handicap Vision to Signal Appeasement and Elicit Sympathy	Hasson, 2009
Anger 	Research Needed	Alerts of Impending Threat, Communicates Dominance	Marsh, Ambady, & Kleck, 2005 Wilkowski & Meier, 2010
Fear 	Widened Eyes Increase Visual Field and Speed Up Eye Movements	Alerts of Possible Threat, and Appeases Potential Aggressors	Marsh et al., 2005 Ohman & Mineka, 2001 Susskind et al., 2008
Surprise 	Widened Eyes Increase Visual Field to See Unexpected Stimulus	Research Needed	Ekman, 1989
Disgust 	Constricted Orifices Reduce Inhalation of Possible Contaminants	Warns About Aversive Foods, as Well as Distasteful Ideas and Behaviors	Rozin et al. 1994, Chapman, Kim, Susskind, & Anderson, 2009
Pride 	Boosts Testosterone and Increases Lung Capacity to Prepare for Agonistic Encounters	Communicates Heightened Social Status	Carney, Cuddy, & Yap, 2010 Shariff & Tracy, 2009 Tracy & Matsumoto, 2008
Shame 	Reduces/Hides Bodily Targets From Potential Attack	Communicates Lessened Social Status, Desire to Appease	Keltner & Harker, 1998 Shariff & Tracy, 2009 Tracy & Matsumoto, 2008
Embarrassment 	Reduces/Hides Bodily Targets From Potential Attack	Communicates Lessened Social Status, Desire to Appease	Keltner & Buswell, 1997

Figure 2. Example figures that reinforce the common belief of a one-to-one mapping between a single emotion category and a single facial configuration.

A. Keltner et al., in press, Table 2, formatting modified. **B**. Shariff & Tracy, 2011, Figure 2. Permissions pending.

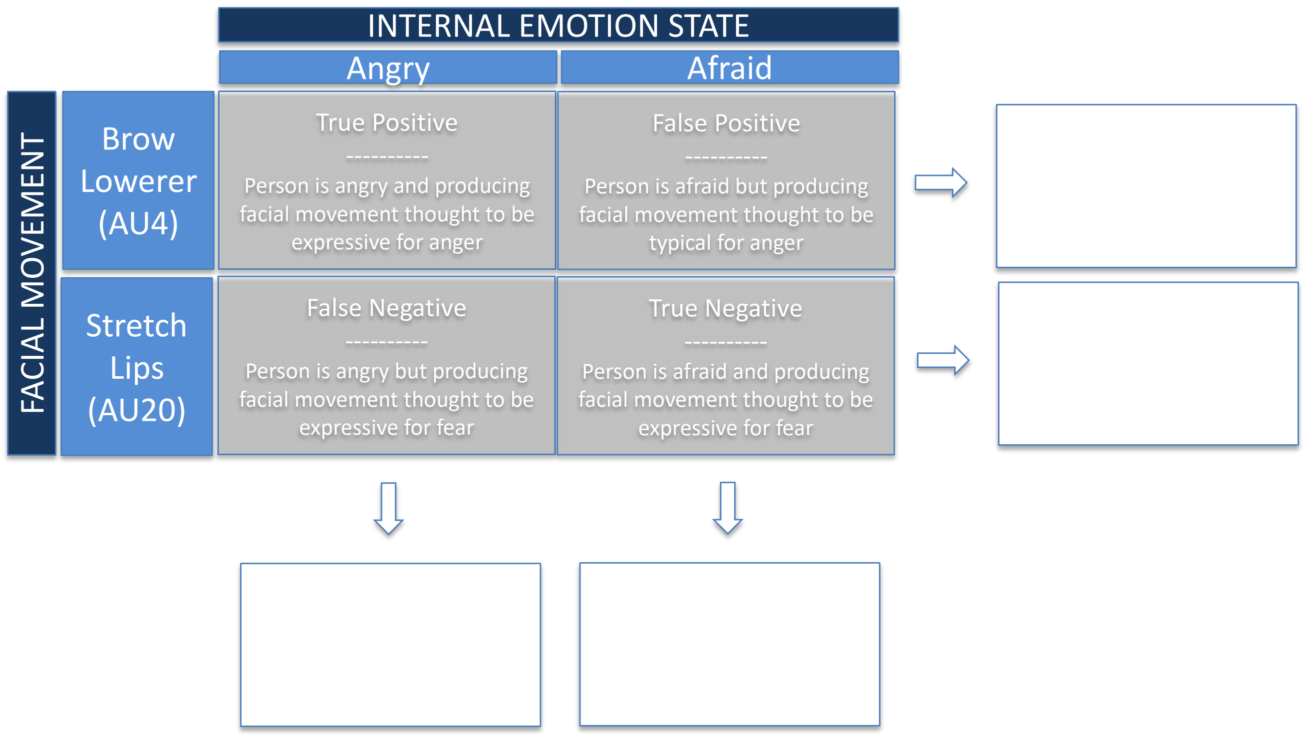


Figure 3. Reliability and specificity in relation to forward and reverse inference.

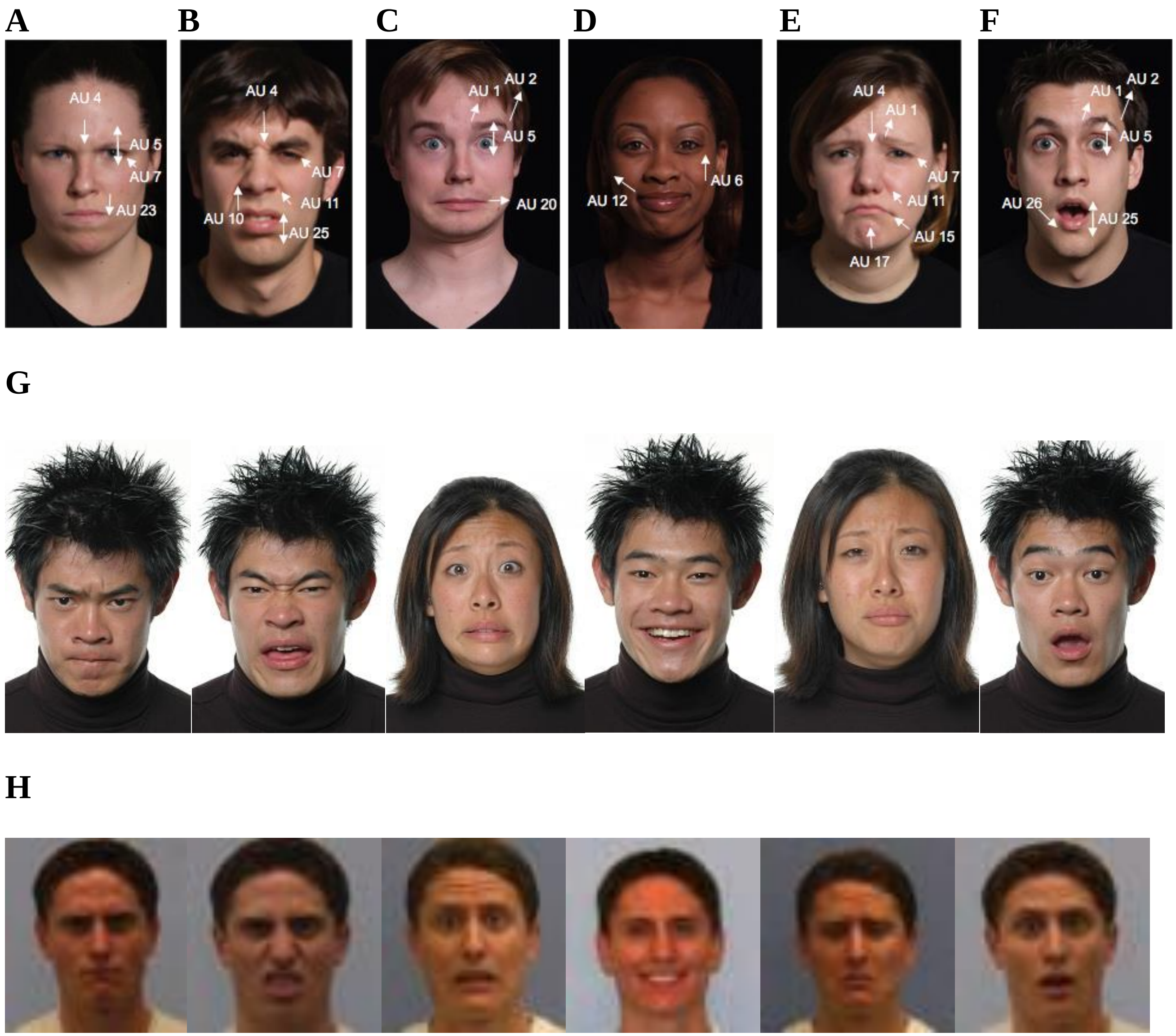

A
AU 4
AU 5
AU 7
AU 23
B
AU 4
AU 7
AU 11
AU 10
AU 25
C
AU 1
AU 2
AU 5
AU 20
D
AU 6
AU 12
E
AU 4
AU 1
AU 7
AU 11
AU 15
AU 17
F
AU 1
AU 2
AU 5
AU 26
AU 25
G
H

Figure 4. Facial action ensembles for commonsense facial configurations. Facial action coding system (FACS) codes are shown for six expressive configurations in adults. **A** is proposed expression for anger and corresponds to prescribed EMFACS code for anger (AUs 4, 5, 7, and 23). **B** is proposed expression for disgust and corresponds to prescribed EMFACS code for disgust (AU 10). **C** is proposed expression for fear and corresponds to prescribed EMFACS code for fear (AUs 1, 2, and 5 or 5 and 20). **D** is proposed expression for happiness and corresponds to prescribed EMFACS code for the so-called Duchenne smile (AUs 6 and 12). **E** is proposed expression for sadness and corresponds to prescribed EMFACS code for sadness (AUs 1, 4, 11 and 15 or 1, 4, 15 and 17). **F** is proposed expression for surprise and corresponds to prescribed EMFACS code for surprise (AUs 1, 2, 5, and 26). It was originally proposed that infants express emotions with the same facial configurations as adults. Later research revealed morphological differences between the proposed expressive configurations for adults and infants. Only three out of a possible nineteen proposed configurations for negative emotions from the infant coding scheme were the same as the configurations proposed for adults (Oster et al., 1992). **G**, adapted from Keltner et al., in press, Table 2. **H**. adapted from Shariff & Tracy, 2011, Figure 2. Permissions pending.

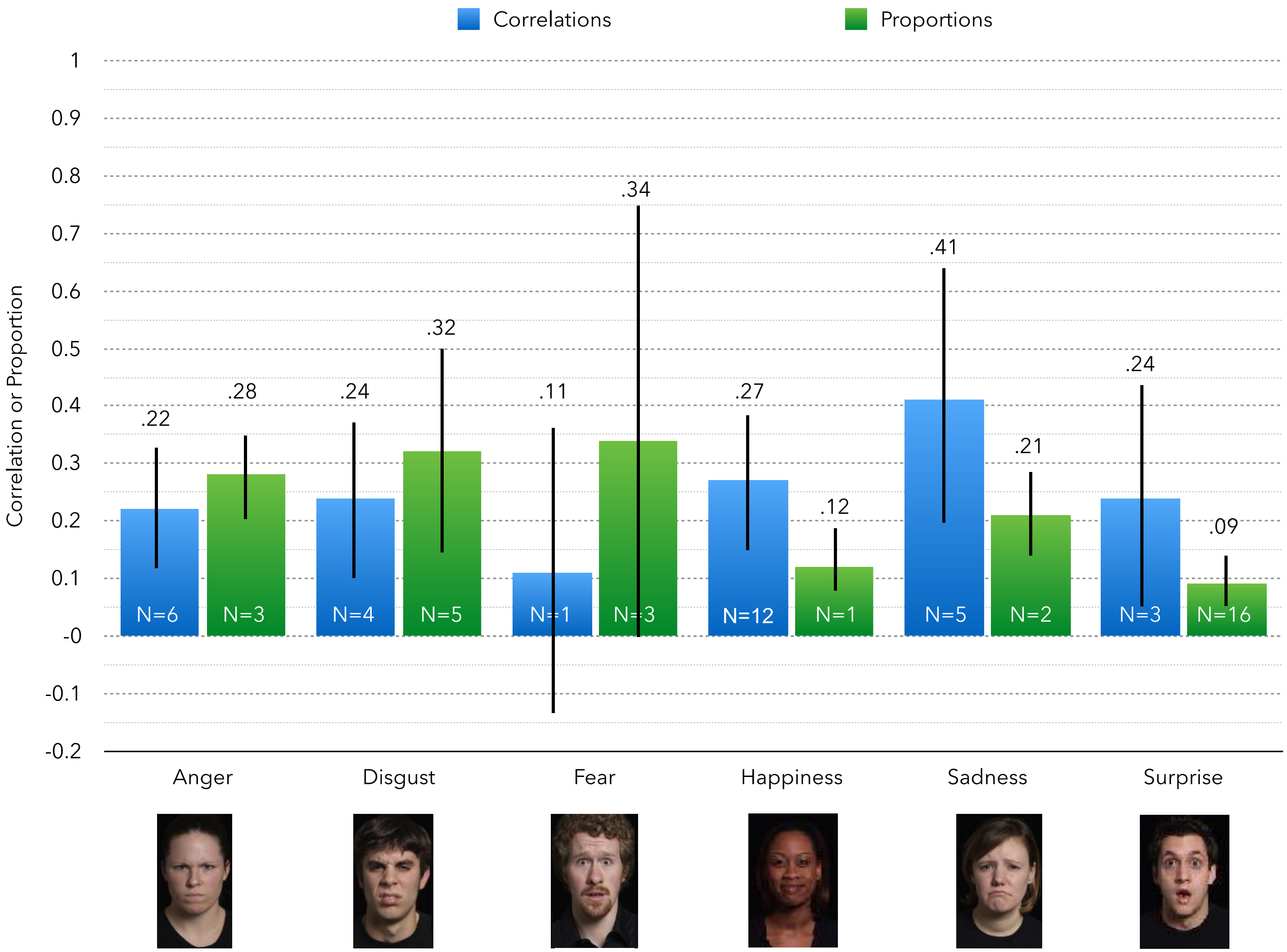

Correlations
Proportions
Correlation or Proportion
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
-0
-0.1
-0.2
.22
.28
.24
.32
.11
.34
.27
.12
.41
.21
.24
.09
N=6
N=3
N=4
N=5
N=1
N=3
N=12
N=1
N=5
N=2
N=3
N=16
Anger
Disgust
Fear
Happiness
Sadness
Surprise

Figure 5. Meta-analysis of facial movements during emotional episodes.

Effect sizes are computed as correlations or proportions (as reported in the original experiments). Results shown here included experiments that reported a correspondence between a facial configuration and its hypothesized emotion category combined with those that reported a correspondence between a partial configuration (any individual AU comprising that facial configuration) and the relevant emotion category; meta-analytic summaries for entire ensembles of AUs *only* (the facial configurations specified in Figure 4) were even lower than those that appear here. Colored bars show correlations and proportions as averages across the studies (number of studies given by N), vertical black line indicates the 95% confidence interval.

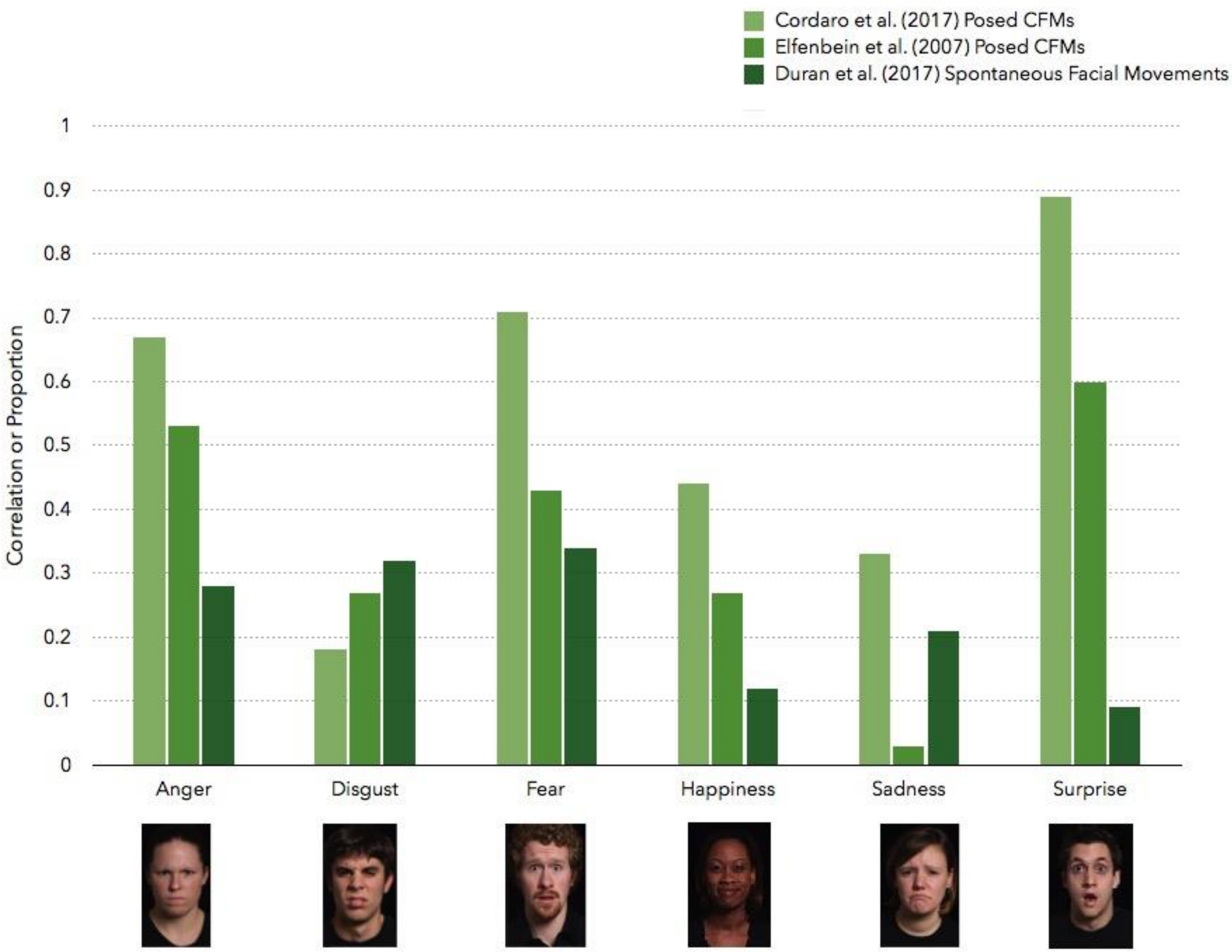

Cordaro et al. (2017) Posed CFMs
Elfenbein et al. (2007) Posed CFMs
Duran et al. (2017) Spontaneous Facial Movements
Correlation or Proportion
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Anger
Disgust
Fear
Happiness
Sadness
Surprise

Figure 6: Comparing posed and spontaneous facial movements.

Results from Table 6, Cordaro et al. (2017), percent of overlap between the hypothesized expressive form for each emotion category and the "International Core Patterns" derived from participants' expressive poses. For Elfenbein et al. (2007), proportion of Gabonese participants who posed the hypothesized universal facial configuration for each emotion category, reliability for the anger category is for AU4 + AU5 only. Proportion data from Duran et al., (2017) are also presented for comparison.

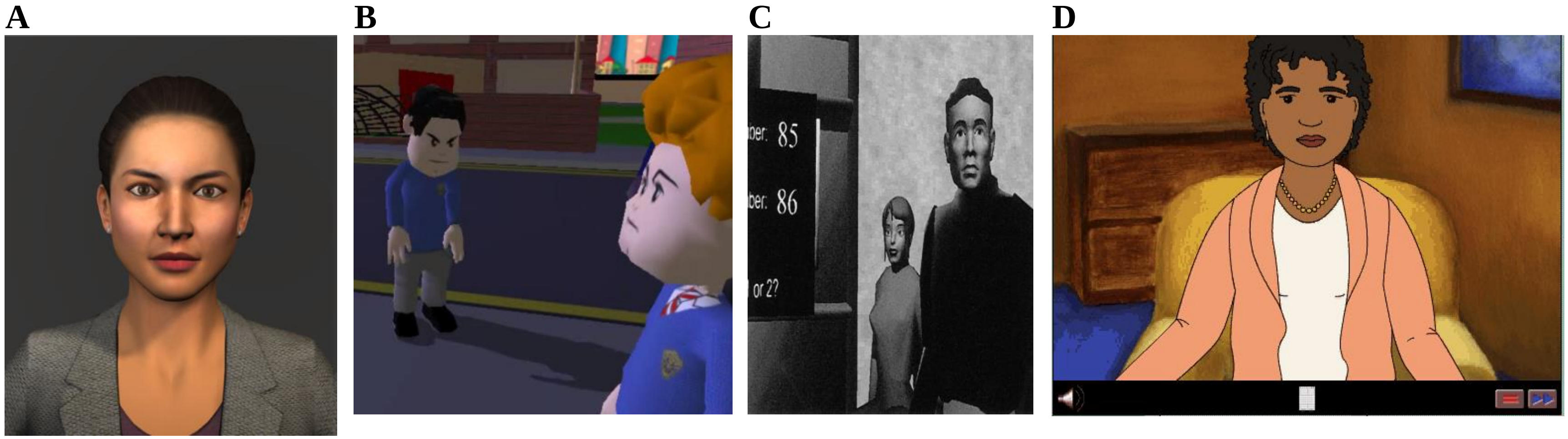


Figure 7. Virtual humans: Examples.(**A**) Feng et al, 2017; (**B**) Zoll et al., 2006; (**C**) Hoyt et al., 2003; (**D**) Marsella et al. 2000. Permissions pending.

A

B

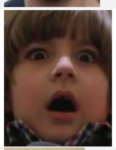
	% EMOTION LABELS OFFERED							
	Anger	Disgust	Fear	Happiness	Sadness	Surprise	Non-Affective	Action
	39.92		7.19		7.93		12.92	3.96
	11.38	24.94		2.63	10.05		12.39	
	4.14	4.01	34.75	4.12	4.13	18.61	2.73	
				48.84		3.17	1.84	10.01
	8.31		8.60		37.41		13.40	
	2.16	6.02	12.23	6.01		31.90		1.82

*Figure 8 Emotion perception findings.***A**. Average effect sizes for perceptions of facial configurations from Elfenbein & Ambady (2002), in which 95% of the articles summarized used choice-from-array to measure participants' emotion inferences. **B**. Free-labeling of facial configurations across five language groups from Srinivasan & Martinez (2018). IDs chosen represent the best match to the commonsense facial configurations in Figure 4 based on AUs present. No configuration discovered in this study exactly match the AU configurations proposed by Darwin or documented in prior research. Proportion of times participants offered emotion category labels (or their synonyms) are reported. According to standard scientific criteria, universal expressions of emotion should elicit agreement rates that are considerably higher than those reported here, generally in the 70 ± 90% range, even when methodological constraints are relaxed (Haidt & Keltner, 1999). Specificity data were not available for the Elfenbein & Ambady (2002) meta-analysis.

A


Bahinemo of
Papua New
Guinea
Sorenson (1975)
Fore of Papua
New Guinea
Ekman,
Sorenson &
Friesen (1969)*
Sadong of Borneo*
Ekman et al. (1969)
Sadong of Borneo
Sorenson (1975)
Fore of Papua
New Guinea
Ekman & Freisen
(1971)
Dani of Indonesia
(Western New Guinea)
Ekman (1972)
Fore of Papua
New Guinea (2
samples)
Sorenson (1975)
Dani of Indonesia
(Western New Guinea)
Ekman, Heider et al.
(unpublished)

B

Figure 9. Map of cross-cultural studies of emotion perception in small-scale societies.
People in small scale societies typically live in groupings of several hundred to several thousand that maintain autonomy in social, political and economic spheres. **A.** Epoch 1 studies, published between 1969 and 1975, were geographically constrained to societies in the Pacific area. **B.**

Epoch 2 studies, published between 2008 and 2017, sample from a broader geographic range including Africa and South America, and are more diverse in the ecological and social contexts of the societies tested. This type of diversity is a necessary condition for discovering the extent of cultural variation in psychological phenomena (Medin et al., 2017). Reproduced with permission from Gendron et al. (2018)

Supplementary On-Line Materials

Box 1: Google Search Evidence Demonstrating the Existence of a Common View of Emotional Expressions

Box 2: Theoretical Frameworks and Their Relation to Emotion Categories

Box 3: The Current Study of Context Effects in Emotion Perception

Box 4: Origins of the Hypothesized Emotional Expressions For Anger, Disgust, Fear, Happiness, Sadness and Surprise

Box 5: The Face: Anatomy of Facial Muscles and Facial Electromyography

Box 6: A Summary of Computer Vision Algorithms for Automatically Detecting Facial Actions

Box 7: Variations in Facial Movements

Box 8: Meta-Analytic Evidence of Autonomic Nervous System Changes During Emotion

Box 9: Emotion Episodes and Their Affective Features

Box 10: Learning to Express and Perceive Emotion

Box 11: Research with Virtual Humans

Box 12: The In-Group Advantage in Emotion Perception

Box 13: Some Details of the Emotion Perception Studies By Crivelli and Colleagues

Box 14: The Power of Words in Emotion Perception Experiments

Box 15: The Habituation Task Used in Studies of Emotion Perception in Infants

Box 16: Information Theory as Applied to Emotion Communication

Box 1: Google Search Evidence Demonstrating the Existence of a Common View of Emotional Expressions

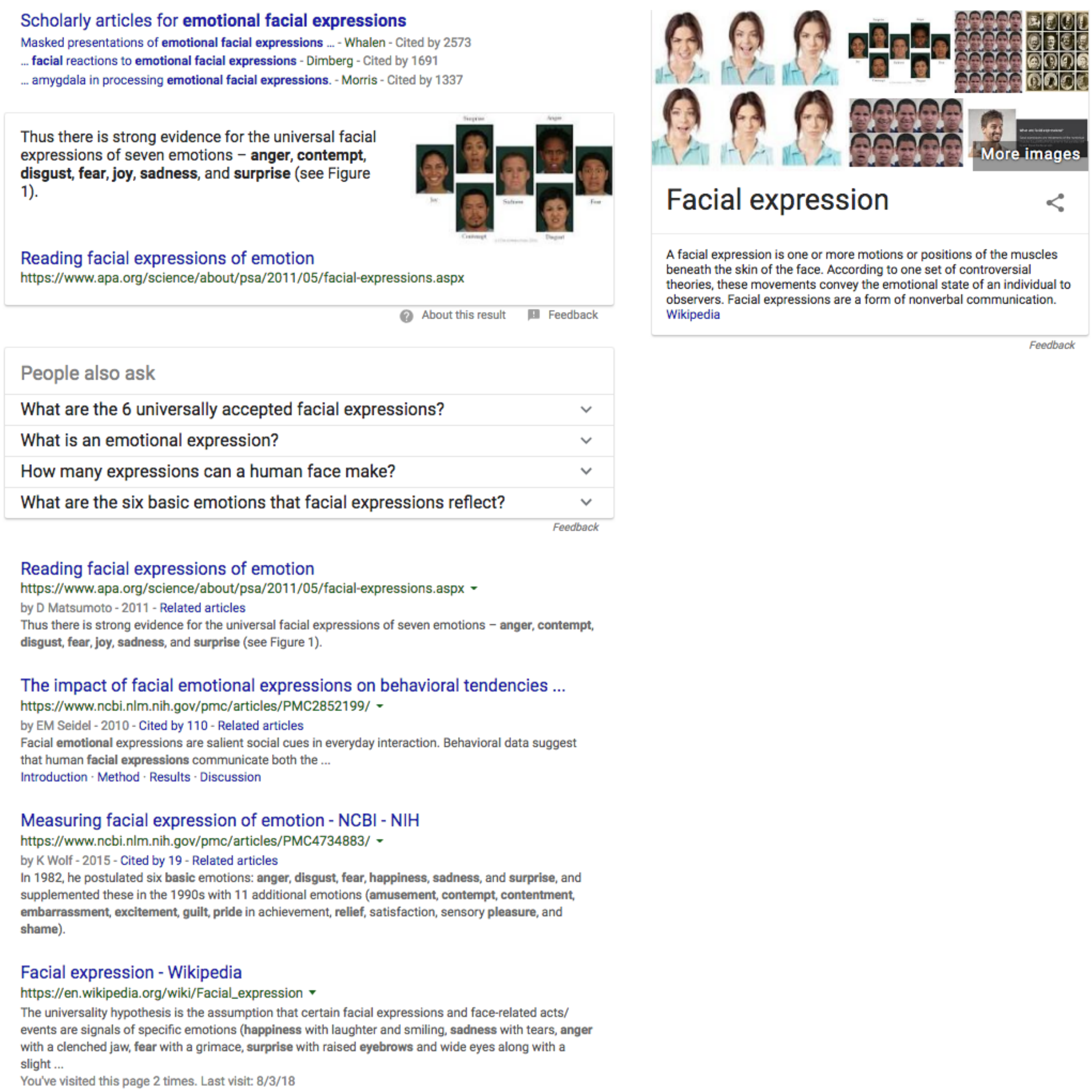
This box provides perhaps the simplest characterization of the "common view" of emotional expressions: how it is represented in the internet. It is important to note that we are not suggesting that the common view is necessarily the view that individual scientists personally hold, nor the view that all laypeople personally hold. Instead, it is the view that most laypeople think science supports, and the view that many scientists, intentionally or unintentionally, have perpetuated in the literature. It is the modal view that computer scientists, educators, and psychiatrists draw from, in the assumption that it is the view currently best supported by scientific evidence; and this is the assumption we evaluate in our paper.

The items retrieved from a Google search represent the consensus of our culture in many ways, and certainly represent what most people take as a summary of the conclusions best supported by current science, broadly conceived. Since different search histories and browsers may produce somewhat different results, we asked 23 students and post-doctoral fellows to search the term "emotional facial expressions," and send us a screenshot of their first page of results.

This search produced web pages very similar to the small reproduction below. In all pages that included images (21/23; like the one at the upper right in the example), the first and most prominent image shows the 6 "basic" facial expressions of emotion; in several this was the only image produced. In all searches that included the box "People also ask" (7/23), the first item listed was "What are the 6 universally accepted facial expressions?"

In 22/23 searches, the very first hit, titled "Reading facial expressions of emotion," brings up an APA article (https://www.apa.org/science/about/psa/2011/05/facial-expressions.aspx) that leads with a section on the universality of facial expressions; Figure 1 in that article is a picture of the facial configurations that are hypothesized to express anger, disgust, fear, happiness, sadness, surprise and contempt and the APA article concludes with "Because facial expressions of emotion are part of our evolutionary history and are a biologically innate ability, we all have the ability to read them." The second common hit (15/23; this and subsequent hits get truncated because not all searches had screens that permitted a long list) is to a website (https://thoughtcatalog.com/january-nelson/2018/06/list-of-emotions/) that lists the 6 basic emotions together with their facial expressions (the source of the most common image retrieved by the search). The next most common hit in the list (14/23, usually third or fourth entry) is a 2018 news from Science Daily titled, "How emotions in facial expressions are understood." It mentions by name only the following emotions: fear, happiness, surprise, anger, sadness. It covers a research paper that showed participants "images of faces expressing the six emotions and one neutral expression". The next most common hit is typically a research paper; in our sample it was https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5152920/ (10/23 hits) which is an article investigating emotion perception from faces, using as stimuli only photos of the six basic emotions. Another common research article here was https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2852199/, whose very first sentence in the abstract is, "Emotional faces communicate both the emotional state and behavioral intentions of an individual."

These searches are complemented by the Wikipedia entry for "facial expressions," as well as by later search hits that paint a more nuanced picture and nearly always explicitly note that there is considerable controversy about how reliably we can judge emotion from faces, what categories of emotional expression there are, and whether emotional facial expressions are culturally universal. Our conclusion from this google search is twofold. First, there is overwhelming evidence that the modal view that is available on the internet with a casual search represents the common view: that specific facial expressions correspond to specific emotions (usually about 6), and that these can be reliably perceived. Second, closer scrutiny notes that the common view has strong critiques. These two facts motivated our paper and the aim to produce the most thorough review of the evidence.



Box Figure 1-1. Screenshot of the first page of a google search for "emotional facial expressions."

Box 2: Theoretical Frameworks and Their Relation to Emotion Categories

An instance of emotion can be described by a variety of features: *physical features* (such as patterns of expressive facial movements, vocal acoustics, autonomic nervous system changes, and neural activity), *affective features* that capture what the instance feels like (e.g., how pleasant or unpleasant the episode feels, how arousing it feels; Barrett & Bliss-Moreau, 2009; Russell & Barrett, 1999), *appraisal features* that refer to how the situation is experienced (e.g., whether the situation is experienced as novel or familiar, as conducive to one's immediate goals or not, and so on; Barrett, Mesquita et al., 2007; Clore & Ortony, 2008, 2013; Scherer et al., 2017) and *functional features* that refer to the goals that a person is attempting to meet (e.g., to avoid a predator, to get closer to someone, to win a competition, etc.; e.g., Adolphs, 2017; Lazarus, 1993). An emotion category is a grouping of emotional episodes that share some feature or set of features in common.

Many debates about the nature of emotion boil down to disagreements about *the nature of the similarities* shared by instances of the same emotion category and the *degree of variation* in the relevant features, as well as potential similarity and differences in features across emotion categories. These debates can be summarized by two dimensions, as depicted in Figure 1: The horizontal dimension represents hypotheses about the surface similarities across instances of the same emotion category, such as being expressed by similar patterns of facial movements. And the vertical dimension represents hypotheses about the deep similarities across instances of the same emotion category, such as sharing the same neural circuitry.

Similarities and Variation in Surface Features

In this paper, we are primarily concerned with the horizontal dimension representing the similarity and variation in surface features, such as the facial movements that express instances of emotion. At one end of the "surface" continuum, basic emotion approaches (e.g., Ekman, 1972; Tracy & Randles, 2011) and the discrete emotion approach (e.g., Izard, 2007, 2010; Izard, Woodburn & Finlon, 2010), in their original form, propose that anger, sadness, fear, disgust, happiness and surprise are each spontaneously expressed with universal configurations of unique facial movements (as depicted in Figure 1). These approaches acknowledge some variation in the facial movements that express the instances of each category, but they are hypothesized to result from processes that are external to the processes that caused the emotional instance, such as display rules, cultural learning, emotion regulation, and so on. Therefore, instances within the same emotion category would be expected to be relatively similar in their expressions (high reliability) across contexts, people and time, and instances of different emotion categories are expected to be relatively unique in their expressions (high specificity). Updated basic emotion theory allows for more variation in expressions within a category, assuming that instances of an emotion category vary around one supposed "best instance" (the prototype) that is either most typical or most frequent in nature.[43] Nonetheless, each instance of an emotion category is hypothesized to share enough of a characteristic facial expression that is *consistently present and recognizably different* from the patterns found in other emotion categories (for specific quotations, see Ekman, 1992, p. 550; Ekman & Cordaro, 2001, p. 364; Levenson, 2011, p. 379; Scarantino & Griffiths, pp. 448-449).

Intermediate along the horizontal dimension is the component process model of emotion which proposes that sequences of evaluations, referred to as appraisal checks, drive the dynamics of facial movements that express emotion (Scherer, Mortillaro & Mehu, 2017; Scherer et al., 2018). This approach belongs to a collection of causal appraisal approaches, which characterize appraisals as more than descriptive features of an emotional instance – they are considered to be actual causal cognitive mechanisms that produce instances of emotion, including how a situation is experienced and the associated facial expressions as components of the instance. Causal appraisal approaches acknowledge the possibility that, in principle, different temporally extended patterns of appraisals can cause instances of the same emotion category, and therefore is consistent with a larger variety of facial configurations that can express the emotion (for discussion, see Barrett, 2017c; Gross & Barrett, 2011; Scherer et al., 2017). In practice, however, this possible variation remains largely unexplored in scientific experiments.

At the other end of the continuum are approaches that explicitly predict substantial, situated variation within categories and similarity between categories, e.g., the theory of constructed emotion (Barrett, 2012, 2013, 2017a, b); core affect theory (Russell, 2003); functional approaches to emotion (relevant references, see Adolphs, 2017; Anderson & Adolphs, 2018; Campos et al., 1994); and descriptive appraisal approaches (e.g., Ortony & Clore, 2013; for a discussion, see Barrett, 2017c; Gross & Barrett, 2011). That is, they predict, in advance, that instances of the same emotion category are highly variable in their expressions (a specific set of facial movements have low reliability as expressions across people, situations and cultures) and instances of different emotion categories or even non-emotion categories are similar in their expressions (a given set of facial movements have low specificity as expressions of a single emotion category). For example, the theory of constructed emotion proposes that facial expressions of emotion are intrinsically constructed in a context-dependent way that has been learned in a particular culture.

Similarities and Differences in Deep Features

At the top of the "deep" continuum, basic emotion approaches propose that anger, sadness, fear, disgust, happiness and surprise are each caused by their own set of dedicated neural circuits (Tracy & Randles, 2011), a hypothetical affect program (Ekman & Cordaro, 2011), or a set of computations (Bach & Dayan, 2017). Some approaches allow for considerable variation in the physical causes of emotion and instead propose that each category shares a universal, functional similarity across situations (e.g., fear is the desire to escape from a predator; Adolphs, 2017; Anderson & Adolphs, 2018; Campos et al., 1994). At the bottom end of the continuum, theoretical approaches hypothesize that an emotion category does not exist in the brain as a fixed neural circuit, a fixed computation, or a fixed function. Instead, the human brain is thought to construct an emotion category on the fly, as needed for a situation-specific goal, with the help of the emotion concepts that person has acquired using the language they speak (Barrett, 2017a, b). This type of ad hoc category is called a goal-based category, because the similarity of its instances is based on the goal that the instances serve in a particular situation at a particular moment in time (Barsalou, 1983, 1985; Barsalou et al., 2003).

Importantly, most of these approaches explicitly anchor their hypotheses in evolutionary considerations, so it is misleading to refer to any one approach as an "evolutionary" approach. Most assume that emotion categories are psychological as well as biological categories. And both draw inspiration from Charles Darwin (albeit from different books, making very different assumptions about the nature of biological categories; for discussion, see Barrett 2017a, b).

Emotion Concepts

An emotion concept is a mental representation of an emotion category. Theoretical hypotheses about the proposed degree of variation in surface and deep features of an emotion category strongly relate to the proposed nature of emotion concepts. Approaches that hypothesize strong reliability and specificity in emotional expressions and other physical features shared by the instances of an emotion category, along with a deep causal similarity, propose that emotion categories are natural kind or *Aristotelian* categories (the red zone in Figure 1); correspondingly, instances of different categories can be distinguished by their physical features, like the facial movements expressing emotion. If an emotion category has a classical structure, then its corresponding concept reads like a dictionary definition that is stored in memory, describing its necessary and sufficient features.

A variety of theoretical approaches propose that emotion categories are prototype categories (the yellow zone in Figure 1), whose instances share some family resemblance. Of all the features that might describe the category, each instance might contain only a sample (resulting is more within-category variation and more between-category similarity than is true for Aristotelian categories). The corresponding emotion concept (its prototype) might be the most frequent instance found in the category, or its most typical instance (i.e., it is the instance that has all or most of the category's distinguishing features). Or the prototype might be a theory that describes the most typical instance (e.g., Clore & Ortony, 1991). The hypothesis that emotion categories are structured as prototypes is consistent with a variety of theoretical approaches in the science of emotion, including basic emotion approaches (e.g., Cowen & Keltner, 2017; Ekman & Cordaro, 2011), appraisal approaches (e.g., Shaver et al., 1987) psychological construction approach (e.g., Russell, 2003) and functional approaches (e.g., Campos et al., 1994; Campos, Campos & Barrett, 1989). The assumption, however, is that there is a single representation – a single prototype – for each category.

More recently, it has been proposed that emotion categories are goal-based, conceptual categories (green zone, Figure 1). The instances of a given emotion category are thought to share a common set of features *within a specific situation*, but these features (including the goal or function of the category) will change from situation to situation (Barrett, 2006, 2012, 2013, 2017a, b; Barrett, Wilson-Mendenhall, & Barsalou, 2015; LeBois et al., 2018; Wilson-Mendenhall et al., 2011). The hypothesis is that emotion categories, like all abstract categories, do not have conceptual cores (Barrett, Wilson-Mendenhall, & Barsalou, 2015; Wilson-Mendenhall et al., 2011), meaning that emotion concepts are constructed on the spot, as needed (i.e., they are ad hoc concepts).

Unlike the typological thinking that supports classical and prototype categories, conceptual categories are rooted in population thinking, or the idea that a biological category is populated with context-dependent, variable instances, so that any summary of the category (like a prototype) is an abstraction. The proposal that emotion categories are goal-based, conceptual categories, derives from Darwin's use of population thinking in On the Origin of Species (1859/2001; see Mayr, 2004), as well as Barsalou's research on grounded concepts (Barsalou, 1983, 1985, 2008; Barsalou et al., 2003), where the prototype of a category is context-dependent, and represents the ideal instance that best suits the function or goal in a specific situation, whether or not it actually exists in nature (e.g., Barsalou, 1993; Voorspoels, Vanpaemel, & Storms, 2011). Correspondingly, the hypothesis is that the similarity in an emotion category is not fixed or static – it varies from situation to situation because the similarity of its instances is based on the goal that the instances serve in a particular situation at a particular moment in time.

References

Adolphs, R. (2017). How should neuroscience study emotions? by distinguishing emotion states, concepts, and experiences. *Soc Cogn Affect Neurosci, 12*(1), 24-31. doi:10.1093/scan/nsw153

Anderson & Adolphs, 2018;

Bach, D. R., & Dayan, P. (2017). Algorithms for survival: a comparative perspective on emotions. *Nature Reviews Neuroscience, 18*, 311. doi:10.1038/nrn.2017.35

Barrett, L.F. (2006). Solving the emotion paradox: Categorization and the experience of emotion. *Personality and Social Psychology Review, 10*, 20-46.

Barrett, L. F. (2012). Emotions are real. *Emotion, 12*, 413-429

Barrett, L. F. (2013). Psychological construction: A Darwinian approach to the science of emotion. *Emotion Review, 5*, 379-389.

Barrett, L. F. (2017a). *How emotions are made: The secret life of the brain*. New York: Houghton-Mifflin-Harcourt.

Barrett, L. F. (2017b). The theory of constructed emotion: An active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience, 12,* 1-23. https://doi.org/10.1093/scan/nsw154

Barrett, L. F. (2017c). Functionalism cannot save the classical view of emotion. *Social Cognitive and Affective Neuroscience*, *12*, 34-36. https://doi.org/10.1093/scan/nsw156

Barrett, L. F., & Bliss-Moreau, E. (2009). Affect as a psychological primitive. *Advances in Experimental Social Psychology*, *41*, 167-218.

Barrett, L. F., Mesquita, B., Ochsner, K. N., & Gross, J. J. (2007). The experience of emotion. *Annual Review of Psychology, 58*, 373-403. PMC1934613

Barrett, L. F., Wilson-Mendenhall, C. D., & Barsalou, L. W. (2015). The conceptual act theory: A road map. Chapter in L. F. Barrett and J. A. Russell (Eds.), *The psychological construction of emotion* (p. 83-110). New York: Guilford.

Barsalou, L. W. (1983). Ad hoc categories. *Memory and Cognition, 11*(3), 211-227.

Barsalou, L. W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *J Exp Psychol Learn Mem Cogn, 11*(4), 629-654.

Barsalou, L. W., Simmons, W. K., Barbey, A. K., & Wilson, C. D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences, 7(*2), 84-91.

Campos, J. J., Campos, R. G., & Barrett, K. C. (1989). Emergent themes in the study of emotional development and emotion regulation. *Developmental Psychology, 25*(3), 394.

Campos, J. J., Mumme, D., Kermoian, R., & Campos, R. G. (1994). A functionalist perspective on the nature of emotion. *Japanese Journal of Research on Emotions, 2*(1), 1-20.

Clore, G. L. & Ortony, A. (1991). What more is there to emotion concepts than prototypes? *Journal of Personality and Social Psychology, 60*, 48-50.

Clore GL, Ortony A. (2013). Psychological Construction in the OCC Model of Emotion. *Emotion Review,* 5(4):335-343. doi:10.1177/1754073913489751.

Clore, G. L., & Ortony, A. (2008). Appraisal theories: How cognition shapes affect into emotion. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of Emotions* (3rd ed., pp. 628-642). New York, NY: Guilford Press.

Cowen, A. S. & Keltner, D. (2017). Self-report captures 27 distinct categories of emotion bridged by continuous gradients. *Proceedings of the National Academy of Sciences*, www.pnas.org/cgi/doi/10.1073/pnas.1702247114
Darwin's

Darwin, C. (1859/1964). *On the Origin of Species*. (A facsimile of the first edition. ed.). Cambridge, MA.: Harvard University Press.

Ekman, P. (1972). Universals And Cultural Differences In facial Expressions Of Emotions. In J. Cole (Ed.), *Nebraska Symposium on Motivation*, 1971 (pp. 207-283). Lincoln, Nebraska: University of Nebraska Press.

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

Ekman, P., & Cordaro, D. (2011). What is meant by calling emotions basic. *Emotion Review,*

Gross, J. J., & Barrett, L. F. (2011). Emotion generation and emotion regulation: One or two depends on your point of view. *Emotion Review, 3*, 8-16. PMC3072688

Izard, C. E. (2007). Basic emotions, natural kinds, emotion schemas, and a new paradigm. *Perspectives on Psychological Science, 2*, 260-280.

Izard, C. E. (2010). The many meanings/aspects of emotion: Definitions, functions, activation, and regulation. *Emotion Review, 2(4),* 363-370.

Izard, C. E., Woodburn, E. M., & Finlon, K. J. (2010). Extending emotion science to the study of discrete emotions in infants. *Emotion Review, 2(2),* 134-136.

Lazarus, R. (1991). *Emotion and Adaptation*. NY, Oxford University Press.

Lebois, L. A. M., Wilson-Mendenhall, C. D., Simmons, W. K., Barrett, L. F., & Barsalou, L.W. (2018). Learning situated emotions. *Neuropsychologia*. https://doi.org/10.1016/j.neuropsychologia.2018.01.008

Levenson, R. W. (2011). Basic emotion questions. *Emotion Review, 3(4)*, 379-386.

Mayr, E. (2004). *What makes biology unique? Considerations on the autonomy of a scientific discipline*. New York: Cambridge University Press.

Russell, J. A. (2003). Core affect and the psychological construction of emotion. *Psychological Review, 110(1)*, 145-172.

Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology, 76*, 805-819.

Scarantino, A., & Griffiths, P. (2011). Don't give up on basic emotions. *Emotion Review, 3(4)*, 444-454.

Scherer, K. R., Mortillaro, M., & Mehu, M. (2017). Facial expression is driven by appraisal and generates appraisal inferences. In J-M. Fernandez-Dols & J. A. Russell (Eds.), *The Science of Facial Expression* (p. 353-374). New York: Oxford U Press.

Scherer, K. R., Mortillaro, M., Rotondi, I., Sergi, I., & Trznadel, S. (2018). Appraisal-driven facial actions as building blocks for emotion inference. *Journal of Personality and Social Psychology, 114*(3), 358-379.

Shaver, P, Schwartz, J., Kirson, D, & O'Connor, C. (1987). Emotion knowledge: Further exploration of a prototype approach. *Journal of Personality and Social Psychology, 52,*1061-1086.

Tracy, J., & Randles, D. (2011). Four models of basic emotions: A review of Ekman and Cordaro, Izard, Levenson, and Panksepp and Watt. *Emotion Review, 3(4)*, 397-405.

Voorspoels, W., Vanpaemel, W., & Storms, G. (2011). A formal idea-based account of typicality. *Psychon Bull Rev, 18,* 1006-1014.

Wilson-Mendenhall, C. D., Barrett, L. F., Simmons, W. K., Barsalou, L. W. (2011). Grounding emotion in situated conceptualization. *Neuropsychologia, 49,* 1105-1127.

---

[43] The word "theory" is used here in its everyday definition to mean "a group of ideas" rather than its scientific definition to mean "a comprehensive explanation of some aspect of nature that is supported by a vast body of evidence" (National Academies of Science, http://www.nas.edu/evolution/TheoryOrFact.html).

Box 3: The Current Study of Context Effects in Emotion Perception

There is growing evidence that both the facial movements that express emotions and the emotional meaning inferred for those movements are strongly influenced by the contexts in which they occur. In the scientific study of how people infer emotions in the facial movements of other people, three types of context have been increasingly studied: situationally-based context, in which a face is physically presented with other sensory input that has some informational value, such as a body posture or a tone of voice, person-based context, in which the processes inside the brain or body the person inferring emotion in the facial movements of other people, and broader cultural contexts (for summaries, see Aviezer & Hassin, 2017; Barrett, Mesquita, & Gendron, 2011; de Gelder, 2016; Gendron, Mesquita & Barrett, 2013; Hess & Hareli, 2017; Wieser & Brosch, 2012).[1] In the majority of studies, other cues dominate emotion-related inferences, such that the emotional meaning of face is interpreted in line with its context, rather than its hypothesized emotional meaning as portrayed in Figure 4 (e.g., Aviezer et al., 2008, 2012; van den Stock et al., 2007; Wallbott, 1988; Wood, Martin, Alibali, & Niedenthal, 2018). Studies of how the context influence the activation of facial muscle movements are less frequent, but still exist (but for several clever examples, see Fernandez-Dols & Ruiz-Belda, 1997; Fridlund, 1991; Ruiz-Belda et al., 2003). Nonetheless, when context is properly acknowledged and assessed, scientific findings run contrary to the notion that facial muscle movements are universal expressions of emotion containing all of the information that is necessary and sufficient to communicate emotional states. They are consistent with functionalist, constructionist and behavioral ecology theories of emotion (Box 2). Contextual influences here are consistent with

[1] Gendron et al. (2013) was first written and submitted for publication in 2010.

evidence showing that context is intrinsically involved in the most basic aspects of movement and object perception.

The classic demonstration of how context impacts the interpretation of facial movements is the Kuleshov effect. In the early 20th century, the Soviet filmmaker Lev Kuleshov demonstrated that that the emotional meaning of an actor's neutral face changes depending on what was viewed immediately before: videos designed to evoke pleasantness (a little girl playing with a doll) or lust (a woman on a divan), unpleasantness (a dead woman in a coffin), and hungry (a bowl of soup) (Barratt et al., 2016; Mobbs et al., 2006; see Calbi et al., 2017 for a recent study as well as older references). These findings are echoed in more recent studies by the US psychologist Jim Russell and may explain the recent emergence of "resting bitch face" and "backpfeifengesicht" (literally, a face in need of a punch) in social media.

When scientists ask participants to pose an emotional expression in the absence of context (or in a singular, impoverished; e.g., Cordaro et al., 2017) or to infer emotion in a facial configuration that is absent any context except that provided by the experimental task, they are typically taking a reductionist approach to discover the emotional meaning of facial movements alone. Participants may instead be communicating a culturally established expectation, stereotype, or meme that may or may not hold in everyday contexts that are richer in the multimodal information that they carry (Srinivasan & Martinez, 2018). We focus our attention on one another's faces when we communicate, but this is not evidence that emotional information is carried solely, or even primarily, in the face alone (Aviezer et al., 2012). In real life, faces don't appear in isolation. Instead, they appear in a multi-sensory context that includes a body, a broader situational arrangement, and often a voice, smells, and so on. These additional sources of information are only considered "context" when the starting assumption is that the face is primary in communication emotion. In our view, this "context" in all its forms must be explicitly measured and modeled to achieve an ethology of emotional communication in the wild (Martinez, 2017b).

References

Aviezer, H., & Hassin, R. R. (2017). Inherently ambiguous: An argument for contextualized emotion perception. In J-M. Fernandez-Dols & J. A. Russell (Eds.), *The science of facial expression* (p. 333-349). NY: Oxford U Press.

Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., . . . Bentin, S. (2008). Angry, disgusted, or afraid? Studies on the malleability of emotion perception. *Psychological Science, 19*(7), 724-732.

Aviezer, H., Trope, Y., & Todorov, A. (2012). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science, 338*(6111), 1225-1229. doi:10.1126/science.1224313

Barratt, D., Rédei, A. C., Innes-Ker, A., & van de Weijer, J. (2016). Does the Kuleshov effect really exist? Revisiting a classic film experiment on facial expressions and emotional contexts. *Perception, 45*, 847–874.

Barrett, L.F., Mesquita, B., & Gendron, M. (2011). Context in emotion perception. *Current Directions in Psychological Science, 20*, 286-290.

Calbi, M., Heimann, K., Barratt, D., Siri, F., Umilta, M. A., & Gallese, V. (2017). How Context Influences Our Perception of Emotional Faces: A Behavioral Study on the Kuleshov Effect. *Front Psychol, 8*, 1684. doi:10.3389/fpsyg.2017.01684

Cordaro, D. T., Sun, R., Keltner, D., Kamble, S., Huddar, N., & McNeil, G. (2017). Universals and Cultural Variations in 22 Emotional Expressions Across Five Cultures. *Emotion*. doi:10.1037/emo0000302

de Gelder, B. (2016). *Emotions and the body*. New York: Oxford University Press.

Fernandez-Dols, J-M., Sanchez, F., Carrera, P., & Ruiz-Belda, M-A. (1997). Are spontaneous expressions and emotions linked? An experimental test of coherence. *Journal of Nonverbal Behavior, 21*, 163-177.

Fridlund, A. J. (1991). Sociality of solitary smiling: Potentiation by an implicit audience. *Journal of Personality and Social Psychology, 60(2),* 229-240.

Gendron, M., Mesquita, B., & Barrett, L. F. (2013). Emotion perception: Putting a face in context. Chapter in D. Reisberg (Ed.), *Oxford Handbook of Cognitive Psychology* (p. 539-556). New York: Oxford University Press.

Hess, U., & Hareli, S. (2017). The social signal value of emotions: The role of contextual factors in social inferences drawn from emotion displays. In J-M. Fernandez-Dols & J. A. Russell (Eds.), The science of facial expression (p. 375-392). NY: Oxford U Press

Martinez, A. M. (2017b). Visual perception of facial expressions of emotion. *Current Opinion in Psychology, 17*, 27-33.

Mobbs, D., Weiskopf, N., Lau, H., Featherstone, E., Dolan, R., & Frith, C. (2006). The Kuleshov Effect: The influence of contextual framing on emotional attributions. *Social Cognitive and Affective Neuroscience, 1*, 95–106.

Ruiz-Belda, M. A., Fernandez-Dols, J-M., Carrera, P., & Barchard, K. (2003). Spontaneous facial expressions of happy bowlers and soccer fans. *Cognition and Emotion, 17*, 315-326.

Srinivasan, R., and Martinez, A.M. (2018). Cross-cultural and cultural-specific production and perception of facial expressions of emotion in the wild. IEEE Transactions on Affective Computing. ArXiv preprint, available at https://arxiv.org/pdf/1808.04399.pdf

Van den Stock, J., Righart, R., & De Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion, 7(3),* 487.

Wallbott, H. G. (1988). In and out of context: Influences of facial expression and context information on emotion attributions. British Journal of Social Psychology, 27, 357-369.

Wieser, M. J., & Brosch, T. (2012). Faces in context: a review and systematization of contextual influences on affective face processing. *Front Psychol, 3*, 471. doi:10.3389/fpsyg.2012.00471

Wood, A., Martin, J. D., Alibali, M. W., & Niedenthal, P. M. (2018). A sad thumbs up: incongruent gestures and disrupted sensorimotor activity both slow processing of facial expressions. *Cognition and Emotion*, 1-14.

Box 4: The Origin of the Proposed Facial Expressions

The belief that certain configurations of facial movements express emotion and therefore display a person's inner state of mind reached prominence in the paintings by Rembrandt in the 17th century. Different sets of universal expressive forms were proposed by various artists, such as in the drawings of Charles Le Brun, the 17th century French painter for King Louis XIV of France (Montagu, 1994), in the facial movements found in Hindu Navarasa dance (https://www.youtube.com/watch?v=1uIan6u3UrQ), in the heads depicted by the 18th century German sculptor Franz Xaver Messerschmidt (https://www.pinterest.com/tkearneyg/franz-xaver-messerschmidt/), in moving photographs by the 19th century Czech neuroscientist Jan Evangelista Purkinje (Wade, 2016), in the prescriptions of the 19th century French dramatist François Delsarte (Stebbins & Delsarte, 1887/2013), and in the photographs taken by the French physician Guillaume-Benjamin-Amand Duchenne, which were highlighted in the writings of Charles Darwin (Darwin, 1872). These varying proposals were united by the assumption that each emotion category was expressed by one specific facial configuration, making it possible to infer one (emotional state) from the other (the configuration of facial movements).

How did the science of emotion end up focusing almost exclusively on the facial configurations displayed in Figure 4? The story of their origin has been discussed in Ekman, Friesen & Ellsworth (1972), Gendron & Barrett (2017) and Russell (1994). Darwin did not discover the facial configurations through careful observation in the same way that such observations led him to discover the idea of natural selection. Instead, he stipulated them based on drawing by Bell (1806) and photographs by Duchenne (1990/1862), continuing the tradition of others who, before him, stipulated other facial movements as the expressions of emotions (see introduction in main paper).

Darwin conducted two informal studies of emotional expressions. In *The Expression of the Emotions in Man and Animals*, he described an informal survey about these facial movements that he believed expressed emotions in a way that is shared with other animal species and therefore in a specific and universal way around the world. He provided his colleagues with verbal descriptions of specific expressive forms along with the emotion category he thought they expressed, and he asked his colleagues (living in various parts of the world) whether they believed his hypotheses were true. Darwin also conducted a second informal study that he described in a letter to a colleague, in which he presented 24 participants with 11 static photographs of facial configurations elicited by electrical stimulation of facial muscles (detailed in Snyder, Kaufman, Harrison, & Maruff, 2010). These photographs were taken by Duchenne, who believed that facial muscles produce expressions that reveal a person's inner state. Duchenne created over 60 photographs of facial configurations that are often referred to as "induced emotional expressions," but actually the photographs capture exaggerated facial muscle contractions elicited by external electrical stimulation. Darwin's description of the study suggests he asked his participants to freely describe the emotion presenting in each photo (i.e., he used a free-labeling response method).

Research that proceeded in the early 20th century attempted to replicate and extend Darwin's second study using photographs of faces posed in exaggerated configurations. Others went about the task of detailing the specific facial movements that constitute the configurations that were believed to be emotional expressions. Few research studies evaluated whether emotional expressions made in everyday life actually conform to these portrayals. The emotion perception studies during this period generally showed that perceivers were highly variable in the in the emotional causes they inferred for each configuration, with little reliability and specificity.

(The one exception was studies using a choice-from-array approach involving brief vignettes describing each emotion category, referred to as the Dashiell method). This larger body of research gave rise to a scientific era that was guided by the hypothesis emotions and their expressions were socially constructed and culturally variable.

In the early 1960's the U.S. psychologist Paul Ekman and his colleagues resurrected a research paradigm to test Darwin's original ideas about emotional expressions. They took a large set of photographs developed by Sylvan Tomkins, in which actors posed what they believed to be the expression of anger, contempt, disgust, fear, happiness, interest, sadness, shame and surprise categories. Participants viewed the candidate poses and based on their reliability in choosing the expected emotion word to label each photo, a final set was chosen. The focus on six categories -- anger, disgust, fear, happiness, sadness, and surprise was accidental, not based on theoretical considerations – portrayals of contempt, interest and shame expressions were not labeled reliably and so were not initially studied. Ekman and colleagues also incorporated work by the Swedish anatomist Carl-Herman Hjortsjö, who catalogued Duchenne's facial configurations and stipulated them to be emotional expressions (Hjortsjö, 1969).

This is the origin of the facial configurations in Figure 4, which are used in the majority of experiments of emotion perception and that correspondingly constitute commonsense beliefs about how certain emotion categories are specifically expressed. Many different sets of facial poses haven been developed over the years, a selection of which can be found here:

https://rystoli.github.io/FSTC.html

http://cbcsl.ece.ohio-state.edu/downloads.html

http://cseweb.ucsd.edu/~gary/CAFE/

https://ieeexplore.ieee.org/document/8013713/

References

Bell C. (1806). *Essays on the Anatomy of Expression in Painting*. London: Longman, Hurst,Rees & Orme.

Darwin, C. (1872/1965). *The Expression of the Emotions in Man and Animals* (Chicago: University of Chicago Press).

Duchenne CB (1990/1862). *The Mechanism of Human Facial Expression* (Renard, Paris); reprinted (1990) (Cambridge Univ Press, London).

Ekman, P., Friesen, W., & Ellsworth, P. (1972). *Emotion in the human face: guidelines for research and an integration of findings*. Elmsford, N.Y.: Pergamon Press.

Gendron, M., & Barrett, L. A. (2017). Facing the past: A history of the face in psychological research on emotion perception. In J-M. Fernandez-Dols & J. A. Russell (Eds.), *The science of facial expression* (p. 15-36). NY: Oxford U Press.

Hjortsjö, C. H. (1969). *Man's face and mimic language*. Studen litteratur.

Montagu, J. (1994). *The expression of the passions: The origin and influence of Charles Le Brun's lecture on general and particular expression*. London: Yale University Press.

Russell, J. A. (1994). Is there universal recognition of emotion from facial expressions? A review of the cross-cultural studies. *Psychological Bulletin, 115(1)*, 102-141.

Snyder, P. J., Kaufman, R., Harrison, J., & Maruff, P. (2010). Charles Darwin's emotional expression "experiment" and his contribution to modern neuropharmacology. *Journal of the History of the Neurosciences*, 19:2, 158-170, doi: 10.1080/09647040903506679

Stebbins, G. & Delsarte, F. (1887/2013). Delsarte system of expression. E. S. Werner.

Wade, N. J. (2016). Faces and photography in 19th-century visual science. *Perception, 45*, 1008-1035.

Box 5: The Face: Anatomy of Facial Muscles and Facial Electromyography

The Anatomy of Facial Muscles

To understand how people make certain patterns of facial movements, it's helpful to consider how facial muscles develop and are structured. Healthy humans have a common set of 17 facial muscles on each side of the face that contract and relax in patterns (Rinn, 1984) [2]. Facial muscles develop in utero. By 36 weeks, almost all of the muscles used to produce facial movements are formed (Gasser, 1967), although their morphology is not identical to an adult human. The muscles are controlled by cranial nerves V and VII which become functional a bit earlier, by 11 weeks in utero (Reissland et al., 2011). Facial muscle movements support many functions, such as sucking movements necessary for feeding and tongue movements necessary for speech (Haywood & Getchell, 2014). As a consequence, the development of facial muscles influences a wide range of behaviors and capacities over and above expressing emotion.

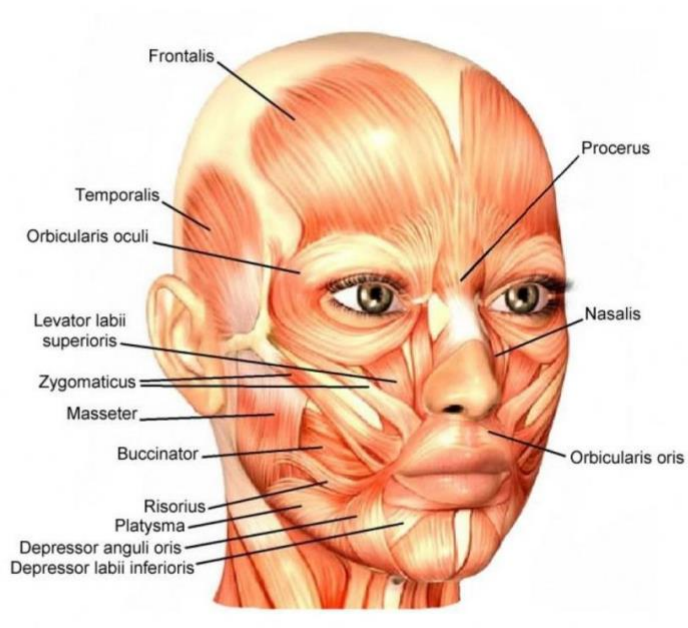
To create facial movements that are visible to the naked eye, facial muscles contract, moving skin into folds and wrinkles on an underlying skeletal structure. These are the movements people observe in one another, that are captured in photographs and videos, and that express emotions. These visible facial articulations are called Action Units (AUs). Clear and consistent data are currently not available to indicate when facial expressions are first formed. While fetuses in utero definitely make some facial movements, it is currently entirely unknown if these have anything to do with emotion-like states. Some scientists claim that young infants produce facial movements to pain that resemble those of adults (Izard et al., 1987), even though the underlying anatomy is not morphologically identical (Camras & Shutter, 2010).

---

[2] There are somewhat different numbers often given, because a single muscle may comprise more than one functional unit.

Individual differences in facial anatomy, as well as in the brain's control of facial muscles, cause variation in the details of how facial movements are executed at the muscular level and how they look to the naked eye. People vary in the underlying bone structure of the face and details of the skin, the structure and strength of their facial muscles (Pessa et al., 1988), the dynamics of facial muscle movements, and consequently they vary in how their facial movements look to a human observer (e.g., Farahvash et al., 2010; Shim et al., 2008; Shimada & Gasser, 1989). In addition, some people have strong asymmetries for one side of the face or the other, and some people lack certain smaller muscles altogether (Waller, 2008). In fact, if you inserted your exact facial muscles into a different face (someone with a different bone structure or someone much older or younger than you, or whose face is thinner or fatter), the resulting muscle movements would look different than they do on your face. And even when facial movements look the same to the naked eye, there may be differences in their execution under the skin. As human perceivers, we see stable facial behaviors (i.e., a frown) when in reality, under the skin, there is more variation that meets the eye. A facial behavior, like frowning, or scowling, is, in fact, a category of variable instances. When you watch a frown unfold in the same person on two different occasions, the exact muscle contractions that curl the upper lip and turn down the corners of the mouth can subtly vary from one instance to another. What to the naked eye looks like the same frown in two different people can result from different patterns of underlying muscle contractions.

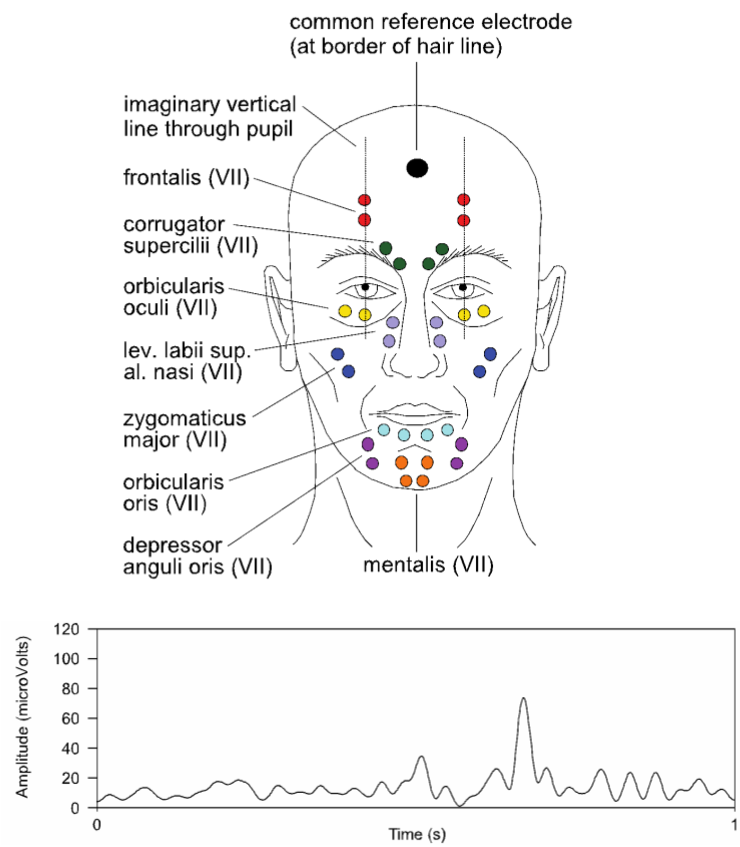
Put simply: something as seemingly simple as a single facial movement is best understood as a conceptual category, resulting from a variable set of more basic, variable physical changes. The same is true for all motor movements, for smells and for sounds. For example, the sound of a "b" is acoustically different when heard in the words "bad" and "bed", yet human brains wire themselves to hear both as the sound of "b" (i.e., the sound of a "b" is a category; Barsalou 1992).



Box Figure 5-1. The muscles of the face. From www.hdanatomy.com. Permission pending.

Measuring Facial Movements with Facial Electromyography (fEMG)

The most sensitive and direct way to measure facial muscle movements is to record the electrical activity of facial muscles as they contract. This is called facial electromyography, or facial EMG, and is done by placing electrodes in or on the muscles of the face. Actually inserting thin electrodes into the muscles provides the most sensitive and specific electrical recordings, but this approach is rarely used since it is more invasive. It is more common to use surface recordings from electrodes placed on the skin that measure the changes in electrical potential of muscle depolarization. EMG recordings of facial muscle movements were first made in the late 1950s and were later used for studying emotion and affect in the 1970s (Tassinary et al., 2007). The specificity with which individual muscles can be distinguished from one another depends on how many electrodes are put on the face. Typically, only three to six electrodes are put on the face, although many more than that can be uncomfortable for a test subject. This means that many facial muscle movements are not measured in most studies that use facial EMG. Current research suggests that specific patterns of facial EMG activity reliably distinguish between pleasant vs. unpleasant states, as well as the intensity of the states along with how social the situation is, but that they do not reliably distinguish between different individual categories of emotion (Cacioppo et al., 2000).



Box Figure 5-2. Measuring facial electromyography (fEMG). Left: Common electrode locations, showing some of the facial muscles whose activity they can measure. VII indicates that all the muscles are controlled by the facial nerve (7th cranial nerve). Right: example of a recorded EMG signal. The trace plots the absolute value of the change in electrical activity (in microvolts) versus time (in seconds). Reproduced with permission from van Boxtel (2010). See Figure 4 for a summary of how combinations of these muscles contribute to facial action units in FACS coding.

References

Barsalou, L.W. (1992). Frames, concepts, and conceptual fields. In: Frames, Fields, and Contrasts: new essays in semantic and lexical organization, pp. 21-74. Lehrer, A. and Kittay, E.F., eds. Hillsdale, NJ: Lawrence Erlbaum.

Camras, L., and Shutter, J.M. (2010). Emotional facial expressions in infancy. Emotion Review 2, 120-129.

Cacioppo, J. T., Berntson, G. G., Larsen, J. T., Poehlmann, K. M., & Ito, T. A. (2000). The psychophysiology of emotion. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of emotions* (Vol. 2, pp. 173–191). New York, NY: Guilford Press.

Gasser, R.F. (1967). The development of the facial muscles in man. Developmental Dynamics 120, 357-375.

Farahvash, M.R., Abianeh, S. H., Farahvash, B., Farahvash, Y., Yagoobi, A., & Nazparvar, B. (2010). Anatomic variations of midfacial muscles and nasolabial crease: A survey on 52 hemifacial dissections in fresh Persian cadavers. *Aesthetic Surgery Journal*, 30(1) 17–21.

Haywood, K.M., and Getchell, N. (2014). Life Span Motor Development, 6th Edition, (eBook: Human Kinetics).

Izard, C.E., Hembree, E.A., and Huebner, R.R. (1987). Infants' emotion expressions to acute pain: developmental change and stability of individual differences. Developmental Psychology 23, 105-113.

Pessa, J.E., Zadoo, V.P., Adrian, E.K., Yuan, C.H., Aydelotte, J., and J.R, G. (1988). Variability of the midfacial muscles: analysis of 50 hemifacial cadaver dissections. Plastic and Reconstructive Surgery 102, 1888-1893.

Reissland, N., Francis, B., Mason, J., and Lincoln, K. (2011). Do facial expressions develop before birth? PLoS One 6, e24081.

Rinn, W.E. (1984). The neuropsychology of facial expression: a review of the neurological and psychological mechanisms for producing facial expressions. Psychological Bullettin 95, 52-77.

Shim, K-S., Hu, K-S., Kwak, H-H., Koh, K-S., Fontaine, C., & Kim, H-J. (2008). An anatomical study of the insertion of the zygomaticus major muscle in humans focused on the muscle arrangement at the corner of the mouth. *Plastic and Reconstructive Surgery, 121*, 466-473.

Shimada, K., and Gasser, R.F. (1989). Variations in the facial muscles at the angle of the mouth. Clinical Anatomy 2: 129-134.

Tassinary, L.G: Chapter in: Cacioppo, J.T., Tassinary, L.G., and Berntson, G.G. (2007). Handbook of Psychophysiology, 3rd edition, (Cambridge, England: Cambridge University Press).

Van Boxtel, A. (2010). Facial EMG as a tool for inferring affective states. Proceedings of Measuring Behavior 2010. Eds. Spink, Grieco, Krips, Loijens, Noldus, Zimmerman. Eindhoven, the Netherlands (August 24-27, 2010).

Waller, B.M. (2008). Selection for universal facial emotion. Emotion 8, 435-439.

Box 6: A Summary of Computer Vision Algorithms for Automatically Detecting Facial Actions

The goal of computer vision algorithms is to identify a functional mapping between to variables x and y, where x is the image of a face and y the list of active action units (AUs) observed in the face in x, i.e., AUs=f(face image). There are two main solutions to the problem of identifying f(.). One is to manually derive functions that we believe do a good job at identifying AUs in images. The other approach is to learn f(.) from the data – from pairs of samples x and y. This second approach is called **machine learning**.

Manually-Derived Functions

There are three types of manually-derived functions for solving AUs=f(face image).

1. *Optical flow*: Optical flow is the apparent motion of landmark points on an object, such as a face (Kroeger et al., 2016), e.g., how does the landmark point defining the left corner of the mouth appear to move when we activate AU 12 (lip corner puller). A dense optical flow approach is when most of the pixels defining a face are landmark points. A computer vision algorithm can use dense optical flow to identify AUs on a face by computing the apparent movement of a set of facial landmark points between a generic neutral face and the given image (Donato et al., 1999; Yacoob & Davis, 1996; Martinez, 2003). One way to improve the performance of this algorithm is to use a person-specific approach, where the neutral face used to compute the optical flow is an image of the same individual as that in the image we need to analyze. Advantages: This approach generally worked well in lab conditions (>80%) and where the face is seen frontally and is not occluded. The algorithm's performance improves (>90%) when we use a person-specific approach. Disadvantages: Both a person's

neutral face (at rest) and the image of the apex of a facial configuration must first be known. Detection of the apex is a problem that has received scarce attention in computer vision. This, combined with the need for non-occluded, frontal images, has limited the use of this approach to images filmed in controlled, lab conditions. The method is also limited to people for whom we have a neutral face available or one can be detected in a video sequence.

2. *Linear filters*: Another solution is to use a linear filter h(.). A linear filter (such as a Gabor filter) is a functional mapping that acts locally on an image and has the linear constraints. The functional mapping is given by a convolution of the image with the filter h(.). A convolution can be intuitively understood as the multiplication of the pixels of the filter with a local patch of the image of the same size as the filter, with this multiplication being applied to every possible placing of the filter on the image. The response of these local multiplications is generally different when the AU is active in a face than when it is not. Because of this, local filters have been extensively used to identify AUs in images of faces (Benitez-Quiroz et al., 2016; De la Torre & Cohn, 2011; Liu & Wechsler, 2002; Lyons et al., 1999; Tian et al., 2002). Advantages: Convolutions are fast operators. This makes for fast (i.e., processing at >30 frames per second) and accurate algorithms (>90%) detection of AUs in images filmed in the lab as well as in the wild. To date, this is one of the most successful approaches to identifying some (but not all) AUs (Benitez-Quiroz et al., 2017a). Disadvantages: Local filters only work when the texture in the patch of the image is well defined. When the local patch is smooth (e.g., middle of the cheeks), the response of a filter is similar for active versus inactive AUs.

3. *Shape descriptors*: The shape of an object is defined as the geometric properties of the object when all information related to the object's position, scale and rotation have been eliminated

(Hamsici & Martinez, 2009). Shape is especially useful to define AUs that deform major components of the face, e.g., the lips, eyelids, eyebrows, nose, and jaw line. Shape can also be used to define the wrinkles seen in the forehead when activating AUs 1 (inner brow raiser) and 2 (outer brow raiser) as well as those around the nose caused by AUs 9 (nose wrinkler) and 10 (upper lid raiser). These descriptors have been successfully used to detect these and other AUs in images filmed in the lab and in the wild (Benitez-Quiroz et al., 2016; Martinez & Du, 2012; Neth & Martinez, 2010; Kotsia et al., 2008). Advantages: Works in local patches with little texture and contrast. Can be applied to low-resolution images and under varying pose. Disadvantages: Not all image changes caused by AUs define an easily detectable shape change, e.g., AUs 11 (nasolabial deepener) and 14 (dimpler).

Machine Learning-Derived Functions

There are two types of machine-learning functions for solving AUs=f(face image).

1. *Probabilistic algorithms*: Statistical learning theory and statistical pattern recognition are algorithms that learn from sample pairs: $S=\{(x_1,y_1),\ldots,(x_n,y_n)\}$, where $y_i=f(x_i)$, and f(.) is generally a probability density function (pdf) or a mixture of pdfs (McLachlan & Peel, 2004). Intuitively, a pdf is a function that give the relative likelihood of a value of x_i to belong to a value of y_i. Given the training set S, we can estimate these likelihoods and, hence, the pdf. This approach works best when combined with the computer vision features defined above (optical flow, linear filters, and shape) (Benitez-Quiroz et al., 2016; Zafeiriou et al., 2016; Corneanu et al., 2016). Advantages: To date this is the most successful approach to the recognition of AUs in the lab and in the wild (Benitez-Quiroz et al., 2016, 2017a, 2019). Disadvantages: The training set S needs to include images under a variety of image conditions (illumination, pose) as well as people of diverse

ethnicities, races and skin colors. The lack of diversity in these datasets has resulted in technology that is based to minorities (Buolamwini, 2018).

2. *Deep learning*: An alternative to the above approach is to learn the function f(.) using regression analysis. Regression analysis identifies relationships between predictor variables and dependent variables, much like probabilistic algorithms. The difference here is that we need not use a probabilistic model (Belkin & Niyogi, 2003). Rather, we assume the functional mapping provides a direct relationship between x and y (see, for example, Martinez, 2017a, for some examples). In this approach, the goal is to estimate the parameters of the manifold (function). If the function is linear, we can use linear least-squares. If the function is non-linear (which is typically the case), we need to use a non-linear optimization approach. One famous solution is gradient descent (Rumelhart et al., 1988). Additionally, when the number of parameters to be estimated is very large, this approach is called *deep learning*. Deep learning has been extremely successful in many computer vision applications. Missing from the list of successful applications, however, is the recognition of AUs. The reason for this is simple: to estimate a very large number of parameters in f(.), we require an equally large number of sample pairs (x_i,y_i), $i=1,\ldots,n$, with n large. As we discussed in the paper though, manually annotating a large number of images (x_i) with their corresponding AUs (y_i) would require years and millions of dollars to complete. A solution that is being attempted is to use automatic annotations instead (Benitez-Quiroz et al., 2017b; Zhao et al., 2016b). That is, we can use current algorithms to automatically annotate AUs in a large number of images. One then uses these annotations to train a deep learning algorithm that is robust to errors in the training data. We are awaiting additional experiments to determine if this approach will succeed. One

solution may be to use deep learning to generate realistically-looking samples, typically called deep fakes. A recent paper (Pumarola et al., 2018) shows this should be possible.

Improving Automatic Annotations

Researchers in computer vision are studying ways of improving the performance of current algorithms to map facial movements to AU codes (Corneanu et al., 2016). One promising approach identifies dependencies in facial movements (Zhao et al., 2016). Facial movements typically do not occur in isolation. Try this: Face a mirror and try to move the outer corners of your eyebrows (AU2) while keeping every other facial muscle at rest. You will notice this is a difficult task. Algorithms can learn these dependencies to improve the automatic annotation of AUs of any of the algorithms described above (Benitez-Quiroz et al., 2017b). For example, if we uncover that AUs 1 and 2 are co-articulated quite often and that AU 1 almost never co-occurs with AU 23, we can use this knowledge to improve the annotations of AU 1 once we know whether a face has AU 2 and/or 23 active. Box Figure 1 shows the dependencies obtained on a small set of posed facial configurations taken in controlled, lab conditions. Learning AU dependencies is important to understand the evolution of AU production from birth to adulthood. For example, if these dependencies are small in babies and large in adults, it suggests a learning or developmental process in the production of facial configurations; no changes may indicate an innate system that is anatomically constrained and possibly available at birth. In addition, if these dependencies vary across cultures, it would point to a cultural influence in the malleability of facial movements; no cultural variation would suggest that there are universal (possibly anatomical or neural) constraints on the facial configurations that a human face can produce. These hypotheses about AU dependencies await future research.

Research to Validate Automated Coding of AUs Relies on Supervised Learning

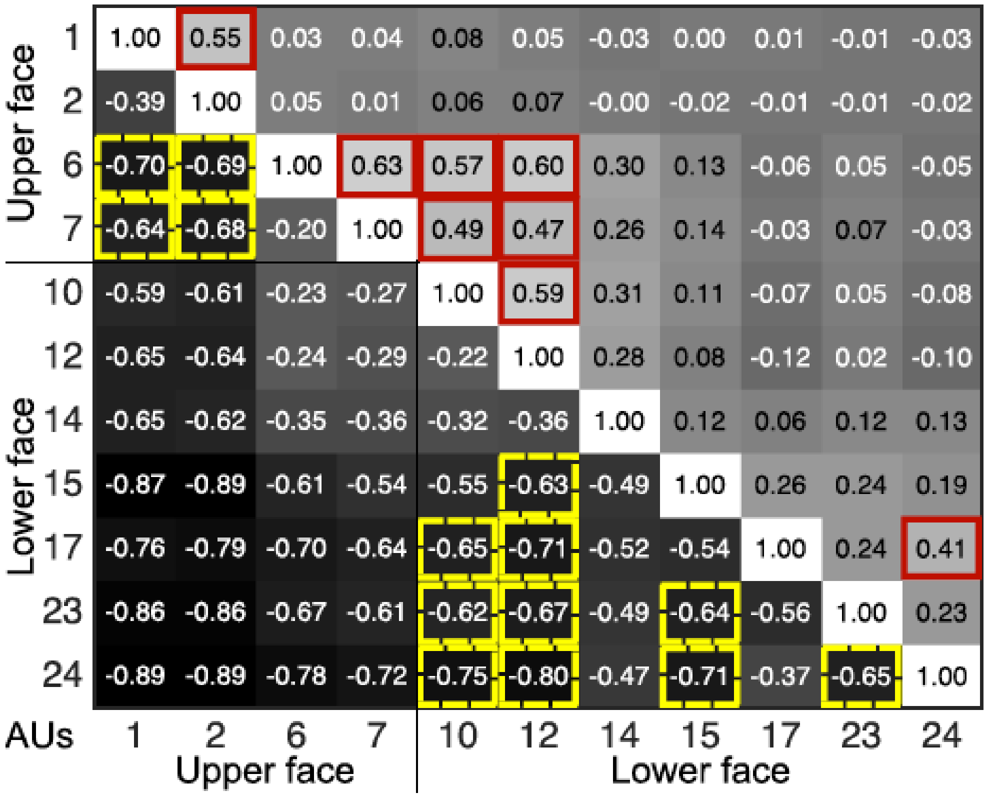
For the above algorithms to work, we need the training set S defined above. That is, we need a set of images whose AUs have been coded by an expert FACS coder (i.e., the images are manually annotated). Many databases have been collected of posed and spontaneous facial configurations in lab conditions (for example, see Lucey et al., 2010; Mavadati et al., 2013). These databases typically have between a few to several hundred training images. Recently, a large database of facial expressions in the wild was collected and made available to researchers (Benitez-Quiroz et al., 2016). This large dataset was used in the EmotioNet Challenge[3] (Benitez-Quiroz et al., 2017a), as discussed in the paper. To evaluate the accuracy of AU annotation of these computer vision algorithms, part of the dataset is used for training an algorithm and an independent set of images is used for testing the accuracy of the algorithm. Accuracy of annotations as well as true and false positive and true and false negative are usually reported.

Temporal Information

Temporal features have also been used to identify AUs (Bartlett et al., 2014). Unfortunately, these algorithms are not based on a natural model of facial movement dynamics (Cohn & Schmidt, 2004; Zuckerman et al., 1976). Instead, these algorithms attempt to identify correlations between temporal *image* features with the presence of AUs. Ideally, we would like to have algorithms that use the same temporal features employed by humans. Unfortunately, these temporal features are, for the most part, unknown at present. Hence, there is a need to better understand the dynamics of facial movements and how these dynamics aid detection and contribute to their inferred psychological meaning. This is an area of future research for both psychologists and computer scientists.

---

[3] http://cbcsl.ece.ohio-state.edu/EmotionNetChallenge/index.html



Box Figure 6-1. *Co-articulation of action units*. Positive correlations (red) and negative correlations (yellow) between facial actions, estimated from a set of 350,000 frames of facial movements. Adapted from Zhao et al. (2016). This approach uses dependencies to predict the presence or absence of AUs before they are detected, improving the accuracy of the algorithms to detect facial movements and map them to AUs (Benitez-Quiroz et al., 2017b). These algorithms have also been recently tested in less constrained conditions (i.e., in the wild; Zafeiriou et al., 2016).

References

Bartlett, M. S., Littlewort, G. C., Frank, M. G., & Lee, K. (2014). Automatic decoding of facial movements reveals deceptive pain expressions. Current Biology, 24(7), 738-743.

Belkin, M., & Niyogi, P. (2003). Laplacian eigenmaps for dimensionality reduction and data representation. Neural Computation, 15(6), 1373-1396.

Benitez-Quiroz, C. F., Srinivasan, R., & Martinez, A. M. (2016). Emotionet: An accurate, real-time algorithm for the automatic annotation of a million facial expressions in the wild. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 5562-5570).

Benitez-Quiroz, C. F., Srinivasan, R., Feng, Q., Wang, Y., & Martinez, A. M. (2017a). EmotioNet Challenge: Recognition of facial expressions of emotion in the wild. arXiv preprint arXiv:1703.01210.

Benitez-Quiroz, C. F., Wang, Y., & Martinez, A. M. (2017b). Recognition of Action Units in the Wild with Deep Nets and a New Global-Local Loss. In Proc. International Conference on Computer Vision (pp. 3990-3999).

Benitez-Quiroz, F., Srinivasan, R., & Martinez, A. M. (2019). Discriminant Functional Learning of Color Features for the Recognition of Facial Action Units and their Intensities. IEEE Transactions on Pattern Analysis and Machine Intelligence.

Buolamwini, J., & Gebru, T. (2018, January). Gender shades: Intersectional accuracy disparities in commercial gender classification. In Conference on Fairness, Accountability and Transparency (pp. 77-91).

Cohn, J. F., & Schmidt, K. L. (2004). The timing of facial motion in posed and spontaneous smiles. International Journal of Wavelets, Multiresolution and Information Processing, 2, 121-132.

Corneanu, C. A., Simón, M. O., Cohn, J. F., & Guerrero, S. E. (2016). Survey on rgb, 3d, thermal, and multimodal approaches for facial expression recognition: History, trends, and affect-related applications. IEEE Transactions on Pattern Analysis and Machine Intelligence, 38(8), 1548-1568.

De la Torre, F., & Cohn, J. F. (2011). Facial expression analysis. In Visual analysis of humans (pp. 377-409). Springer, London.

Donato, G., Bartlett, M. S., Hager, J. C., Ekman, P., & Sejnowski, T. J. (1999). Classifying facial actions. IEEE Transactions on Pattern Analysis and Machine Intelligence, 21(10), 974.

Hamsici, O. C., & Martinez, A. M. (2009). Rotation invariant kernels and their application to shape analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, 31(11), 1985-1999.

Kotsia, I., Zafeiriou, S., & Pitas, I. (2008). Texture and shape information fusion for facial expression and facial action unit recognition. Pattern Recognition, 41(3), 833-851.

Liu, C., & Wechsler, H. (2002). Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. IEEE Transactions on Image processing, 11(4), 467-476.

Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., & Matthews, I. (2010, June). The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on (pp. 94-101). IEEE.

Lyons, M. J., Budynek, J., & Akamatsu, S. (1999). Automatic classification of single facial images. IEEE Transactions on Pattern Analysis and Machine Intelligence, 21(12), 1357-1362.

Martinez, A. M. (2003). Matching expression variant faces. Vision Research, 43(9), 1047-1060.

Martinez, A., & Du, S. (2012). A model of the perception of facial expressions of emotion by humans: Research overview and perspectives. Journal of Machine Learning Research, 13(May), 1589-1608.

Martinez, A. M. (2017a). Computational models of face perception. Current Directions in Psychological Science, 26(3), 263-269.

Martinez, A. M. (2017b). Visual perception of facial expressions of emotion. Current Opinion in Psychology, 17, 27-33.

Mavadati, S. M., Mahoor, M. H., Bartlett, K., Trinh, P., & Cohn, J. F. (2013). Disfa: A spontaneous facial action intensity database. IEEE Transactions on Affective Computing, 4(2), 151-160.

McLachlan, G., & Peel, D. (2000). Finite mixture models. Willey, series in probability and statistics.

Neth, D., & Martinez, A. M. (2010). A computational shape-based model of anger and sadness justifies a configural representation of faces. Vision Research, 50(17), 1693-1711.

Pumarola, A., Agudo, A., Martinez, A. M., Sanfeliu, A., & Moreno-Noguer, F. (2018, July). Ganimation: Anatomically-aware facial animation from a single image. In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 818-833).

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. Nature, 323, pp. 533-536, 533.

Tian, Y. L., Kanade, T., & Cohn, J. F. (2002, May). Evaluation of Gabor-wavelet-based facial action unit recognition in image sequences of increasing complexity. In Automatic Face and Gesture Recognition, 2002. Proceedings. Fifth IEEE International Conference on (pp. 229-234). IEEE.

Yacoob, Y., & Davis, L. S. (1996). Recognizing human facial expressions from long image sequences using optical flow. IEEE Transactions on Pattern Analysis and Machine Intelligence, 18(6), 636-642.

Zafeiriou, S., Papaioannou, A., Kotsia, I., Nicolaou, M., & Zhao, G. (2016). Facial Affect``In-The-Wild. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Workshops (pp. 36-47).

Zhao, K., Chu, W. S., De la Torre, F., Cohn, J. F., & Zhang, H. (2016). Joint patch and multi-label learning for facial action unit and holistic expression recognition. IEEE Transactions on Image Processing, 25(8), 3931-3946.

Zhao, K., Chu, W. S., & Martinez, A. M. (2018). Learning Facial Action Units From Web Images With Scalable Weakly Supervised Clustering. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 2090-2099).

Zuckerman, M., Hall, J. A., DeFrank, R. S., & Rosenthal, R. (1976). Encoding and decoding of spontaneous and posed facial expressions. Journal of Personality and Social Psychology, 34(5), 966.

Box 7: Variations in Facial Movements

In principle, a human face can make a multitude of movement patterns: 16 million different combinations are possible, in principle, assuming each of the 24 AUs corresponding to muscle movements could move independently (ignoring temporal dynamics; Martinez, 2017b). In addition, facial muscles can contract with different intensities and varying time to peak contraction (Jack & Schyns, 2017), further increasing the number of movement patterns a face can generate, in principle. In practice, however, a much smaller subset of combinations is likely because of the way that brain controls the face as well as various anatomical constraints (e.g., some muscles are more or less likely to move together because of their relative position to one another, how they are attached to facial bones, or how they are innervated by nerves).

Several obstacles make it challenging to scientifically observe which configurations are made by healthy people and with what frequency. The most serious obstacle is one of naming: *many published studies equate facial movements with emotional expressions rather than treating their correspondence as a hypothesis to be tested*. For example, facial AU 4, 5, 7 and 23 are referred to as an expression of anger rather than as a lowered brow (AU 4), raised upper lid (AU 5), tightened lid (AU 7) and tightened lip (AU 23). This conflation is sometimes built into an experiment itself, particularly if FACS coding is not used (for a good example, see Calvo et al., 2014). Some studies measure the presence or absence of facial movements with a less systematic coding system called the EMFACS (Emotion FACS), in which coders decide whether a pre-specified group of AUs (i.e., the stipulated expression for each emotion category) are active en masse (i.e., coders identify the presence or absence of the entire configuration), rather than *independently* identify the presence or absence of each facial AU. Hence, EMFACS is less reliable than the FACS [4] because it encourages coders to make mental inferences about the meaning of the muscle movements while they are describing the muscle movements, thereby conflating scientific measurements (which facial muscles moved) with their interpretation (which emotions were expressed).

Approach

To date, no research has systematically or fully cataloged the number of configurations that are biologically possible nor which are routinely made in the wild and with what frequency, but one study makes a start, and also offers a clear example of combining automated-human FACS coding (Srinivasan & Martinez, 2018). Over seven million images were mined from the internet by first identifying all the nouns, verbs, adjectives and adverbs, along with their semantic and lexical relations in the English dictionary that many people think of as emotions.[5] These words were then translated into Spanish, Mandarin Chinese, Farsi, Arabic and Russian. The words were then used to identify and download images of human faces using a variety of online search engines. The configuration of AUs in each image was automatically FACS coded using a computer vision algorithm (Benitez-Quiroz et al., 2016; Benitez-Quiroz et al., 2017b; Benitez-Quiroz et al., 2019). Human FACS coders then manually verified the accuracy of the results provided by the automated analysis for a subset of images.

---

[4] EMFACS identifies a prescribed set of AUs that are thought to express emotion rather than coding the presence or absence of each AU one at a time. Anyone who is trained to use FACS can also use EMFACS; there is no special training. It is important to take note of which coding is used, because EMFACS is less reliable (Rosenberg, Erika. Frequently Asked Questions. erikarosenberg.com/faq) and potentially more prone to bias.

[5] WordNet (Miller, 1995) defines a structure—a graph—that identifies synonyms as well as superordinate and subordinate concepts for each word. A subordinate is a word with a more specific meaning (e.g., *despair* a subordinate of *sadness*). A superordinate is a word with more general meaning (e.g., *emotion* is a superordinate of *sadness*).

Facial Configurations Discovered

Thirty-five configurations of AUs were observed as common to the images mined in all six languages (see Box Table 7-1). This amounts to 22% of the facial configurations that were identified in the seven million images (1.87% of the seven million images contained these 35 configurations). Only eight additional facial configurations were identified that were common in the images mined in one or more, but not all, of these languages.[6]

The results provide the very first attempt at cross-cultural assessment of facial configurations observed in the wild (rather than deliberately posed) and invite a variety of interpretations. One possibility is that these findings support the hypothesis of a small number of facial configurations that are available for emotional expression within and across cultures (although admittedly significantly larger than the six proposed configurations in Figure 4, e.g., Martinez, 2017a). This interpretation is cautioned by several considerations, however.

First, the scientific approach taken by Srinivasan & Martinez (2018) is likely to have missed some facial configurations. Current algorithms only identify 16 of the 24 possible AUs.[7] Human coders only verified the AUs detected by the algorithm, meaning that AUs that were actually present in the images but went undetected by the algorithm were missed entirely. More generally, as illustrated by the EmotioNet Challenge (http://cbcsl.ece.ohio-state.edu/EmotionNetChallenge/), current algorithms do not have high accuracy for detecting facial movements in the wild. In addition, English words were used to mine for images, but the other languages sampled (Spanish, Mandarin Chinese, Farsi, Arabic and Russian) contain

[6] Of these, 2 were observed in a single language, 2 in two languages, 3 in three languages and 1 in four languages.
[7] These are the most frequently observed AUs. Those AUs that are not currently detected are infrequent.

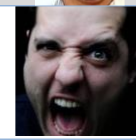
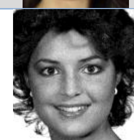
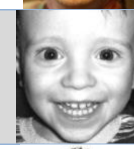
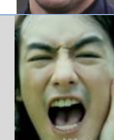
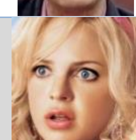
indigenous words for emotions that do not correspond easily to single English words (Smith, 2016), and therefore may be associated with AU configurations that were not sampled in the first place.

Second, the images used by Srinivasan & Martinez (2018) were derived from the internet. Internet images, while better than posed faces, do not substitute for scientific observations of facial movements in the real world. The internet is a curated version of reality. Some common facial configurations are likely missing because they are rarely uploaded to the internet, and some configurations commonly found on the internet may not be commonly observed in the real world.

For these reasons, and also because the cultures sampled have some contact with practices and norms of the U.S., the Srinivasan & Martinez (2018) study does not, on its own, confirm that the 35 identified facial configurations are, in fact, universal. Instead, the study suggest that these facial configurations are commonly used to express instances of emotion across a number of languages in industrialized nations. Furthermore, even if future studies reveal that some or all of these 35 facial configurations are indeed universally made, this does not automatically mean that each is an *innate* expression with a unique emotional meaning; the universality of these expressions may be a result of cultural norms people learn as children (see Box 10).

Box Table 7-1: Thirty-five unique combinations of facial actions observed in cultures whose primary languages are English, Spanish, Mandarin Chinese, Farsi, and Russian (from Srinivasan & Martinez, 2018).

ID	AUs	Examples	ID	AUs	Examples	ID	AUs	Examples

1	4		13	4, 7, 9, 10, 17		25	1, 2, 5, 25, 26	
2	5		14	9, 10, 15		26	4, 7, 9, 25, 26	
3	2, 4		15	12, 25		27	12, 25, 26	
4	4, 7		16	6, 12, 25		28	2, 12, 25, 26	
5	12		17	2, 5, 12, 25		29	5, 12, 25, 26	
6	2, 12		18	1, 2, 5, 12, 25		30	2, 5, 12, 25, 26	
7	5, 12		19	6, 12, 25		31	1, 2, 5, 12, 25, 26	
8	1, 2, 5, 12		20	10, 12, 25, 26		32	6, 12, 25, 26	
9	6, 12		21	1, 2, 25, 26		33	7, 9, 20, 25, 26	
10	4, 15		22	1, 4, 25, 26		34	1, 2, 5, 20, 25, 26	
11	1, 4, 15		23	5, 25, 26		35	1, 4, 5, 20, 25, 26	
12	4, 7, 17		24	2, 5, 25, 26				

Note. ID is the unique identification number given to each facial configuration. An example of each is shown. AUs are the active facial action units that describe each configuration. IDs 1, 2, 10, 11, and 21 were most frequently labeled by participants as expressions of sadness. IDs 5 through 9, 15 through 19, and 26 through 32 were most frequently labeled as expressions of happiness. IDs 20 and 22 through 24 were most frequently labeled as expressions of surprise. IDs 4, 12, 14, 25, and 35 were most frequently labeled as expressions of anger. IDs 33 and 34 were most frequently labeled as expressions of fear. ID 13 was most frequently labeled as an expression of disgust. No consistent labels were offered for ID 3. No configuration exactly matches the AU configurations proposed by Darwin or documented in prior research (for AU comparisons, see Cordaro et al., 2017; Ekman et al., 1968; Matsumoto et al., 2008).

References

Benitez-Quiroz, C. F., Srinivasan, R., & Martinez, A. M. (2016). *Emotionet: An accurate, real-time algorithm for the automatic annotation of a million facial expressions in the wild*. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 5562-5570).

Benitez-Quiroz, F., Srinivasan, R., & Martinez, A. M. (2019). Discriminant Functional Learning of Color Features for the Recognition of Facial Action Units and their Intensities. *IEEE transactions on pattern analysis and machine intelligence*.

Benitez-Quiroz, C. F., Wang, Y., & Martinez, A. M. (2017b). *Recognition of action units in the wild with deep nets and a new global-Local loss*. In Proceedings of the International Conference on Computer Vision.

Jack, R. E., & Schyns, P. G. (2017). Toward a social psychophysics of face communication. *Annual Review of Psychology, 68*, 269-297. doi:10.1146/annurev-psych-010416-044242

Martinez, A. M. (2017b). Visual perception of facial expressions of emotion. *Current Opinion in Psychology, 17*, 27-33.

Srinivasan, R., and Martinez, A.M. (2018). Cross-cultural and cultural-specific production and perception of facial expressions of emotion in the wild. IEEE Transactions on Affective Computing. ArXiv preprint, available at https://arxiv.org/pdf/1808.04399.pdf

Box 8: Meta-Analytic Evidence of Autonomic Nervous System Changes During Emotion

There have been four statistical summaries of scientific findings (called meta-analyses) from experiments designed to measure autonomic nervous system (ANS) changes during emotional episodes (Cacioppo et al., 2000; Siegel et al., 2018; Stemmler, 2004; for a discussion, see Quigley & Barrett, 2014). All of them, including the most comprehensive meta-analysis published to date covering over 200 published experiments involving more than 20,000 participants (Siegel et al., 2018) replicate the same results: ANS changes are neither consistent nor specific for any emotion category. Combining different measurements into a pattern performed no better at distinguishing one emotion category from another than did individual measures (Siegel et al., 2018). Some review articles have tried to make a case for the existence of emotion-specific ANS fingerprints (e.g., Friedman, 2010; Kreibig, 2010), but the meta-analyses are consistent in their findings that different emotion categories cannot be distinguished from one another by changes in heart rate, respiration rate, skin conductance, or any other measure of the autonomic nervous system, alone or in combination; said another way, ANS measures individually, or in combination, are neither consistent nor specific for emotion categories.

The variety in emotion-related ANS changes is consistent with the writings of William James (1890), population views of emotion (Box 2), with Darwin's articulation of *population thinking* in *On the Origin of Species* (1859/2001) (for a discussion, see Barrett, 2017) and with the US physiologist Paul Obrist's findings that peripheral physiological changes are tied to the metabolic demands associated with action (e.g., cardiosomatic coupling; Obrist, Webb, Sutterer, & Howard, 1970) or anticipated action (e.g., supra-metabolic activity; Obrist, 1981; Sterling, 2012; Turner & Carroll, 1985). Because all animals (including humans) behave in a variety of context-sensitive ways, crying, shouting, smiling, freezing and laughing in anger will each be supported by a distinct pattern of ANS change. In this view, ANS variation is not a bug to be explained away as error or designated as epiphenomenal to the nature of emotion. Substantial variation in ANS patterns within an emotion category is a feature that should be expected because it confers evolutionary advantage (for a discussion on how evolution selects for variation in emotion, see Barrett, 2017a).

The variation in emotion categories observed for ANS changes is consistent with the evidence for the brain basis of emotion in humans (Barrett, 2017b; Clark-Polner et al., 2016) and in non-human animals (for a discussion, see Barrett, 2017b; Barrett & Satpute, 2017). For example, even studies of "fear" learning in rodents find evidence of variability in ANS responses and neural circuitry (e.g., Barrett & Finlay, in press; Gross & Canteras, 2012; Iwata & LeDoux, 1988; Tovote et al., 2015). Behavioral experiments also clearly show that the variation within emotion categories is meaningfully tied to context and situational factors and is not merely due to variability in the experimental method. A growing number of studies of emotion are designed to explicitly model and capture heterogeneity within emotion categories both within individuals and across cultures (e.g., Ceulemans, Kuppens, & Van Mechelen, 2012; Gendron, Roberson, van der Vyver, & Barrett, 2014; 2014; Hortensius, Schutter, & Harmon-Jones, 2011; Kuppens, Van Mechelen, & Rijmen, 2008; Kuppens et al., 2007; Nezlek, Vansteelandt, Van Mechelen, & Kuppens, 2008; Stemmler, Aue, & Wacker, 2007; Wilson-Mendenhall, Barrett, & Barsalou, 2013; 2015; Wilson-Mendenhall, Barrett, Simmons, & Barsalou, 2011).

References

Barrett, L. F. (2017a). *How emotions are made: The secret life of the brain*. New York: Houghton-Mifflin-Harcourt.

Barrett, L. F. (2017b). The theory of constructed emotion: An active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience, 12,* 1-23. https://doi.org/10.1093/scan/nsw154

Barrett, L. F. & Finlay, B. L. (in press). Concepts, goals and the control of survival-related behaviors. *Current Opinion in the Behavioral Sciences*.

Barrett, L. F., & Satpute, A. B. (2017). Historical pitfalls and new directions in the neuroscience of emotion. *Neuroscience Letters*.

Cacioppo, J. T., Berntson, G. G., Larsen, J. H., Poehlmann, K. M., & Ito, T. A. (2000). The psychophysiology of emotion. In R. Lewis & J. M. Haviland-Jones (Eds.), *The handbook of emotion* (2 ed., pp. 173-191). New York: Guilford Press.

Ceulemans, E., Kuppens, P., & Mechelen, I. V. (2012). Capturing the Structure of Distinct Types of Individual Differences in the Situation-specific Experience of Emotions: The Case of Anger. *European Journal of Personality, 26*(5), 484-495.

Clark-Polner, E., Johnson, T., & Barrett, L. F. (2017). Multivoxel pattern analysis does not provide evidence to support the existence of basic emotions. *Cerebral Cortex, 27*, 1944-1948.

Darwin, C. (1859/1964). *On the Origin of Species*. (A facsimile of the first edition. ed.). Cambridge, MA.: Harvard University Press

Friedman, B. H. (2010). Feelings and the body: the Jamesian perspective on autonomic specificity of emotion. *Biological Psychology, 84(3)*, 383-393.

Gendron, M., Roberson, D., van der Vyver, J. M., & Barrett, L. F. (2014a). Perceptions of emotion from facial expressions are not culturally universal: evidence from a remote culture. *Emotion, 14(2)*, 251-262. doi:10.1037/a0036052

Gendron, M., Roberson, D., van der Vyver, J. M., & Barrett, L. F. (2014b). Cultural relativity in perceiving emotion from vocalizations. *Psychological Science, 25*, 911-920.

Gross, C. T., & Canteras, N. S. (2012). The many paths to fear. *Nature Reviews Neuroscience, 13(9)*, 651-658.

Hortensius, R., Schutter, D. J., & Harmon-Jones, E. (2012). When anger leads to aggression: induction of relative left frontal cortical activity with transcranial direct current stimulation increases the anger–aggression relationship. *Soc Cogn Affect Neurosci, 7(3)*, 342-347.

Iwata, J., & LeDoux, J. E. (1988). Dissociation of associative and nonassociative concomitants of classical fear conditioning in the freely behaving rat. *Behavioral Neuroscience, 102(1)*, 66-76.

James, W. (1890/2007). *The Principles of Psychology* (Vol. 1). New York: Dover.

Kreibig, S. D. (2010). Autonomic nervous system activity in emotion: a review. *Biological Psychology, 84(3)*, 394-421. doi:10.1016/j.biopsycho.2010.03.010

Kuppens, P., Van Mechelen, I., & Rijmen, F. (2008). Toward disentangling sources of individual differences in appraisal and anger. *Journal of Personality, 76(4)*, 969-1000.

Kuppens, P., Van Mechelen, I., Smits, D. J. M., De Boeck, P., & Ceulemans, E. (2007). Individual differences in patterns of appraisal and anger experience. *Cognition & Emotion, 21(4)*, 689-713. doi:Doi 10.1080/02699930600859219

Nezlek, J. B., Vansteelandt, K., Van Mechelen, I., & Kuppens, P. (2008). Appraisal-emotion relationships in daily life. *Emotion, 8(1)*, 145-150.

Obrist, P. A. (1981). *Cardiovascular psychophysiology: A perspective.* New York: Plenum.

Obrist, P. A., Webb, R. A., Sutterer, J. R., & Howard, J. L. (1970). The cardiac-somatic relationship: some reformulations. *Psychophysiology, 6(5)*, 569-587.

Quigley, K. S., & Barrett, L. F. (2014). Is there reliability and specificity of autonomic changes during emotional episodes? Guidance from the conceptual act theory and psychophysiology. *Biological Psychology, 98*,82-94.

Siegel, E. H., Sands, M. K., Van den Noortgate, W., Condon, P., Chang, Y., Dy, J., Quigley, K. S., & Barrett, L. F. (2018). Emotion fingerprints or emotion populations? A meta-analytic investigation of autonomic features of emotion categories. *Psychological Bulletin, 144(4)*, 343-393.

Stemmler, G. (2004). Physiological processes during emotion. In P. Philippot & R. S. Feldman (Eds.), *The regulation of emotion* (pp. 33–70). Mahwah, NJ: Erlbaum & Associates.

Stemmler, G., Aue, T., & Wacker, J. (2007). Anger and fear: Separable effects of emotion and motivational direction on somatovisceral responses. *International Journal of Psychophysiology, 66(2)*, 141-153.

Sterling, P. (2012). Allostasis: a model of predictive regulation. *Physiology & Behavior, 106(1)*, 5-15.

Tovote, P., Fadok, J. P., & Lüthi, A. (2015). Neuronal circuits for fear and anxiety. *Nature Reviews Neuroscience, 16(6)*, 317-331.

Turner, J. R., & Carroll, D. (1985). Heart rate and oxygen consumption during mental arithmetic, a video game, and graded exercise: Further evidence of metabolically exaggerated cardiac adjustments? *Psychophysiology, 22,* 261–267.

Wilson-Mendenhall, C., Barrett, L. F., & Barsalou, L. W. (2013). Neural evidence that human emotions share core affective properties. *Psychological Science, 24,* 947-956. PMC4015729

Wilson-Mendenhall, C.D., Barrett, L.F., & Barsalou, L.W. (2015). Variety in emotional life: Within-category typicality of emotional experiences is associated with neural activity in large-scale brain networks. *Social Cognitive and Affective Neuroscience*, *10*, 62-71.

Wilson-Mendenhall, C. D., Barrett, L. F., Simmons, W. K., Barsalou, L. W. (2011). Grounding emotion in situated conceptualization. *Neuropsychologia, 49,* 1105-1127.

Box 9: Emotional Episodes and Their Affective Features

In English, the word ‘‘affect’’ means ‘‘to produce a change.’’ To be affected by something is to be influenced by it. In science, and particularly in psychology, ‘‘affect’’ refers to a special kind of influence—something’s ability to influence your mind in a way that is linked to changes in your body. Sometimes “affect” is used as a cautious term, to mean anything emotional. It allows people to refer to emotions in general terms, without specifying exactly what an emotion is or how it should be defined. Sometimes “affect” is used to refer specifically to emotional experiences – to be affected is to feel something (e.g., Panksepp, 1998). In modern psychological usage, ‘‘affect’’ refers to the mental counterpart of internal bodily sensations, whether or not those sensations are associated with emotions. Historically, “affect” referred to simple feelings that are part of every waking moment of your life (Wundt, 1998b/1897; for a discussion, see Barrett & Bliss-Moreau, 2009). This allows us to clarify some persistent confusions that muddle the scientific study of emotions and their expressions.

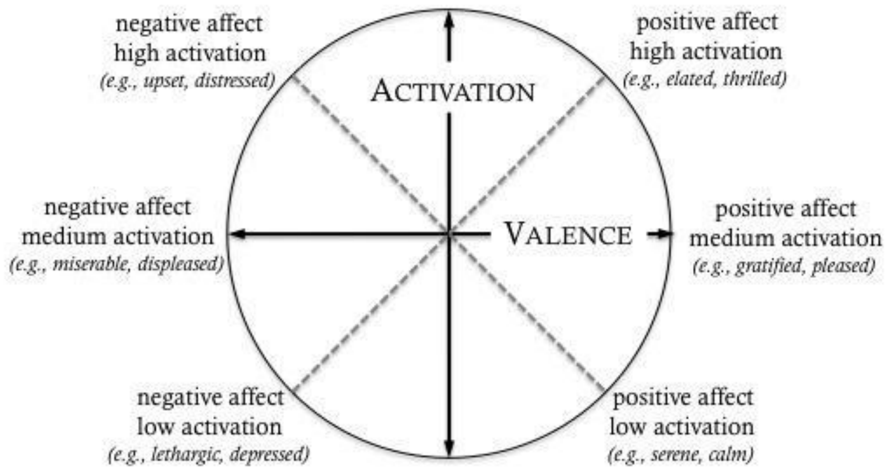
Affective Features

Affect is a property of consciousness (e.g., Edelman & Tononi, 2000; James, 1890; Searle, 1992, 2004) and is not perfectly synonymous with emotion. More than a century of research has revealed that affect, whether part of an emotional episode or otherwise, can be described as a single point in a space with at least two features: valence (pleasure to displeasure), arousal (high to low), although other features are sometimes discussed (interpersonal closeness to distance, dominance to submissiveness, etc.; Russell & Barrett, 1999). Valence and arousal are not ingredients of affect – they are descriptions of feeling in a given moment. They are descriptive features of emotional instances (Box 2), but they are not specific to emotion. Nor are causal processes or mechanisms that cause anything. These two features of affect form a low-dimensional, circular space that describes how a person feels at any moment in time, as in Box Figure 9-1. Valence and arousal are not independent from one another (as one feature changes, so does the other; for a discussion, see Barrett & Bliss-Moreau, 2009; Kuppens et al., 2013).

Thus far, what we understand about the brain-basis of affect (e.g., Lindquist et al., 2016) is consistent with the hypothesis that affective feelings are associated with a wide range of psychological phenomena (reviewed in Barrett & Bliss-Moreau, 2009). Valence and arousal are likely low dimensional representations of internal bodily sensations (referred to as interoception; Craig, 2015) that result from the constantly changing state of the body's internal systems, such as the autonomic nervous system, the immune system, the endocrine system, and so on (referred to as allostasis; Sterling, 2012). Somehow, bodily sensations, which are physical, are transformed into affective feelings, which are mental. Scientists don't yet know how this transformation happens, but many studies suggest that it does. For the moment, it remains one of the mysteries of consciousness.

Emotional Episodes

Affect provides a quick summary of the physiological state of the body, like a barometer, without much detail (Barrett & Bliss-Moreau, 2009). Therefore, affect alone does not indicate what to do next, or how to act, other than to approach something or avoid it (Davidson, 1992; Lang et al., 1993). Many scientists propose that emotional episodes, on the other hand, are specific instances of affect that involve very specific intentions to act. Their specific relation to affect depends on how emotion is defined (various theoretical proposals are presented in Box 2).



Box Figure 9-1. The affective circumplex. Hedonic valence is represented on the horizontal axis and arousal on the vertical axis. Reprinted with permission from Barrett & Bliss-Moreau (2009).

References

Barrett, L. F., & Bliss-Moreau, E. (2009). Affect as a psychological primitive. *Advances in Experimental Social Psychology*, *41*, 167-218.

Craig, A. D. (2015). *How do you feel?: an interoceptive moment with your neurobiological self*. Princeton University Press.

Davidson, R. J. (1992). Anterior cerebral asymmetry and the nature of emotion. Brain Cogn, 20(1), 125-151.

Edelman, G. M., & Tononi, G. (2000). *A universe of consciousness: How matter becomes imagination*. New York: Basic books.

James, W. (1890/2007). *The Principles of Psychology* (Vol. 1). New York: Dover.

Kuppens, P., Tuerlinckx, F., Russell, J. A and Barrett, L. F. (2013). The relationship between valence and arousal in subjective experience. *Psychological Bulletin, 139*, 917-940.

Lang, P.J., Greenwald, M.K., Bradley, M.M., & Hamm, A.O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology, 30*, 261–273.

Lindquist, K. A., Satpute, A. B., Wager, T. D., Weber, J., & Barrett, L. F. (2016). The brain basis of positive and negative affect: Evidence from a meta-analysis of the human neuroimaging literature. *Cerebral Cortex*. 26, 1910-22.

Panksepp, J. (1998). *Affective neuroscience: The foundations of human and animal emotions*. New York: Oxford University Press.

Russell, J. A., & Barrett, L. F. (1999). Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology, 76,* 805-819.

Searle, J. R. (1992). The rediscovery of the mind: MIT press.

Searle, J. R. (2004). Mind: a brief introduction: Oxford University Press.

Sterling, P. (2012). Allostasis: a model of predictive regulation. *Physiology & Behavior*, *106(1)*, 5-15.

Wundt, W. (1998/1897). *Outlines of Psychology* In: Judd, C.H. (Trans.). Leipzig: Wilhelm Engelmann.

Box 10: Learning to Express and Perceive Emotions

Expressing Emotion

Increasingly, scientists have been riveted by discoveries that very young human infants are powerful learners who quickly learn and make use of cues from their environments. When children begin to produce facial movements during emotional episodes with some regularity, they appear to reflect the influences of learning about cultural and social expectations (for review, see Kärtner, Holodynski, & Wörmann, 2013). For example, social smiling is often described as a purely maturational phenomenon, but recent cross-cultural studies suggest its development is influenced by sociocultural factors (see Wörmann, Holodynski, Kärtner, & Keller, 2012). Children are sensitive to variations in the distribution of multimodal sensory changes in their environments, and this variation becomes reflected in the perceptual categories children form for facial movements (Pollak & Kistler, 2002). Infants generalize from small samples to larger populations and gauge their inferences depending upon whether sampling strategies are strong or weak (Denison, Reed & Xu, 2013; Gweon, Tenenbaum & Schulz, 2010). Young children are also able detect complex probabilistic information in their environments, as well as implicitly learn and reproduce the underlying statistical distributions (Plate et al., 2017). These studies suggest that the patterns of facial communication are readily learned by children from their social environments, likely beginning shortly after birth. As noted by Sullivan and Lewis (2003), how specific patterns of facial movements become associated with specific contexts is still unresolved.

There is some evidence that infants learn to express instances of emotion with facial movements as they begin to acquire the emotion concepts of their cultural context. For example, recent studies suggest the facial configuration proposed as the expression of disgust (see Figure

4) is learned in middle childhood as children learn their culture's concept for disgust (Stevenson, Oaten, Case, Repacholi & Wagland, 2010; Widen & Russell, 2013).[8] Further details remain unknown however, due to a gap in the scientific literature. Published studies that carefully describe a child's facial muscle movements during emotional episodes have not yet directly measured individual differences in emotion concept learning, nor variation in parenting or family environment that serves as the context to learn emotion concepts. Studies that have examined the link between how children moves their faces during emotional episodes and how their parents act during emotional situations unfortunately do not precisely quantify the facial muscle movements.

Emotion Perception

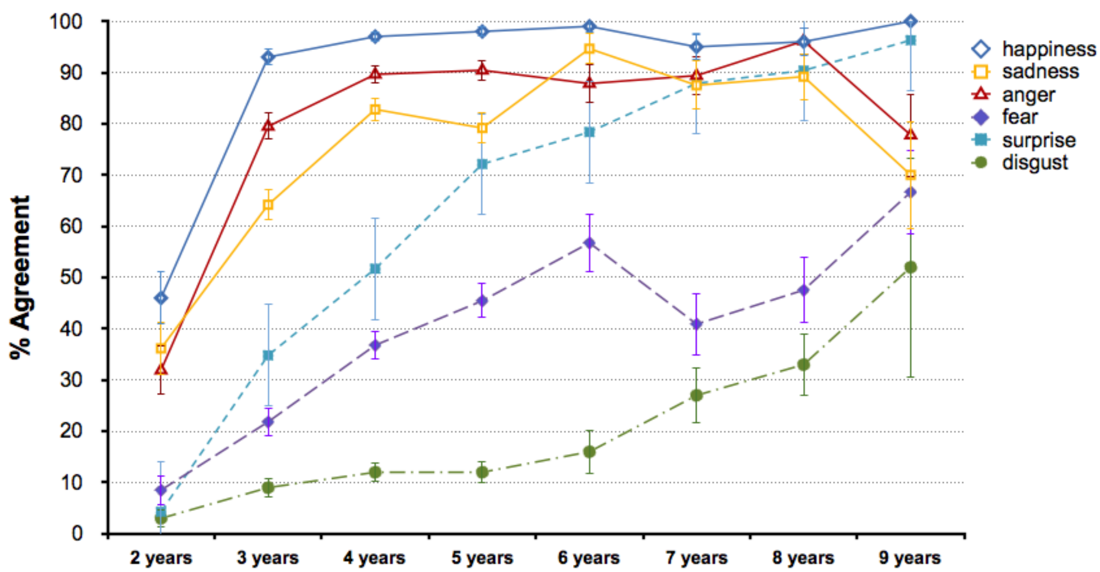
Infants begin by implicitly perceiving broad affective distinctions (e.g., positive-negative; approach-avoid; pleasant-unpleasant) as discussed in the main text of the paper. Additional research is needed to understand the developmental trajectory of infants' implicit emotion perceptions. More is known about the development of children's explicit emotion perception capacities. Numerous experiments now suggest that very young children do not explicitly understand the emotional meaning of facial movements, but that this is a skill they learn over the course of development (Pollak & Kistler, 2002). That is, newborns appear to learn to detect faces and facial movements in the first few days of life, but their capacity to associate these movements with the emotional meaning (i.e., to see facial movements as emotional expressions) is learned during childhood.

---

[8] These studies focus on learning the association between emotion words and facial movements. However, as we describe below, other theorists (e.g., Röttger-Rössler et al. 2015) have proposed that infants begin learning associations between faces and other nonverbal emotion components during early interactions with their caregivers. In this manner, caregivers might shape infants' nonverbal emotion concepts. For example, adults might respond to infants' diffuse responses to emotion eliciting situations by producing specific facial movements and behavioral responses of their own. Thus, probabilistic learning from the environment could result in the infant learning emotion categories, which might be considered "concepts" as well (Pollak, 2009). Still, there is little direct evidence with respect to distinguishing among different negative emotions concepts in infancy.

One careful line of research provides robust, consistent evidence for this hypothesis (reviewed in Widen, 2016). In this research, young children are shown the proposed expressive forms in Figure 4 and are asked to freely label them as emotional expressions by nominating emotion words. Almost twenty experiments now suggest that North American toddlers make broad affective distinctions (e.g., feels good-feels bad) when explicitly labeling the facial configurations, but these perceptions narrow and become more emotionally differentiated (more **emotionally granular**) as children's emotion concepts develop (again, see Widen, 2016). Summarizing across eleven different published studies observing over 1,000 children ranging in age from two to nine years old, it is possible to discern an average developmental trajectory from affect perception to emotion perception (see Box Figure 10-1). Children's perceptions begin to differentiate around the age of two years of age and continue becoming more specific until the middle school years when they look more adult-like (for a review, see Widen, 2016). Even elementary school-aged children commonly label disgust as "anger", and when asked to select disgust from standardized sets of facial configurations, improvement with age is very gradual, with only about 50% of nine year olds offering the word "disgusted" (or a close synonym) for the proposed disgust expression (Widen & Russell 2013). There is growing evidence that

Nonetheless, some scientists argue that more specific aspects of emotion perception are operating much earlier in development than the above studies seem to suggest (e.g., Izard et al., 2010; Leppänen & Nelson, 2009). For example, Grossman (2010) claims that by age three months infants can distinguish proposed expressive configurations for happiness, surprise, and anger, and that by seven months they can discriminate the proposed expressive configurations for fear, sadness, and interest. These claims suggest the interesting possibility that emotion perception might be observed much earlier than two years of age using methods that do not require children to overtly label faces depicting emotional expressions. They are also consistent with other evidence suggesting that infants and young children may implicitly perceive emotion in facial movements when assessed with the habituation task or a perceptual matching task. Left uncertain is whether, beyond discriminating between different facial movements, infants understand the emotional meaning that is typically inferred from these cues. Once again, we see that the choice of stimuli and task design can yield substantially different conclusions.



Box Figure 10-1. Developmental trajectories for emotion concept acquisition. N = 1065. Adapted from Widen (2016) with permission From an early age, children used the expected emotion labels (with standard errors) for the smiling facial configuration ("happiness"), the scowling configuration ("anger"), and the frowning configuration ("sadness") but the expected emotion labels for the other facial configurations gradually increased with age. Data from 11 studies were aggregated (for details, see Widen, 2016). The N for each age group was: two years (N=94), three years (N=229), four years (N=299), five years (N=209), six years (N=74), seven years (N=66), eight years (N=61), and nine years (n=33).

References

Denison, S., Reed, C. & Xu, F. (2013) The emergence of probabilistic reasoning in very young infants: evidence from 4.5- and 6-month-old infants. *Developmental Psychology, 49(2)*, 243-249.

Grossmann, T. (2010). The development of emotion perception in face and voice during infancy. *Restor Neurol Neurosci, 28(2):*219-36

Gweon, H., Tenenbaum, J., & Schulz, L. E. (2010). Infants consider both the sample and the sampling process in inductive generalization. *Proceedings of the National Academy of Sciences, 107(20)*, 9066-9071

Izard, C. E., Woodburn, E. M., & Finlon, K. J. (2010). Extending emotion science to the study of discrete emotions in infants. *Emotion Review, 2(2),* 134-136.

Kärtner, J., Holodynski, M., & Wörmann, V. (2013). Parental ethnotheories, social practice and the culture-specific development of social smiling in infants. *Mind, Culture, and Activity, 20(1)*, 79-95.

Leppänen, J. M., & Nelson, C. A. (2009). Tuning the developing brain to social signals of emotions. *Nature Reviews Neuroscience, 10,* 37-47.

Plate, RC. Fulvio, JM.; Shutts, K; Green, CS; Pollak, SD. (2018). Probability learning: Changes in behavior across time and development. *Child Development, 89*, 205-218.

Pollak, S. D., & Kistler, D. J. (2002). Early experience is associated with the development of categorical representations for facial expressions of emotion. *Proceedings of the National Academy of Sciences, 99(13),* 9072-9076.

Stevenson, R. J., Oaten, M. J., Case, T. I., Repacholi, B. M., & Wagland, P. (2010). Children's response to adult disgust elicitors: Development and acquisition. *Developmental Psychology, 46(1),* 165.

Sullivan, M. W., & Lewis, M. (2003). Contextual determinants of anger and other negative expressions in young infants. *Developmental Psychology, 39(4)*, 693.

Widen, S. C. & Russell, J. A. (2013). Children's recognition of disgust in others. *Psychological Bulletin, 139,* 271-299.

Widen, S. C. (2016). The development of children's concepts of emotion. In L. F. Barrett, M. Lewis, J. M., & Haviland-Jones (Eds.), *Handbook of emotions* (4th ed., pp. 307-318). New York: Guilford Press.

Wörmann, V., Holodynski, M., Kärtner, J., & Keller, H. (2012). A cross-cultural comparison of the development of the social smile: A longitudinal study of maternal and infant imitation in 6-and 12-week-old infants. *Infant Behavior and Development, 35(3)*, 335-347.

Box 11: Research with Virtual Humans

Virtual humans (Rickel et al., 2002) (or Embodied Conversational Agents, Cassell et al., 2000) are software-based artifacts that not only look like people but are also engineered to interact with us using the same verbal and nonverbal behavior that we use to interact with each other. To support interaction, virtual humans are designed with a range of technologies that allow them to interpret a human's behavior, maintain beliefs about the interaction, reason about how to achieve its goals given those beliefs and express itself to achieve those goals. This requires the researcher to engineer virtual humans to simulate perception, beliefs, goals and behavior, bringing together theory and scientific insights spanning traditional scientific disciplines. For example, to model emotion expression and perception in a virtual human, a designer may consider the interactions between the simulations of emotion with those for perception, decision-making, behavior and consequently social interaction (Rickel et al., 2002; Becker-Asano & Wachsmuth, 2008).

In the process of crafting a virtual human, a designer is exploring challenges from an engineering perspective that are closely tied to fundamental psychological questions about the mechanisms and processes that underlie mental events, such as how should a virtual human's internal states such as emotions and communicative intentions map to its expressive behavior in the context of a social interaction? A designer must also grapple with questions about the mechanisms and processes that infer emotional meaning from facial movements (i.e., emotion perception), such as how should a virtual human interpret its perceptions of another's facial movements and what role should its prior beliefs, goals, emotions and the context of the interaction play in that interpretation? Virtual human designs vary considerably in the details of the mechanisms and processes used to perform these tasks. They may seek to simulate how these inference processes work, for example how do prior beliefs bias interpretations. They may use machine learning approaches to acquire a model from human behavioral data. At another extreme, the designer may more simply use a canned or fixed approach that makes specific inferences in specific situations.

There are some unique challenges when considering how to engineer emotion expressions and perceptions, because often the goal is not to faithfully render human behavior as it occurs in everyday life, but to implement them in a virtual human to achieve some other scientific or applied goal. For example, the design of a child's virtual tutor (Lester et al. 1997) may employ exaggerated, unambiguous or supernormal nonverbal behavior to motivate the student. Nonetheless, in other cases, psychological realism in the virtual human's nonverbal behavior, as well as accurate inferences about the human's nonverbal behavior, is often desirable, such as in systems that are designed to train social skills, where inferring the psychological meaning of movements, like facial movements, is crucial (Kron et al., 2017). In some cases, psychological realism is paramount, as when virtual humans are designed to elicit behavior in a human interaction partner that allows the diagnosis of depression (Devault et al, 2015).

Mapping the virtual human's simulated internal states to its movements

Often the realization of expressive behavior in a virtual human involves two connected computational models. There is the model of the emotion elicitation that determines the virtual human's internal, mental states, in essence simulating how the virtual human's experiences elicit the internal emotional states it "feels." In addition, there is a model that maps those states to the behaviors or physical movements such as facial actions designed to express those states. Some studies also endow virtual humans with the capacity to regulate their emotion simulations

(Marsella & Gratch, 2009), allowing the face to dynamically portray episodes of both authentically "felt" emotion and strategically intended emotional expressions.

***Simulating emotional states*.** The design of the computational models of emotion elicitation used in these systems (Dias & Paiva, 2005; Marsella & Gratch, 2009; Becker-Asano & Wachsmuth, 2008) have been most heavily influenced by appraisal theories (See Box 2). Appraisal theories are popular in computational models in virtual human research because each appraisal (e.g., pleasantness, novelty, etc.) maps readily to formalisms used in Artificial Intelligence to model internal states such as beliefs, desires and intentions.[9]

The computational models of emotion elicitation used in virtual humans have not been empirically validated against human data, except for a few notable exceptions. For example, one study compared human self-reports of their changing emotional experience over the course of playing the two-person competitive game of Battleship to the predictions of alternative computational models of emotion (Gratch et al., 2009). A related study (Marsella et al, 2009) evaluated a computational model, EMA, (EMotion and Adaptation, inspired by the appraisal model developed by Lazarus, 1991) that could simulate an emotional episode and regulate the simulation. Again, the game of Battleship was used. In both cases, the subjective reports of emotional experience at key points in the unfolding game were compared to the model's predictions for those experiences (i.e., what the model predicted the real human would feel and how the person would regulate those feelings). In both cases evidence was found supporting aspects of the various models while also identifying some discrepancies between subjects' subjective report and model predictions. For example, EMA correctly predicted the changes in

[9] Interestingly, many of the computational models have been based on appraisal models, such as Ortony, Clore & Collins (1988). However, the computational work has treated it as a causal model, so that the appraisals are the sole cause of emotions, even though the work of Ortony et al was designed as a descriptive model.

subjective reports of emotion intensity, as well as changes in subject's reports of their effort to win and the importance of winning, in the service of regulating their emotions. However, it incorrectly predicted that subjects would engage in wishful thinking, increasing their expectations of eventually winning, in the face of evidence that they were actually losing. In fact, when faced with losing, subjects tended to lower their expectations of winning, perhaps again as a means to down regulate the eventual negative emotions that would result from losing.

*Mapping states to movements***.** Once an emotion simulation is implemented, another model is used to express the simulated emotions in the virtual human's movements, including facial actions. Expression stereotypes (see Figure 4) have frequently been used to map the simulated emotions provided by these computational models to facial actions made by virtual humans. The component process model of emotion (Scherer et al., 2017) has also been explored by tying the virtual human's facial actions to the simulated appraisals generated by its emotion model (Paleari et al, 2007; Malatesta et al, 2009).

Data-driven techniques provide an alternative to theory driven approaches. One approach to using data is to depend on other people's beliefs about which facial movements express emotion, rather than to rely on scientists. For example, posed expressions from actors can be digitally scanned, creating visually realistic faces with posed facial actions (Alexander et al., 2009) that in appearance can be indistinguishable from humans.[10] However, the goal here is typically grounded in perception, as opposed to physical realism. Specifically, the facial movements and their dynamics, as expressions of emotion, are based on beliefs, stereotypes or artistic interpretation, with an underlying expectation that a human user will infer emotional

[10] Digital scanning enables film makers to fool audiences so that, for example, they can bring an actor back to life, twenty years after his death, as was done in the recent Hollywood film, Rogue One, when Peter Cushing (who played a general in the Empire) made an appearance.

meanings from those movements. The main weakness in these approaches is ecological – as we discuss in the main text of the paper, people's beliefs about emotional expressions are often better thought of as stereotypes and may not accurately capture how they actually move their faces when expressing emotion in real life. An alternative data-driven approach, one with greater potential for **ecological validity**, is to record people's facial movements while they are engaged in situations that are assumed to evoke emotion, either in the lab or in everyday life, but this requires inferring which emotional episodes are being created (for a discussion on the difficulties in measuring an emotional episode, see discussion in the main paper).

To date, there is no strong empirical evidence that compares these alternative models of expressing emotions in terms of the psychological realism of the facial movements they generate.

An additional challenge is the mismatch between abstract details provided by a psychological hypothesis and the far more specific engineering details about the appearance of the character's face, skin, as well as realizing realistic dynamics for the individual facial actions required to animate a virtual human's face. Those dynamics include not only timing of each action unit, but also its duration and the rates at which it moves. Furthermore, the dynamics of the action units must influence other elements, such as the bunching and wrinkling of skin. Beyond realizing one configuration of facial movements, there is also sequencing of the movements over time, such as when initial surprise may transition into an instance of fear or anger. Such dynamics encompass not only the facial movements themselves but also the dynamics of the underlying mental states that are being expressed. These factors are all critical to getting a realistic effect. That is, basic emotion theory and the componential model of emotion describe their hypotheses in terms of abstract (non-physical) processes, not the physical mechanisms by which movements or simulations are realized. This requires designers to make inferences as they build computational models. As a consequence, in a given study, it is difficult to infer whether the results pertain to the engineering choices made to realize a psychological theory or more generally to the theory itself.

More broadly, mapping a virtual human's simulated internal states to its movements is, in fact, more complex than simply connecting the emotion simulation (as an output of an emotion elicitation model) to physical movements (as an output of an expression model). Virtual humans are designed to engage us in face-to-face interaction. As in human-to-human social interactions, facial movements in a virtual human-to-human interaction serve a variety of functions, such as establishing rapport by mirroring the human participants' behavior (Huang et al., 2011; Tickle-Degnen & Rosenthal, 1990), emphasizing something with shared attention (Ekman, 1979), regulating turn-taking in the interaction, greetings, communicating attitudes, and so on (Cafaro et al., 2017).

Endowing a virtual human with the capacity to infer emotions from facial movements

Virtual humans often are designed to infer psychological meaning in the nonverbal movement of real humans and use those inferences to inform how they interact. This involves two connected technologies: technologies to sense facial movements, and a model of emotion perception that determines the virtual human's emotion inferences about what those states mean.

In drawing emotional inferences about a human's facial movements, virtual human designs have typically exploited standard technologies that rely on data-driven machine learning techniques discussed earlier. For example, the SimSensei virtual human system was designed to interview patients suffering from depression and post-traumatic stress syndrome (PTSD), as well as nonpatients, asking people questions such as "What is your dream job?", "What do you do now?" and "Tell me about something you did recently that you really enjoyed?" (Devault et al., 2015). A range of sensing technologies were used to measure, for example, smile intensity/duration (OKAO[11]), facial actions & emotion (FACET[12]) and acoustic/vocal qualities[13].

Wörtwein & Scherer (2017) reports on another study that used the SimSensei virtual human. The focus of the study was to discover which questions asked by the virtual human and which nonverbal actions emitted by the participants were most diagnostic of whether or not a given test-subject was suffering from PTSD. Questions were ranked in terms of information gained[14] to discover the most diagnostic nonverbal actions associated with that information. For example, depending on the question, reduced variability of movement (they moved the facial feature less), or reduced average displacement for facial actions AU1, AU2, AU9, AU17, AU18 and AU26 was found to be correlated with PTSD. Interestingly, which facial actions were more diagnostic depended on the question, leading the researchers to suggest that context matters in the sense that the specific context of the question being asked determined what facial actions were most diagnostic of PTSD.

More recent work has begun to explore how a virtual human's model of the human can help bias the inferences of the human's mental states. For example, to the extent the virtual human has a model of the goals and beliefs of the human, it can use a theory of emotion[15] to predict how the human may emotionally react and use those predictions to refine its interpretation of the human's facial movements as well as use the movements to refine its model of the human's beliefs and goals (Alfonso et al., 2015). In addition, the virtual human can

[11] //www.omron.com/r d/coretech/vision/okao.html
[12] http : //www.emotient.com/
[13] http : //www.cogitocorp.com/
[14] See the discussion of information theory in Box 16.
[15] In the two works we cite here (Alfonso et al, 2015; Yongsatianchot & Marsella, 2016), variants of appraisal theories were used to provide the virtual agent with a folk theory of emotion. Alfonso et al (2015) used a model inspired by Smith & Lazarus (1990) and Roseman (2001) while Yongsatianchot & Marsella (2016) used a model inspired by Ortony et al., (1988)

acquire and refine over the course of an interaction how a specific human tends to move his face during a situation thought to evoke emotions (Yongsatianchot & Marsella, 2016).

Should we care about the impact of virtual human technology?

Virtual humans provide an important tool for studying the human perception of emotion in the laboratory. However, they are likely to have an additional, broader social impact. As these systems increasingly play a role in our day-to-day lives, they are likely to have a significant impact on our culture. For example, major corporations such as Microsoft and Amazon are exploring the use of virtual humans in a range of business and home applications, suggesting that virtual humans increasingly will integrate with our everyday life. A valuable lesson is to be learned from the study of how television watching hinders children's ability to infer emotion in facial movements in natural settings (Coats et al., 1999). We might similarly expect the design of virtual humans to impact how we perceive emotion in one another. One might even expect the impact to be more profound and immediate than passive watching of television since virtual humans are designed to engage us in an interaction.

More specifically, virtual humans are being designed to train social skills that have significant real-world consequences, such as medical student's bedside manner (Kron et al, 2017) or a soldier's ability to cross-culturally negotiate (Kim et al, 2009; Traum et al., 2003). These systems either implicitly or explicitly seek to teach how to express emotion and infer emotions from expressions, in highly critical situations such as informing patients that they have a life-threatening disease. As part of the learning process, the system can try to assess and seek to improve the learner's ability to infer emotion from the virtual human's expression as well as to express their emotion in ways the system design deems is appropriate. Clearly, the models of expression production and recognition that are incorporated into the system will impact what is learned.

References

Alexander, O., Rogers, M., Lambeth, W., Chiang, M. and Debevec, P. (2009) Creating a Photoreal Digital Actor: The Digital Emily Project, In IEEE European Conference on Visual Media Production (CVMP).

Alfonso, B., Pynadath, D., Lhommet, M. and Marsella, S. (2015), "Emotional Perception for Updating Agents' Beliefs", in proceedings of the *2015 International Conference on Affective Computing and Intelligent Interaction (ACII)*, IEEE, pp. 201-207.

Becker-Asano, C. and Wachsmuth, I. Affect simulation with primary and secondary emotions. In Proceedings of the Eighth International Conference on Intelligent Virtual Agents (Tokyo). Springer, 2008, 15–28.

Cafaro, A., Pelachaud, C. & Marsella. S. (2017) Nonverbal behavior in multimodal performances, editor Phil Cohen, Springer.

Cassell, J., Sullivan, J., Prevost, S. & Churchill, E., Eds. (2000). Embodied Conversational Agents. Cambridge, MA, MIT Press.

Coats, E. J., Feldman, R. S., & Philippot , P. (1999). The influence of television on children's nonverbal behavior. In P. Philippot & R. S. Feldman(Eds.), The social context of nonverbal behavior (pp. 156–181). New York: Cambridge University Press.

DeVault, D., Artstein, R., Benn, G., Dey, T., Fast, E., Gainer, A., Georgila, K., Gratch, J., Hartholt, A., Lhommet, M., Lucas, GM., Marsella, S., Morbini, F., Nazarian, A., Scherer, S., Stratou, Suri, A., Traum, D., Wood, R., Xu, Y. Rizzo, S. and Morency, LP. (2014) "SimSensei Kiosk: A Virtual Human Interviewer for Healthcare Decision Support", in The 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS).

Dias, J. and Paiva, A. (2005) Feeling and reasoning: A computational model for emotional agents. In Proceedings of the 12th Portuguese Conference on Artificial Intelligence (Covilhã, Portugal). Springer, 127–140.

Ekman, P. (1979). About Brows: emotional and conversational signals. In von Cranach, Foppa, Lepenies & Ploog (eds) Human ethology, Cambridge University Press.

Gratch, J., Marsella. S., Wang, N. and Stankovic, B. (2009) Assessing the validity of appraisal-based models of emotion, in Proceedings of Affective Computing and Intelligent Interaction, ACII, (Amsterdam, Netherlands.

Hoyt, C., Blascovich, J., & Swinth, K. (2003). Social inhibition in immersive virtual environments. *Presence*, *12*(2), 183-195.

Huang, L., Morency, L.P. and Gratch. J. (2011). Virtual rapport 2.0. In Proceedings of the 10th international conference on Intelligent virtual agents (IVA'11), Hannes Högni Vilhjálmsson, Stefan Kopp, Stacy Marsella, and Kristinn R. Thórisson (Eds.). Springer-Verlag, Berlin, Heidelberg, 68-79.

Kim, J., R. Hill, P. Durlach, H. Lane, E. Forbell, M. Core, S. Marsella, D. Pynadath & Hart, J. (2009). "BiLAT: A Game-Based Environment for Practicing Negotiation in a Cultural Context. Special Issue on Ill-Defined Domains", International Journal of AI in Education, vol. 19, no. 3, pp. 289-308

Kron, et al. (2017) Using a computer simulation for teaching communication skills: A blinded multisite mixed methods randomized controlled trial. Patient Education and Counseling 100 748–759.

Lester, J., Converse, S., Kahler, S., Barlow, S.T., Stone, B. & Bhogal, R. (1997). The persona effect: affective impact of animated pedagogical agents. In Proceedings of the ACM SIGCHI Conference on Human factors in computing systems (CHI '97). ACM, New York, NY, USA, 359-366.

Malatesta, L., Raouzaiou, A., Karpouzis, K., Kollias, S.D. (2009). Towards modeling embodied conversational agent character profiles using appraisal theory predictions in expression synthesis. Applied Intelligence. 30(1), 58-64.

Marsella, S., Johnson, W.L., and LaBore, C. 2000. Interactive pedagogical drama. In Proceedings of the fourth international conference on Autonomous agents (AGENTS '00). ACM, New York, NY, USA, 301-308.

Marsella, S. and Gratch, J. , Computational Models of Emotion as Psychological Tools. In Feldman Barrett, Lewis and Haviliand-Jones (eds) Handbook of Emotions, Guilford, 2016.

Marsella, S. and Gratch, J. EMA: A process model of appraisal dynamics. Journal of Cognitive Systems Research 10, 1 (Mar. 2009), 70–90.

Marsella, S., Gratch, J., Wang, N. and Stankovic, B. (2009) Assessing the validity of a computational model of emotional coping, in Proceedings of Affective Computing and Intelligent Interaction, ACII, (Amsterdam, Netherlands).

Ortony, A., Clore, G. & Collins. (1988). The Cognitive Structure of Emotion. Cambridge University Press.

Paleari, M., Grizard, A., Lisetti, C.L. (2007). Adapting psychologically grounded facial emotional expressions to different anthropomorphic embodiment platforms. In: 20th International Florida Artificial Intelligence Research Society Conference (FLAIRS 2007), May 7-9, 2007, Key West, USA.

Rickel, R., Marsella, S., Gratch, J., Hill, R., Traum, D. and Swartout, B. (2002) "Towards a New Generation of Virtual Humans for Interactive Experiences," in IEEE Intelligent Systems July/August, pp. 32-38.

Roseman, I. J. (2001). A model of appraisal in the emotion system: Integrating theory, research, and applications. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Series in affective science. Appraisal processes in emotion: Theory, methods, research* (pp. 68-91). New York, NY, US: Oxford University Press.

Scherer, K. R., Mortillaro, M., & Mehu, M. (2017). Facial expression is driven by appraisal and generates appraisal inferences. In J-M. Fernandez-Dols & J. A. Russell (Eds.), The science of facial expression (p. 353-374). NY: Oxford U Press.

Smith, C. & Lazarus, R. (1990) “Emotion and Adaptation,” in *Handbook of Personality: Theory and Research*, L. A. Pervin, Ed., 1990, pp. 609–637.

Tickle-Degnen, L., & Rosenthal, R. (1990). The Nature of Rapport and Its Nonverbal Correlates. Psychological Inquiry, 1(4), 285-293.

Traum, D., Rickel, J., Gratch, J. & Marsella, S. (2003)."Negotiation over Tasks in Hybrid Human-Agent Teams for Simulation-Based Training.", in Proceedings of the Second International Joint Conference on Autonomous Agents and Multiagent Systems, Australia.

Wörtwein, T. & Scherer, S. (2017) What Really Matters - An Information Gain Analysis of Questions and Reactions in Automated PTSD Screenings. Proceedings of Affective Computing and Intelligent Interaction (ACII).

Yongsatianchot, N. & Marsella, S. (2016) "Integrating Model-Based Prediction and facial Expressions in the Perception of Emotion", in International Conference on Artificial General Intelligence, pp. 234-243, Springer.

Box 12: The In-Group Advantage in Emotion Perception

People's perceptions of emotion are in better agreement with scientists' expectations when judging faces from their own culture or heritage, referred to as an **in-group advantage** (Elfenbein, 2017; Elfenbein & Ambady, 2002a, 2002b, 2003), suggesting that people are better able to infer emotional states from configurations of muscle movements when they are more familiar with the structural features of the faces they are asked to label (Neth & Martinez, 2010). This suggests that perceptual learning may play an important role in perceiving emotion. Perceptual learning may be necessary for making sense of even the most basic visual details in photographs. For example, it is only when we are familiar with other people that we realize that different photos of the same person are, in fact, the same person (i.e., perceptual learning and familiarity are necessary to categorize faces by identity; Beale & Keil, 1995; Jenkins, White, Monfort, & Burton, 2011; McKone, Martini & Nakayama, 2001; Viviani, Binda & Bosato, 2007; for a discussion of how familiarity is important for face perception, see Young & Burton, 2017).

References

Beale, J. M., & Keil, F. C., (1995). Categorical effects in the perception of faces. Cognition, 57, 217-239.

Elfenbein, H. A. (2017). Emotional dialects in the language of emotion. In J-M. Fernandez-Dols & J. A. Russell (Eds.), The science of facial expression (p. 479-496). NY: Oxford U Press.

Elfenbein, H. A., & Ambady, N. (2002). Is there an in-group advantage in emotion recognition? Psychol Bull, 128(2), 243-249.

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: a meta-analysis. Psychological Bulletin, 128(2), 203-235.

Elfenbein, H. A., & Ambady, N. (2003). Universals and cultural differences in recognizing emotions. Current Directions in Psychological Science, 12(5), 159-164.

Jenkins, R., White, D., Van Montfort, X., & Mike Burton, A. (2011). Variability in photos of the same face. Cognition, 121(3), 313-323. doi:10.1016/j.cognition.2011.08.001

McKone, E., Martini, P., & Nakayama, K. (2001). Categorical perception of face identity in noise isolates configural processing. Journal of Experimental Psychology: Human Perception and Performance, 27(3), 573-599

Neth, D., & Martinez, A. M. (2010). A computational shape-based model of anger and sadness justifies a configural representation of faces. Vision research, 50(17), 1693-1711.

Viviani, P., Binda, P., & Borsato, T. (2007). Categorical perception of newly learned faces. Visual Cognition, 15(4), 420-467.

Young, A. W., & Burton, A. M. (2018). Are We Face Experts? Trends Cogn Sci, 22(2), 100-110. doi:10.1016/j.tics.2017.11.007

Box 13: Some Details of the Emotion Perception Studies by Crivelli and Colleagues

Crivelli and colleagues have published three articles examining emotion perception in small-scale, remote cultural contexts that are relevant to this paper. If we only pay attention to the *reliability* with which participants infer an emotional cause for the facial configurations in Figure 4, then it appears that several of these experiments provide moderate support for the commonsense hypothesis that anger, disgust, fear, happiness, sadness and surprise each are expressed with the facial configurations in Figure 4. But a closer examination of the specificity, in addition to reliability, makes it clear that they do not.

1. Crivelli, C., Jarillo, S., Russell, J. A., & Fernández-Dols, J. M. (2016). Reading emotions from faces in two indigenous societies. *Journal of Experimental Psychology: General, 145(7)*, 830-843.

In Study 1 of Crivelli, Jarillo, et al. (2016), photographs of six posed facial configurations similar to those in Figure 4 of the main paper were spread before participants from the Trobriand Islands; participants were asked to point to the person who felt a specific emotion (so, this was a choice-from-array task: matching facial configurations to words).

	Reliability	**Specificity**
Smiling	Labeled as happiness (.58 [.39, .74])	Labeled as anger (.20 [.09, .38])
Pouting	Labeled as sadness (.46 [.29, .65])	Labeled as fear (.27 [.13, .46])
Gasping	Labeled as fear (.31 [.16, .50])	Labeled as anger (.30), disgust (.29)
Nose-scrunch	Labeled as disgust (.25)	Labeled as fear (.27), sad (.23), anger (.20)
Scowling	Labeled as anger (.07)	Labeled as disgust (.08)

Results from Table 1 of Crivelli et al., Study 1. Facial configurations are listed in the first column. 95% confidence intervals (CIs) presented in square brackets.

There is only above chance reliability and specificity for smiling as an expression of happiness. This result has an untested alternative explanation: participants were able to pick the correct label based on valence alone (smiling faces depict a pleasant state and the only pleasant word offered was happiness). For a discussion, see main text.

In Crivelli, Jarillo et al. (2016) Study 2, the method was the same (matching faces to words), except that some participants saw posed static photographs of facial configurations and some saw posed dynamic videos. Participants were the Mwani of Mozambique and the experiment was conducted in their native language (Kimwani). There was no difference between the static and dynamic conditions. Only the static condition is discussed here for brevity.

	Reliability	**Specificity**
Smiling	Labeled as happiness (.58)	Labeled as sadness (.21)
Pouting	Labeled as sadness (.16)	Labeled as anger (.16), fear (.16), disgust (.32)
Gasping	Labeled as fear (.58)	Labeled as anger (.21)
Nose-scrunch	Labeled as disgust (.37)	Labeled as sadness (.16), anger (.32), fear (.11)
Scowling	Labeled as anger (.26)	Labeled as happiness (.11), sadness (.16), fear (.11), disgust (.11)

Results from Table 5, of Crivelli et al., Study 2. Facial configurations are listed in the first column. The Specificity column lists results that are not statistically different from those in the Reliability column.

No facial configuration was labeled specifically as hypothesized. The results for the dynamic facial configurations was similar, except that the smiling configuration was specifically labeled as happiness, which is tempered by an alternative explanation (perceiving valence).

2. Crivelli, C., Russell, J. A., Jarillo, S., & Fernández-Dols, J. M. (2016). The fear gasping face as a threat display in a Melanesian society. *Proceedings of the National Academy of Sciences*, 113(44), 12403-12407.

In this study, Trobriand Islanders were asked to point to the posed facial configuration of a person who felt a specific emotion (in the emotion condition) or who was communicating a specific social motive (in the social motive condition). Participants could select a facial configuration from an array of six, pick a card with a black cross meaning “face not present in the array,” or answer that they did not know the response.

	Reliability	**Specificity**
Smiling	Labeled as happiness (1.0)	
Pouting	Labeled as sadness (.53 [.15, .57])	Labeled as hunger (.31), fear (.25), submission (.25)
Gasping	Labeled as fear (.39)	Labeled as anger (.56), threat (.69)
Nose-scrunch	Labeled as disgust (.22)	Labeled as fear (.28), rejection (.56)
Scowling	Labeled as anger (.06)	Labeled as disgust (.36), hunger (.14), help (.19), submission (.14), rejection (.22), about to eat (.19)

Results from Table S1, of Crivelli et al. Facial configurations are listed in the first column. The Specificity column lists results that are not statistically different from those in the Reliability column.

There was strong reliability and above chance specificity for smiling as an expression of happiness, which is tempered by an alternative explanation (perception of valence).

3. Crivelli, C., Russell, J. A., Jarillo, S., & Fernández-Dols, J. M. (2017). Recognizing spontaneous facial expressions of emotion in a small scale society of Papua New Guinea, Emotion, 17(2), 337-347.

In Study 1, Trobriand Islanders were asked to freely label spontaneous expressions of happiness, sadness, anger, surprise and disgust produced by the Fore people of Papua New Guinea; the photographs were labeled by and published by Ekman (1980).

	Reliability	**Specificity**
Smiling	Labeled as happiness (.13 [.04, .29])	Labeled as laughing or smiling (.44)
Pouting	Labeled as sadness (.16 [.06, .32]	Labeled as avoidance (.19)
Startled	Labeled as surprise (.00)	Labeled as avoidance (.19), sadness (.16)
Nose-scrunch	Labeled as disgust (.06 [.01, .21]),	Labeled as avoidance (.22)
Scowling	Labeled as anger (.03 [.00, .17])	Labeled as avoidance (.56)

Results from Table 1, of Crivelli et al., Study 1. Facial configurations are listed in the first column. The Specificity column lists results that are not statistically different from those in the Reliability column.

No spontaneous expression was labeled as predicted by Ekman (1980) or by the commonsense view.

In Study 2 of Crivelli at al. (2017), Trobriand Islanders matched posed facial configurations to emotion labels that were provided by the experimenter (i.e., they performed the matching faces to words method that was used in Crivelli, Jarillo et al., 2016). According to the paper, "The response format consisted of nine written terms. Five of the labels were predicted by Ekman (1980): *mwasawa* (happiness), *ninamwau* (sadness), *leya* (anger), *eyowa lopola* (surprise, startle), and *minena* (disgust). Two of the labels were Study 1's modal categories: *gigila* (laughing, smiling) and *gibulwa* (feels like avoiding social interaction). And, two of the remaining labels were *itwali* (other emotion) and *gala anukwali* (I do not know). On the actual questionnaire, only the Kilivila terms were listed. The items were always presented in the same order: *gilbuwa*, *ninamwau/mwau*, *minena*, *eyowa lopola*, *gigila*, *leya*, *mwasawa*, *itwali*, and *gala anukwali*."

	Reliability	**Specificity**
Smiling	Labeled as happiness (.17 [.06, .37])	Labeled as laughing, smiling (.69), sadness (.13)
Pouting	Labeled as sadness (.29 [.15, .49])	Labeled as anger (.17), surprise (.17)
Startled	Labeled as surprise (.21 [.09, .41])	Labeled as sadness (.21), happiness (.13), avoidance (.17)
Nose-scrunch	Labeled as disgust (.38 [.21, .57])	Labeled as avoidance (.33)
Scowling	Labeled as anger (.13 [.04, .32])	Labeled as sadness (.29), avoidance (.50)

Results from Table 3 of Crivelli et al., Study 2. Facial configurations are listed in the first column. The Specificity column lists results that are not statistically different from those in the Reliability column.

No spontaneous expression was labeled as predicted by the commonsense view. The smiling configuration was consistently chosen as the expression for "happy" just barely above chance, it was also labeled sadness as frequently. And the smiling configuration was also labeled as laughter more reliably than it was labeled "happy," replicating Study 1. Nor did a pouting configuration meet the specificity criterion; proportionally, participants were just as likely to label a pouting face as anger or surprise. The startled configuration (proposed as the surprise expression) was as reliably labeled as sadness, happiness or "feels like avoiding a situation" as it was labeled surprise. A nose-wrinkled configuration (the proposed disgust expression) was consistently labeled as "disgust" better than what would be expected by chance, but it was not specifically (uniquely) labeled as “disgust,” as it was just as often labeled as “feels like avoiding a social interaction.”

Box 14: The Power of Words in Emotion Perception Experiments

The emotion words provided during choice-from-array tasks may have a psychological impact that extends beyond just constraining participants' word choice – they may actually help to create reliable emotion inferences.

Words Support Perception

In people who already possess conceptual knowledge for U.S. concepts of anger, fear, and so on, emotion words appear to encourage participants to see certain emotions in the facial configurations of Figure 4 more so than they would otherwise. Many experiments now show that words have a more basic function in supporting perception, even for unfamiliar objects (Lupyan, Rakison & McClelland, 2007), contradicting the widespread assumption that people simply learn names for categories they already know, and supporting the hypothesis that words shape how categories are learned in the first place (Gelman & Roberts, 2017; Waxman & Gelman, 2010). Consistent with these broader findings, emotion words have been shown to shape how participants perceive and even literally *see* faces (Gendron et al., 2012) because they influence how people encode and remember facial features (Fugate et al., 2010, 2017; Doyle & Lindquist, 2018). Additional evidence also shows that the conceptual knowledge linked to words dynamically shapes the perception of facial configurations: when participants believe that two emotion categories are conceptually more similar to one another, facial configurations depicting those categories were also perceived as more similar, even when controlling for the actual perceptual similarity of the facial configurations (using novel research methods like reverse correlation and computer mouse tracking; Brooks & Freeman, 2018).

Words Invite Concept Learning

Research shows that the words provided in choice-from-array tasks may actually quickly *teach* participants the expected answers in an experiment. Participants in an experiment might see anywhere from a dozen to a hundred proposed expressive configurations, and for each one they see, the same handful of emotion words are presented over and over again. Under these conditions, participants quickly learn which words are supposed to correspond to each facial configuration. For example, children learn to label an artificially constructed facial expression (e.g., a blowfish expression) with the word "pax" in a choice from array task at levels that are comparable to the proposed expressive configurations in Figure 4 (Nelson & Russell, 2016). Participants also use a process of elimination strategy: words that are not chosen on prior trials are selected more frequently, inflating agreement levels (DiGirolamo & Russell, 2017).

There is some evidence that choice-from-array tasks inadvertently allow participants to learn emotion categories during the course of an experiment. In a recent study, emotion categories that are untranslatable with a single word in English, in Mandarin Chinese and in Hadza culture (and that do not exist in those cultures) were presented to participants from those cultures in a choice-from-array task along with contrived cues (in this case, made up vocalizations). Participants free-labeling of the vocalizations indicated that they were unfamiliar with them; they did not label the vocalizations with words for the novel emotion concepts or with words for anger, disgust, fear, and so on. Nonetheless, participants labeled the vocalizations with reliability and specificity when they were offered the novel category words in a choice-from-array task, making those emotion categories and their (completely made up) vocalizations appear universal (Hoemann, Crittenden, Ruark, Gendron, & Barrett, in press; also see Gendron et al., 2015).

Emotion Words, Emotion Concepts and Emotion Perception

Developmental evidence is also consistent with the hypothesis that emotion words and associated conceptual knowledge play a powerful, and perhaps even necessary, role in emotion perception. A careful line of research provides robust, replicable evidence that children implicitly learn the affective meaning of facial movements in infancy, but only learn to explicitly infer an emotional meaning for facial configurations when they acquire the relevant emotion concept (see Box 10). These studies also suggest that emotion words play an important role in the development of emotion perception during early and middle childhood. For example, children between the ages of three and ten find it easier to match an expressive stereotype to an emotion word than to another example of the same stereotype (i.e., children find it easier to match the word “angry” to a scowling face than to perceptually match two scowling faces; see Widen, 2016). In a story-telling task, three and four-year-old children find it harder to state the cause of an expressive stereotype (e.g., a scowling face) than for an emotion word (e.g., angry) or a corresponding behavior (e.g., a scream) (Widen & Russell, 2004). This **label superiority effect** is robust and is observed in a variety of experiments (Balconi & Carrera, 2007; Camras & Allison, 1985; Reichenbach & Masters, 1983; Russell & Widen, 2002a, 2002b; Widen & Russell, 2002, 2004, 2010a, 2010b). Children between the ages of four and ten years of age find it more difficult to freely label an expression stereotype like those in Figure 4 than brief stories describing anger, fear, surprise, disgust, compassion, embarrassment, shame and contempt (where the stories do not contain any emotion label; Widen & Russell, 2010a); for example, children are more likely to freely offer the word “disgusted” (or a synonym) to label a story describing disgust than the stereotyped disgust expression (e.g., see also Camras & Allison, 1985 N. Eisenberg, murphy, shepherd, 1997; for a review see Widen & Russell, 2013a). This phenomenon has been called a **face inferiority effect**. It suggests that the ability to infer emotions emerges later for facial movements than for stories, an effect that could result from a number of factors.

Figure S8-1 may reflect more about children’s’ ability to incorporate facial movements into their emotion concepts, rather than their acquisition of those concepts. This interpretation is suggested by the developmental trajectory of the human visual system during these same years (Mondloch et al., 2003). Adults process the configurations in Figure 4 configurally (Martinez, 2017a; Neth & Martinez, 2009), meaning that they perceive second-order dependencies in image features (first-order dependencies involve the ordering of features (e.g., eyes on top of the nose on top of the mouth); second-order dependencies involve the relative distances between these features). There is some debate over when children become proficient in the visual interpretation of configural features (some evidence suggests at birth (Turati et al., 2010) whereas other evidence suggests around eight years of age (Mondloch et al., 2003; Le Grand et al., 2004; Maurer et al., 2002; Sinha et al., 2006)).[16] This may help to explain, at least in part, why children increasingly improve in their ability to match words to faces as they age. It may also help explain why toddlers (and even older children) have difficulty perceptually matching photos of different people who are posing the proposed expressive configurations for the same emotion category (e.g., different people scowling, different people frowning, etc.; Widen, 2016).[17]

Despite these ambiguities, it is possible that infants begin to learn emotion concepts and infer emotional meaning in facial movements earlier than they can explicitly label those movements with emotion words. One (as yet untested) hypothesis proposes that early in their development, infants hear emotion words being used by their parents and caregivers, and these emotion words serve to scaffold the ability of infants to begin to form emotion categories and learn emotion concepts (this idea is thought to operate more generally for other abstract categories and concepts; see Barrett, 2017a; Barrett et al., 2007; Lindquist & Gendron, 2013).

Words can initiate and scaffold the formation of concepts and categories (Balaban & Waxman, 1997; Waxman & Markow, 1995), particularly when the instances of a concept vary in how they look, sound and feel (as is the case with emotion concepts). These are called **abstract** or **artifact categories**. A growing body of research shows that infants and toddlers use words as

[16] Experience with faces allows children to more quickly learn how to interpret the meaning of facial movements (Oakes & Ellis, 2013); scientists speculate that this experience allows children to differentiate facial information into categories that are functional in their social environments.

[17] It has been suggested that children have a difficult time perceptually matching faces because they are limited in their ability to process faces configurally (Pascalis et al., 2002). Configural processing is necessary to gain expertise in face recognition (Maurer et al., 2002) and plays an important role in the visual perception of emotion (Neth & Martinez, 2009).

a powerful tool for learning artifact or nominal kind categories (i.e., objects are treated as similar for performing some function when they are perceptually dissimilar; e.g., Fulkerson & Waxman, 2007; Landau & Shipley, 2001; Plunkett et al. 2008; Addyman & Mareschal, 2010; Althaus & Westerman, 2016; Baldwin & Markman, 1989; Dewar & Xu, 2007, 2009; Welder & Graham, 2006). It has been proposed that emotion words help children form emotion categories and learn emotion concepts precisely because emotion concepts are abstract concepts (Barrett, 2017; Barrett et al., 2007; Lindquist & Gendron, 2013).

If variation is the norm when it comes to emotion categories, then emotion categories are abstract categories: people can tremble, jump, freeze, scream, gasp, hide, attack, and even laugh in the face of fear; the same appears to be true for anger. Physiological changes such as heart rate, breathing rate or blood pressure increase, decrease, or stay the same across instances of all emotion categories that have been studied (see Box 9). The variation is not random – it is situated -- and it is beyond what can be accounted for by common beliefs about emotion (see Box 2). The fact that the instances of an emotion category, like fear, can vary considerably in their facial movements, their physical changes, and their behaviors implies that when it comes to learning emotion concepts, children are faced with the task of making inferences about deeper commonalities across perceptually variable instances (i.e., they must learn a nominal kind category). That is, they must learn that instances of fear belong to the same category because they serve the same purpose, even if those instances look, sound, and feel different.

Evidence from Congenitally Deaf Children

Another opportunity to study the role of emotion words and their associated concepts in emotion perception comes from observing children who are born deaf. Congenitally deaf children are often born to hearing parents who do not know sign language and who may subsequently struggle to learn it, reducing crucial opportunities to communicate with their infants early in life. Communication with parents and caregivers is the basis of emotion concept learning (Harris, de Rosnay, & Pons, 2016) and language learning (Kuhl, 2014) more generally, suggesting that congenitally deaf children who are born to non-signing, hearing parents might be slower to learn mental words and concepts. This may offer a window to observe the effects of early exposure to language, or its delay, on emotion perception competency. Numerous studies now show that congenitally deaf children who are born to non-signing, hearing parents have no opportunities to engage in or hear conversations about emotion or benefit from emotion labeling. These children are, in fact, slower to learn mental words and concepts because of this lack of access (e.g., Levrez et al., 2012; Remmel & Peters, 2008; Russell et al., 1998; Schick et al., 2007; Steeds, Rowe, & Dowker, 1997). A variety of studies do, in fact, suggest that deaf children have a difficult time inferring emotions from scowls, frowns, smiles and so on, when compared to hearing children or children who are raised by parents who are fluent in sign language (for a review of evidence, see Sidera et al., 2017). This difficulty extends to inferring mental causes for physical movements, more generally (e.g., Ludlow et al., 2012).

Some scientists working in this area hypothesize that deaf children's difficulty inferring emotion from facial movements is an example of their larger difficulties inferring mental events in general, primarily due to a delay in their ability to learn language (e.g., Dyck & Denver, 2003; Ludlow et al., 2010; Schick et al., 2007; Spencer & Marschark, 2010; Walker-Andrews & Lennon, 1991). Hearing children who are delayed in learning language also have difficulties inferring emotion from facial movements (Nelson et al., 2011). Taken together, then, this research is consistent with the hypothesis that emotion perception competency emerges in the context of word and concept learning.

Summary

Taken together, these findings suggest that emotion words are not psychologically inert – they may shape how emotion is inferred in facial movements and encourage participants to assign emotional meaning to facial configurations differently than they would if the words were not present. Since most of the studies that support the common view of emotion perception are choice-from-array tasks that include emotion words, the potency of those words provides an alternative explanation for the hundreds of studies that seem to strongly support the hypothesis that people perceive specific emotions in specific facial configurations with reliability.

References

Addyman, C., & Mareschal, D. (2010). The perceptual origins of the abstract same/different concept in human infants. Animal cognition, 13(6), 817-833.

Althaus, N., & Westermann, G. (2016). Labels constructively shape object categories in 10-month-old infants. Journal of experimental child psychology, 151, 5-17.

Balaban, M. T., & Waxman, S. R. (1997). Do words facilitate object categorization in 9-month-old infants? Journal of experimental child psychology, 64(1), 3-26.

Balconi, M., & Carrera, A. (2007). Emotional representation in facial expression and script: A comparison between normal and autistic children. Research in developmental disabilities, 28(4), 409-422.

Baldwin, D. A., & Markman, E. M. (1989). Establishing word-object relations: A first step. Child development, 381-398.

Barrett, L. F., Lindquist, K. A., & Gendron, M. (2007). Language as context for the perception of emotion. Trends in cognitive sciences, 11(8), 327-332.

Barrett, L. F. (2017a). How emotions are made: The secret life of the brain. New York: Houghton-Mifflin-Harcourt.

Brooks, J. A., & Freeman, J. B. (2018). Conceptual knowledge predicts the representational structure of facial emotion perception. *Nature Human Behavior*, https://doi.org/10.1038/s41562-018-0376-6

Camras, L. A., & Allison, K. (1985). Children's understanding of emotional facial expressions and verbal labels. Journal of nonverbal Behavior, 9(2), 84-94.

Dewar, K., & Xu, F. (2009). Do early nouns refer to kinds or distinct shapes? Evidence from 10-month-old infants. Psychological Science, 20(2), 252-257.

DiGirolamo, M. A., & Russell, J. A. (2017). The emotion seen in a face can be a methodological artifact: The process of elimination hypothesis. Emotion, 17(3), 538-546. doi:10.1037/emo0000247

Doyle, C. M., & Lindquist, K. A. (2018). When a word is worth a thousand pictures: Language shapes perceptual memory for emotion. J Exp Psychol Gen, 147(1), 62-73. doi:10.1037/xge0000361

Dyck, M. J., & Denver, E. (2003). Can the Emotion Recognition Ability of Deaf Children be Enhanced? A Pilot Study. J Deaf Stud Deaf Educ, 8(3), 348-356. doi:10.1093/deafed/eng019

Eisenberg, N., Murphy, B. C., & Shepard, S. (1997). The development of empathic accuracy. Empathic accuracy, 73-116.

Fugate, J. M. B., Gendron, M., Nakashima, S., & Barrett, L. F. (2017). Emotion words: Adding face value. Emotion

Fugate, J., Gouzoules, H., & Barrett, L. F. (2010). Reading chimpanzee faces: evidence for the role of verbal labels in categorical perception of emotion. Emotion, 10(4), 544-554.

Fulkerson, A. L., & Waxman, S. R. (2007). Words (but not tones) facilitate object categorization: Evidence from 6- and 12-month-olds. Cognition, 105(1), 218-228.

Gelman, S. A., & Roberts, S. O. (2017). How language shapes the cultural inheritance of categories. *Proceedings of the National Academy of Sciences*, *114*(30), 7900-7907.

Gendron, M., Lindquist, K., Barsalou, L., & Barrett, L. F. (2012). Emotion words shape emotion percepts. Emotion, 12, 314-325.

Gendron, M., Roberson, D., & Barrett, L. F (2015). Cultural variation in emotion perception is real: A response to Sauter et al. Psychological Science, 26, 357-359.

Harris, P. L., de Rosnay, M., & Pons, F. (2016). Understanding emotion. In L. F. Barrett, M. Lewis, & J. M. Haviland-Jones (Eds.), Handbook of Emotion, 4th Edition. New York: Guilford.

Hoemann, K., Crittenden, A. N., Ruark, G., Gendron, M., & Barrett, L. F. (in press). Context facilitates the cross-cultural perception of emotion. *Emotion*.

Hoemann, K., Xu, Fei, & Barrett, L. F. (in press). Emotion words, emotion concepts, and emotional development in children: A constructionist hypothesis. *Developmental Psychology*.

Kuhl, P. K. (2014). Early language learning and the social brain. Cold Spring Harbor symposia on quantitative biology, 79, 211-220.

Landau, B., & Shipley, E. (2001). Labelling patterns and object naming. Developmental science, 4(1), 109-118.

Le Grand, R. L., Mondloch, C. J., Maurer, D., & Brent, H. P. (2004). Impairment in holistic face processing following early visual deprivation. Psychological Science, 15(11), 762-768.

Levrez, C., Bourdin, B., Le Driant, B., & Baudouin, F. D. (2012). The impact of verbal capacity on theory of mind in deaf and hard of hearing children. American Annals of the Deaf, 157, 66-77.

Lindquist, K.A. & Gendron, M. (2013). What's in a word? Language Constructs emotion perception. Emotion Review, 5, 66-71.

Ludlow, A. K., Heaton, P., & Deruelle, C. (2013). Decoding Actions and Emotions in Deaf Children: Evidence From a Biological Motion Task. Journal of Cognition and Development, 14(4), 561-572. doi:10.1080/15248372.2012.698434

Ludlow, A., Heaton, P., Rosset, D., Hills, P., & Deruelle, C. (2010). Emotion recognition in children with profound and severe deafness: do they have a deficit in perceptual processing? J Clin Exp Neuropsychol, 32(9), 923-928. doi:10.1080/13803391003596447

Lupyan, G., Rakison D. H., & McClelland, J. L. (2007). Language is not just for talking: Redundant labels facilitate learning of novel categories. Psychological Science, 18, 1077-1083.

Martinez, A. M. (2017a). Computational models of face perception. Current Directions in Psychological Science, 26(3), 263-269.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. Trends in cognitive sciences, 6(6), 255-260.

Mondloch, C. J., Geldart, S., Maurer, D., & Le Grand, R. (2003). Developmental changes in face processing skills. Journal of experimental child psychology, 86(1), 67-84.

Nelson, N. L., & Russell, J. A. (2016). Building emotion categories: Children use a process of elimination when they encounter novel expressions. J Exp Child Psychol, 151, 120-130. doi:10.1016/j.jecp.2016.02.012

Nelson, K. E., Welsh, J. A., Trup, E. M. V., & Greenberg, M. T. (2011). Language delays of impoverished preschool children in relation to early academic and emotion recognition skills. First Language, 31, 164–194.

Neth, D., & Martinez, A. M. (2009). Emotion perception in emotionless face images suggests a norm-based representation. Journal of vision, 9(1), 5.

Oakes, L. M., & Ellis, A. E. (2013). An eye-tracking investigation of developmental changes in infants' exploration of upright and inverted human faces. Infancy, 18(1), 134-148.

Pascalis, O., de Haan, M., & Nelson, C. A. (2002). Is face processing species-specific during the first year of life?. Science, 296(5571), 1321-1323.

Plunkett, K., Hu, J. F., & Cohen, L. B. (2008). Labels can override perceptual categories in early infancy. Cognition, 106(2), 665-681.

Reichenbach, L., & Masters, J. C. (1983). Children's use of expressive and contextual cues in judgments of emotion. Child Development, 993-1004.

Remmel, E., & Peters, K. (2008). Theory of mind and language in children with cochlear implants. Journal of Deaf Studies and Deaf Education, 14, 218–236.

Russell, J. A., & Widen, S. C. (2002a). A label superiority effect in children's categorization of facial expressions. Social Development, 11(1), 30-52.

Russell, J. A., & Widen, S. C. (2002b). Words versus faces in evoking preschool children's knowledge of the causes of emotions. International Journal of Behavioral Development, 26(2), 97-103.

Russell, J. A., Hosie, C. D., Gray, C., Scott, C., Hunter, N., Banks, J. S., & Macaualy, M. C. (1998). The development of theory of mind in deaf children. Journal of Child Psychology and Psychiatry, 39, 903–910.

Schick, B., de Villiers, P., de Villiers, J., & Hoffmeister, R. (2007). Language and theory of mind: A study of deaf children. Child Development, 78, 376–396.

Sidera, F., Amado, A., & Martinez, L. (2017). Influences on facial Emotion Recognition in Deaf Children. J Deaf Stud Deaf Educ, 22(2), 164-177. doi:10.1093/deafed/enw072

Sinha, P., Balas, B., Ostrovsky, Y., & Russell, R. (2006). Face recognition by humans: Nineteen results all computer vision researchers should know about. Proceedings of the IEEE, 94(11), 1948-1962.

Spencer, P. E., & Marschark, M. (2010) Evidence-based practice in educating deaf and hard-of-hearing students. New York, NY: Oxford University Press.

Steeds, L., Rowe, K., & Dowker, A. (1997). Deaf children's understanding of beliefs and desires. Journal of Deaf Studies and Deaf Education, 185-195.

Turati, C., Di Giorgio, E., Bardi, L., & Simion, F. (2010). Holistic face processing in newborns, 3-month-old infants and adults: Evidence from the composite face effect. Child Development, 81, 1894-1905.

Walker-Andrews, A. S., & Lennon, E. (1991). Infants' discrimination of vocal expressions: Contributions of auditory and visual information. Infant Behavior and Development, 14(2), 131-142.

Waxman, S. R., & Gelman, S. A. (2010). Different kinds of concepts and different kinds of words: What words do for human cognition. In D. Mareschal, P. C. Quinn & S. E. G. Lea (Eds), The making of human concepts (p. 101-130). New York Oxford University Press.

Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12-to 13-month-old infants. Cognitive psychology, 29(3), 257-302.

Welder, A. N., & Graham, S. A. (2006). Infants' categorization of novel objects with more or less obvious features. Cognitive psychology, 52(1), 57-91.

Widen, S. C., & Russell, J. A. (2002). Gender and preschoolers' perception of emotion. *Merrill-Palmer Quarterly, 48*(3), 248-262. doi: 10.1353/mpq.2002.0013

Widen, S. C., & Russell, J. A. (2004). The relative power of an emotion's facial expression, label, and behavioral consequence to evoke preschoolers' knowledge of its cause. *Cognitive Development, 19*(1), 111-125. doi: 10.1016/j.cogdev.2003.11.004

Widen, S. C., & Russell, J. A. (2010a). Children's scripts for social emotions: Causes and consequences are more central than are facial expressions. *British Journal of Developmental Psychology, 28*(3), 565-581. doi: 10.1348/026151009X457550d

Widen, S. C., & Russell, J. A. (2010b). Differentiation in preschooler's categories of emotion. *Emotion, 10*(5), 651-661. doi: 10.1037/a0019005

Widen, S. C. (2016). The development of children's concepts of emotion. In L. F. Barrett, M. Lewis, J. M., & Haviland-Jones (Eds.), *Handbook of emotions* (4th ed., pp. 307-318). New York: Guilford Press.

Box 15: The Habituation Task Used in Studies of Emotion Perception in Infants

The most popular experimental design that is used to study emotion perception in preverbal infants is called the **habituation task**. In a standard habituation task, an infant is presented with a stimulus that experimenters believe represents an emotion category (say, a scowling face to represent anger) and the infant will look at it. As subsequent stimuli are presented (say, a series of scowling faces), the infant is assumed to identify or categorize each face according to its emotional meaning. If the infant categorizes the stimuli as belonging to the same emotion category, it will look for a shorter amount of time (because the infant is presumed to become bored). Once the infant's looking time drops below a certain threshold, experimenters assume that the child is habituated to the emotion category (i.e., the infant has become uninterested in looking at "scowling faces"). Then, a new stimulus is presented, and looking time is again recorded (e.g., after a viewing a series of scowling faces, the infant is then shown a smiling face). If the infant looks for a *longer* time, scientists infer that she has categorized the stimulus as belonging to an emotion category that is different from before (i.e., the infant is assumed to be interested in novelty; this novelty is supposed to reflect a category difference in the context of the experiment).

The habituation task obviously requires that the experimenter infer what looking times mean. Such inferences call for having strong alternative hypotheses that are, in practice, rarely considered. For example, the proposed expressive forms in Figure 4 differ in their familiarity (e.g., most infants are more familiar with smiling faces than with scowls or frowns). This makes it difficult to know which features of a face are holding an infant's attention (familiarity or novelty), and can lead to potentially incorrect inferences. For example, several studies claim that infants are somehow born prepared to detect fearful faces. But the proposed expression for fear is less familiar than "happy" faces; infants may look longer at them because they are attempting to learn novel stimuli (e.g., Bayet et al., 2017; see also Peltola et al., 2008, 2009).

References:

Bayet, L., Quinn, P. C., Laboissière, R., Caldara, R., Lee, K., & Pascalis, O. (2017). Fearful but not happy expressions boost face detection in human infants. *Proc. R. Soc. B*, *284*(1862), 20171054.

Peltola, M. J., Leppänen, J. M., Palokangas, T., & Hietanen, J. K. (2008). Fearful faces modulate looking duration and attention disengagement in 7-month-old infants. *Developmental Science*, *11*(1), 60-68.

Peltola, M. J., Leppänen, J. M., Vogel-Farley, V. K., Hietanen, J. K., & Nelson, C. A. (2009). Fearful faces but not fearful eyes alone delay attention disengagement in 7-month-old infants. *Emotion*, *9*(4), 560.

Box 16: Information Theory as Applied to Emotion Communication

An information theory approach to understanding emotional communication asks: what is the information that one person's facial movements (the sender) convey to another person (the perceiver) in a particular context and how does that information help achieve synchrony between the sender and perceiver. The questions are cast in terms of how the sender and perceiver share information. The sender is communicating some information with his facial movements (sometimes emotional, sometimes not), and the perceiver uses that information to reduce uncertainty about what the sender is going to do next (in Western cultures, this typically means making a mental inference that allows the perceiver to predict the sender's actions). For recent examples of experiments that use information theory as a scientific tool, see Jack & Schyns, (2017).

This approach begins with the assumption that one person, the sender, is in some state of mind, makes some facial movements, and performs some action (speaks, moves his body, etc.) in a particular situation. The perceiver detects these movements and makes a prediction about what the sender will do next, often by making an inference about the sender's state of mind. Communication of information between the sender and the perceiver is understood as mutual information, which is a symmetric measure of the information shared between two variables, X and Y. Specifically, mutual information is a measure of how much knowing value of one variable, X, tell us, reduces the uncertainty, about the value of the other variable, Y, and vice versa.

Formally, mutual Information is a symmetric measure of the dependency between two random variables that provides a nonparametric way to assess how much knowing some random variable, Y, reduces the uncertainty about another random variable, X. Mutual information is based on Shannon's notion of information which is measured by entropy, $H(X)$, a measure of the uncertainty of a random variable. Mutual Information, $I(X;Y)$, is the reduction in uncertainty that results from knowing Y, I(X;Y) = H(X)-H(X|Y) (see Cover & Thomas 2006). If X and Y are independent events, knowing Y provides no information about X and the entropy H(X) is equal to the entropy H(X,Y) and I(X;Y) is zero. On the other hand, if knowing Y fully determines X and Y fully determines X, then the mutual information reduces to the entropy of X (or Y).

When asking whether facial movements express an instance of a certain emotion category emotions, where X is the sender's state of mind and Y is his facial movements, the forward inference is as follows: How much does X, his state of mind, reduce uncertainty about Y, his facial movements, perhaps measured with facial EMG? In the reverse inference, X is the sender's facial movements and Y is his state of mind: How much does X, his facial movements, again measured with facial EMG, reduce uncertainty about his state of mind, Y?)

References

Chen, C., Crivelli, C., Garrod, O., Schyns, P., Fernández-Dols, J. & Jack, R. (2018) Distinct facial expressions represent pain and pleasure across cultures. Proceedings of the National Academy of Sciences Oct 2018, 115 (43)

Jack, R. and Schyns, P. (2017) Toward a Social Psychophysics of Face Communication Annual Review of Psychology 2017 68:1, 269-297

Thomas M. Cover and Joy A. Thomas. (2006). Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing). Wiley-Interscience, New York, NY, USA.

About the Authors

Lisa Feldman Barrett is University Distinguished Professor of Psychology and Director of the Interdisciplinary Affective Science Laboratory (IASLab) at Northeastern University, with research appointments at Harvard Medical School and Massachusetts General Hospital. Her research focuses on the nature of emotion from both psychological and neuroscience perspectives. Dr. Barrett is the recipient of numerous research awards, including the 2018 APS Lifetime Mentor Award and the NIH Director's Pioneer Award for transformative research. She is an elected fellow of the American Academy of Arts and Sciences, the Royal Society of Canada, the American Association for the Advancement of Science, the Association for Psychological Science, as well as several other scientific societies. She was recently chosen one of the 50 most influential living psychologists in the world by thebestschools.org. Her research has been continuously funded by NIH and NSF for over 20 years. Dr. Barrett also educates lawyers, judges and other legal actors about emotion, neuroscience and the law as part of her work for the Center for Law, Brain and Behavior. In addition to publishing over 200 peer reviewed papers and 50 book chapters, Dr. Barrett has testified before US Congress in support of basic behavioral research funding and has edited five volumes, including the 4th edition of the *Handbook of Emotion*, published by Guilford Press. Her book, *How emotions are made: The secret life of the brain*, is published by Houghton Mifflin Harcourt. More about Dr. Barrett's work can be found at www.affective-science.org and https://lisafeldmanbarrett.com/

Ralph Adolphs is the Bren Chair Professor of Psychology, Neuroscience and Biology at Caltech in Pasadena, California. His graduate studies were in basic cellular electrophysiology (the auditory system of the barn owl), but his post-doctoral work switched to cognitive neuroscience in humans (cognitive neuropsychology with Antonio Damasio at the University of Iowa). His current lab (emotion.caltech.edu) studies emotion and social behavior using a number of approaches, including electrophysiological recordings from the brain in surgical patients, studies of people with autism, and fMRI. He directs the Caltech Brain Imaging Center and chairs the Institutional Review Board at Caltech (which reviews the ethics of human subjects research). He is the co-author (with David Anderson, a Caltech colleague) of the book, "The Neuroscience of Emotion", published by Princeton University Press.

Stacy C Marsella is a Professor in Computer Science and Psychology at Northeastern University in Boston, Massachusetts. His research is in the computational modeling of emotion, cognition and social behavior, both as a basic research method in the study of human behavior as well as the use of these computational models in a range of applications. His current research studies the interplay of emotion and cognition in decision-making as well as the role that beliefs about others and nonverbal behavior play in social interaction. He pursues a highly interdisciplinary approach that integrates methods and models from computer science, psychology and art. His work is also applied to the design of virtual humans. This technology is being applied in education, health interventions and entertainment as well as being used as stimuli for experiments in social psychology. In addition, his work on modeling and simulating human behavior is used in large scale simulations of social behavior. The research is currently funded by NSF, NIH, AFOSR, DARPA and ONR. Dr. Marsella is a member of the Society of Experimental Social Psychologists, International Society for Research on Emotion and the Association for the Advancement of Artificial Intelligence. His work has received numerous awards including the ACM SIGAI Autonomous Agents Research Award and a UNESCO's promising and innovative digital initiatives award.

Aleix M. Martinez is Professor in the Department of Electrical and Computer Engineering at The Ohio State University (OSU), where he is the founder and director of the Computational Biology and Cognitive Science Lab. He is also affiliated with the Department of Biomedical Engineering, to the Center for Cognitive and Brain Sciences, and to the Center for Cognitive and Behavioral Brain Imaging, where he is a member of the executive committee. He has served as an associate editor of IEEE Transactions on Pattern Analysis and Machine Intelligence, IEEE Transactions on Affective Computing, Image and Vision Computing, and Computer Vision and Image Understanding. He has been an area chair for many top conferences and was a Program co-Chair for CVPR 2014. He is also a member of NIH's Cognition and Perception study section. Aleix is most known for being the first to define many problems and solutions in face recognition (e.g., recognition under occlusions, expression, imprecise landmark detection) as well as providing a systematic, in-depth study of the production and perception of facial expressions in the wild. He is also well-known for his work on discriminant analysis (e.g., Bayes optimal solutions, subclass-approaches, optimal kernels), and structure from motion (e.g., using kernel mappings to better model nonrigid changes, noise invariance). Aleix received a Google Faculty Award in 2015.

Seth Pollak is the College of Letters and Science Distinguished Professor of Psychology and Professor of Pediatrics, Anthropology, Neuroscience, and Public Affairs at the University of Wisconsin – Madison. He is also an investigator at the University's Waisman Center for Human Development. Pollak earned dual Ph.D.s from the University of Rochester in brain & cognitive sciences and in child clinical psychology. Professor Pollak has published widely on topics including children's emotional development, the influences of social risk factors on children's brain and behavioral development, and the effects of early experience on children's learning and health. He is a recipient of the American Psychological Association's Distinguished Early Career Award and the Boyd-McCandless Award for Distinguished Contributions to Child Development. Dr. Pollak is a Fellow of the American Association for the Advancement of Science and the Association for Psychological Science, as well as a member of the Society for Research in Child Development's Governing Council.