

# Emotion in Psychology and Machine Learning

Pablo Caceres

April 30, 2020

## Abstract

(1) after decades of debate, researchers in the psychology and the neuroscience of emotion, do not agree on what emotions are and how ought to be studied; (2) machine learning researchers have taken a notoriously practical approach to the task of emotion recognition and generation, where data availability is the main factor determining how emotions are conceptualized.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>What emotions are not</b>	<b>2</b>
<b>3</b>	<b>Theories of emotion: the emotion wars</b>	<b>4</b>
3.1	Basic emotion theory . . . . .	5
3.2	The circumplex theory of core affect . . . . .	6
3.3	The theory of constructed emotion . . . . .	6
3.4	The functionalist theory of emotion . . . . .	7
<b>4</b>	<b>What researchers that study emotion agree about</b>	<b>7</b>
<b>5</b>	<b>The measure of emotion</b>	<b>7</b>

# 1 Introduction

Emotion has been subject of speculation for philosophers, theologians, and scientists for millennia (Barrett, 2017; Sorabji, 2000). In Ancient Greece, the Stoic school of thought regarded emotion as mental judgments and attitudes, particularly the mistaken ones (Sorabji, 2000). Scholastic philosophers like Thomas Aquinas<sup>1</sup> considered emotions as “a movement of the sense appetite caused by imagining good or evil.” (as cited in Lombardo, 2011, p. 20), whereas Enlightenment philosophers like René Descartes (Descartes, 1649/1995) thought of emotions as “perceptions, feelings, or emotions of the soul which we relate specially to it, and which are caused, maintained, and fortified by some movement of the spirits” (p., 344).

Since the advent of scientific psychology in the XIX century, psychologists and philosophers have debated for decades about the meaning and nature of emotions (XXXXX).

# 2 What emotions are not

Although there is no consensus about what emotions are, most authors distinguish emotions from other psychological constructs like personality traits, mood, mental disorders, and affective states. In particular, we focus our attention on the difference between emotion and *affect*, since these concepts are tightly related and therefore easily conflated.

As with emotion, there isn't a universally accepted definition of *affect*. One of the first definitions was proposed by Wilhelm Wundt (XXX), who understood affect as momentary mental state, characterized by feelings of pleasantness/unpleasantness, rousing/subduing, and strain/relaxation (XXXX). These fleeting states can be considered as 'basic' or 'primi-

---

<sup>1</sup>It is important to hold in mind that Aquinas, and intellectuals previous to the modern scientific era more generally, never wrote about the specific concept of "emotion". Rather, as Lombardo (Lombardo, 2011) notice in the case of Thomas Aquinas, they wrote about *appetites*, *passions*, *affections*, and other related concepts.

tive' (i.e., irreducible) ingredients of the human mind. Wundt's perspective has influenced contemporary views under the rubric of *core affect* ("core" in the sense of "basic" or "primitive"). According to Russell (XXXX), *core affect* is "A neurophysiological state that is consciously accessible as a simple, nonreflective feeling that is an integral blend of hedonic (pleasure–displeasure) and arousal (sleepy–activated) values" (p., 147). Hence, affect is a fundamental part of emotional events, but emotions can not be reduced to affect. Take this quote from the Poet Derek Walcott:

The time will come when, with elation, you will greet yourself arriving at your  
own door, in your own mirror, and each will smile at the other's welcome. (CITE  
Love After Love, Poem)

In this example, *elation* is accompanied by the face expression of a smile. Thus, many emotional experiences will involve other forms of behavior in addition to the fleeting states of hedonic valence and arousal. Russell and Barrett (XXXX) make this clearer by distinguishing between *prototypical emotional episodes* and core affect. On their view, prototypical emotional episodes include core affect, overt behavior (e.g., the smile on the mirror when arriving at your own door), attention towards, appraisal of, and attributions towards the instance triggering the emotional episode (i.e., meta-cognition), and the compound of neurochemical and physiological reactions underlying these psychological events. It can be objected that dimensional perspectives on emotion put hedonic valence and arousal at the center of the construct of emotion (XXXX). To this, we say that not even those authors conflate affect with emotion. Among theorist of emotion, Russell and Barrett are probably the ones that more strongly favor the idea of core affect as a fundamental part of the emotion experience, yet, as we detailed above, not even them conflate affect with emotion. This will become clearer in the next section where we review psychological theories of emotion in more detail.

### 3 Theories of emotion: the emotion wars

Controversy has dominated the psychological and neuroscientific landscape of emotion theory in the last decades (XXXX). Several of the disagreement among emotion theorist are deep, as deep as the disagreements about other controversial subjects in psychology, like nature versus nurture (XXX), or distributed versus symbolic cognition (XXXX). Our reading of the literature indicates that the main cleavages dividing the field are around the following issues:

- *Are emotions natural kinds?* Another way to put this is: Are emotions "discovered" or "constructed"?
- *Do emotions have an evolutionary crafted function?* For instance: Does "anger" have an predefined evolutionary function like removing obstacles or facing predators?
- *Do emotions have objectively identifiable markers?* For instance: Are there universally shared facial expressions that serve as objective "markers" of happiness or sadness?
- *How are emotions represented in the human mind?*

These cleavages are not mutually exclusive. For instance, researchers that favor the idea of emotions as natural kinds, also favor the idea of emotions as having evolutionary predetermined functions. Yet, we hope having this questions in mind may serve as a compass to understand the diversity of views in the literature. In our review we identified at least four main traditions in emotion theory: (1) basic emotion theory; (2) The circumplex theory of core affect; (3) the theory of constructed; (4) the functionalist theory of emotion. In what follows, we examine these perspectives and how they relate to the above mentioned issues. Our primary goal is to describe the main ideas of each perspective, rather than exposing the accumulated evidence to declare a "winner" of this debate. Such endeavor is beyond the scope of this review. For broader discussion about it see (CITE, CITE).

TLDR Table

theory	natural kind	evolutionary	objective markers	representation
Basic Emotion Theory	Yes	Yes	Yes	Discrete

### 3.1 Basic emotion theory

In the opening scene of the American crime television drama *Lie to me\**, we see the encounter between a man accused of a crime and Dr. Cal Lightman, a renowned expert in the science of interpreting micro-expression and body language. In the scene, the camera carefully traces every microscopic body movement of the accused: the quivers of his eyebrow, the pressing of his lips, the twitching of his finger, the heavy breathing of his chest. For Dr. Lightman, every micro-expression reveals the thoughts and emotions of the man in front of him, knowledge that Dr. Lightman will weaponize to carve out the truth about the accused past actions.

The science behind Dr. Lightman skills on *Lie to me\**, is based on the work of Paul Ekman (XXXX). Actually, not only the science, but Dr. Lightman itself is loosely based on Paul Ekman persona (XXXX). Of course, *Lie to me\** is not a literal recreation of Ekman work, but the *core idea* behind Ekman work is what serves as inspiration: emotions as *discrete* mental states, that can be cataloged into a finite set of *classes*, and that can be recognized by stereotypical, universally shared, *body responses* and *facial expressions* (XXXX). From this perspective, emotions qualify as *natural kinds*: as a class of natural phenomena that exist independently of human conventions (XXXX). Societies across history may have labeled such phenomena differently, but its existence is not subject to human arbitrariness. In this sense, emotions are no different than elementary particles in physics: as fermions and bosons, emotions are *discovered*, not constructed. Consequently, if emotions are natural kinds, there must be a way to objectively identify them by the means of science, and indeed, methods have been developed along this line of reasoning for such purposes (XXXX). This view is what is referred to in the literature as *basic emotion theory*. Although Ekman's view is probably the most prominent in this tradition, other authors Izard (XXX) and Panksepp & Watt (XXXX), have also proposed variations of the same perspective.

The exact number of basic emotion classes has changed over time, but today is common to see authors referring the following (CITE: Ekman 2011): (1) anger, (2) fear, (3) surprise, (4) sadness, (5) disgust, (6) contempt, (7) happiness. Ekman and colleagues (XXXX) have developed detailed descriptions of this emotion instances, allowing researchers to identify, catalogue, simulate (with actors or computationally), and even trigger emotional experiences in laboratory experiments (XXXX).

It is important to remark that the basic emotion perspective theory does not propose that emotions are monolithic or inflexible. Emotions do have variation, and can be shaped by life experiences and culture (XXXX). Yet, such variation is constrained by a core of evolutionarily crafted basic emotions, which are shared by all humans as species. The reason why, on this view, emotions have such a strong predetermined foundation, is related to their survival advantage over evolutionary scale (XXXX).

For an extended review of the basic emotion theory perspective, see the reviews of Izard (XXXX) and Ekman & Cordaro (XXXX).

### **3.2 The circumplex theory of core affect**

Why discussing a theory of affect in the context of theories of emotion? First, because in practice, affect is often used interchangeably with emotion as a concept. Second, because its historical relevance for the study of emotion. And third, because in many machine learning applications emotion is operationalized as core affect.

The circumplex theory of core affect was introduced by Russell (CITE 1980 paper). Yet, it has

### **3.3 The theory of constructed emotion**

Feldman

### 3.4 The functionalist theory of emotion

## 4 What researchers that study emotion agree about

Considering the variety of perspective on emotion, and the many fundamental disagreements about its nature, is easy to be left with the impression that there is no common ground on which to base applied research efforts. In other words: Are there points of agreement that can be taken as base to guide applied research utilizing emotion as construct? Of course, we can not speak for the emotion research community on this topic. Yet, based on our interpretation of the debate, there at least three points around which most researchers in the field would agree:

- *emotions are real*
- *emotions have significant variability and can be shaped by life-experiences*
- *emotions can be recognized and measured*

## 5 The measure of emotion

### References

- Barrett, L. F. (2017). *How emotions are made: The secret life of the brain*. Houghton Mifflin Harcourt.
- Descartes, R. (1649/1995). The Philosophical Works of Descartes. In *The Passions of the Soul*. Dover Publications.
- Lombardo, N. E. (2011). *The logic of desire: Aquinas on emotion*. CUA Press.
- Sorabji, R. (2000). *Emotion and peace of mind: From stoic agitation to Christian temptation*. Oxford University Press.