INTRODUCTION TO VOLUME I

The Value of History, a Wealth of Theoretical Options, and the Elements of Emotion Theory

Andrea Scarantino

Emotion Theory: The Routledge Comprehensive Guide is the result of 6 years of work by 101 leading emotion theorists in philosophy, classics, psychology, biology, psychiatry, neuroscience, and sociology. It includes around 750,000 words in total, distributed over two volumes and 62 chapters. The two volumes aim to introduce you to:

- The history of emotion theory over the past 25 centuries, mostly in the Western tradition, but with brief forays into the Indian and the Chinese traditions (Volume I, Part I)
- The main contemporary theories in the philosophy of emotions and in affective science (Volume I, Part II)
- The elements out of which emotions are composed and their interpretations in different theories of emotions (Volume I, Part III)
- What we know about the definitions, bodily underpinnings, and functions of 35 specific emotions (Volume II, Part IV)
- Some of the most significant theoretical challenges currently facing emotion theory in a variety of disciplines (Volume II, Part V)

I am very proud of this collection, and grateful beyond measure to the distinguished authors who have contributed to it, dealt with my surely aggravating requests to break everything down into digestible parts, and patiently waited for this mammoth project to come to an end. I believe the two volumes you are now perusing offer an unrivaled introduction to emotion theory in terms of systematicity, comprehensiveness, relevance, and accessibility.

The defining ambition of this project is to speak to two audiences at once: experts and neophytes. To experts, the volumes aim to offer a cutting-edge synthesis of the available philosophical and scientific literature on a variety of critical topics. All chapters provide systematic overviews of decades and in some cases centuries of theoretical investigation, articulating innovative taxonomies of the available theoretical options, with their attendant costs and benefits. To neophytes, the volumes aim to offer a one-stop shop where one can learn the ropes of the field without being hindered by discipline-specific jargon or misled by oversimplifications. When I started my journey in emotion theory, this is exactly what I needed and craved for but could not find, which led me to spend several years trying to find my way in an intricate and terribly confusing literature.

DOI: 10.4324/9781315559940-1

You will be the judge as to whether we have succeeded in achieving both theoretical depth and expository clarity. What I can tell you is that we have tried *very hard* – each chapter has been written and rewritten with the overarching objective of presenting what we know about emotions in a way that combines informativeness and understandability. I hope you will enjoy the fruits of our collective labor.

I will now explain the rationale for each of the three parts of Volume I, give you an overview of the contents of the chapters contained within, and hint at some of the insights about emotions that Volume I aspires to provide.

Part I: History of Emotion Theory

Why should we care about the history of emotion theory, you may wonder? After all, when we study chemical elements like oxygen or astronomical objects like the stars, we do not need to learn about what Lavoisier or Galileo thought about them. If we come upon their theories, we immediately realize that they entertained all sorts of mistaken ideas of no use to modern chemists and astronomers. Things are somewhat different when it comes to emotion concepts. The historical chapters in Part I collectively make a powerful case that emotion concepts are not like the concepts of *oxygen* or *star*, because they systematically mix descriptive and normative concerns.

Chemical and astronomical concepts are paradigmatic *natural kind concepts*, namely concepts which aim to carve nature at its joints by individuating inductively rich categories about which reliable scientific generalizations can be formulated. Oxygen and stars are supposed to exist as they are independently of *any* human practices – what makes them natural kinds depends not on how we think and talk about them, but on their objective properties, which scientists try to discover as they search for laws of nature.

A great many emotion theorists have historically treated emotion concepts as natural kind concepts. Aristotle already told us that one can approach the study of emotions as a *natural scientist*, and define emotions in terms of their bodily manifestations (e.g. anger is the boiling of blood and heat around the heart; *De anima* 403a 16-403b 1). This approach thrived especially in the 17th century at the rise of mechanistic philosophy, when the physiology, the expressions, and (eventually) the neural underpinnings of emotions became central theoretical concerns.

But Aristotle added that there is another way to conceptualize emotions typical of the *dialectic philosopher*, who cares about how emotions result from appraisals of the stimulus situation and dispose one to action (e.g. anger results from appraising stimuli as slights and leads to an impulse for revenge). This second aspect of emotions was important for practical reasons, which is why we find Aristotle's theory of emotions (mostly) in the *Rhetoric*, a treatise meant to help public speakers shape the emotions of their audience strategically, in order to gain persuasive advantages in the context of political oratory and lawsuits. The Middle Ages offer many other prominent examples of the importance of this practice-oriented approach – several of the leading theorists of the time conceptualized emotions in terms of how a sage should manage them to lead a morally virtuous life.

What the history of emotion theory reveals is an essential, but not much explored, duality at the core of emotion concepts: they have been treated concurrently as *natural kind concepts*, used in "projects of inquiry in which extension and intension are altered to preserve inductive and explanatory power", and as *normative kind concepts*, used to "prescribe and condemn … behavior" (Griffiths 2004, 907; see also Hacking 1999). These are, of course, very different projects, animated by distinct objectives which pull in disparate directions.

When emotion concepts are used in projects of inquiry, the objective of the theorist is to conceptualize emotions in ways that are helpful for describing nature, for making predictions, and for explaining behavior. This often leads to focusing on intersubjectively available physical markers of emotions and trying to discover empirically robust generalizations. When emotions are used in normative projects, the objective of the theorist is to conceptualize emotions in commonsensical ways that promote human-centered practices like linguistic communication, morality, aesthetics, religion, therapy, medicine, art, law, and politics.

The problem is that there is no reason to expect that conceptual frameworks optimized in the service of naturalistic description, prediction, and explanation will also best serve whatever human-centered practices motivate our possession of folk emotion concepts (and vice versa). This creates a threat to emotion theory which has emerged with increasing vigor over the past 30 years, but was interestingly anticipated in the first half of the 20th century, namely that folk emotion concepts, although suitable for normative projects, are just too variable in physical terms to function well in projects of inquiry. Despite this skeptical threat, for most of its history emotion theory has been propelled by the implicit assumption that a good theory of emotions in projects of inquiry is a theory of the very things we call "emotions" in everyday life for purposes of praising, blaming, cajoling, controlling, influencing, and educating one another. It is worth noting that most contemporary emotion theorists still hold this view and embody some of the traits of both natural scientists and dialectic philosophers of Aristotelian lore.

As you will notice in the following chapters, the interplay between normative and descriptive projects has been a formidable force shaping emotion theories over centuries and civilizations. You will observe it at work in the way emotion theories in Ancient Greece were influenced in part by moral concerns about virtue and the good life; in the ways Indian and Chinese theories of emotions were driven by a practical interest in therapeutic intervention, moral development, and aesthetic cultivation; in the way medieval theories of emotions were affected by the preoccupation of reconciling ancient Greek moral doctrines with the teachings of Christianity and the avoidance of sin; in the way Renaissance theories of emotion were affected by an optimist outlook on the value of humanity and concerns with the artistic representation of the human body and the preservation of its health through medicine and natural magic; and in the way 17th- and 18th-century theories of emotion were influenced by epistemic concerns with knowledge, moral concerns with freedom of choice, and political concerns with the social order and the wealth of nations.

When science and philosophy parted ways in the mid-19th century, a clearer divide emerged between attempts to study emotions as natural kinds and attempts to study them as normative kinds. But the two projects never fully separated. Affective scientists throughout the 20th century were driven in part by the attempt to make their theories of emotions compatible with commonsense uses of emotion concepts in human-centered practices. Conversely, lots of philosophers aimed to make their theories of emotions empirically informed as they engaged in normative projects in epistemology, ethics, aesthetics, and political philosophy. And very few theorists in either camp felt the need to distinguish between what emotion terms *mean* as ordinarily understood, and what is a *good thing to mean* by emotion terms for theoretical purposes.

Even James, who famously challenged common sense by arguing that we do not run because we are afraid, but are afraid because we run, ended up relying on common sense when it came to justifying his physiological theory of emotions. He stated that emotions are perceptions of bodily changes because a disembodied emotion is for us *inconceivable* – none of us can allegedly think of it. Of course, many of James's critics could easily think of a disembodied emotion, and developed alternative theories on that basis, but the point is that an astronomer

or a chemist would never rely on this style of argument, and defend a theory of oxygen or a theory of stars in light of what ordinary language concept users can *conceive* oxygen or stars to be.

What the exploration of 25 centuries of emotion theory provided by Part I ultimately offers you is a unique opportunity for understanding what desiderata drive researchers to formulate theories of emotions, and how such desiderata have changed across eras, civilizations, social systems, disciplines, and research traditions. You may or may not end this exploration convinced that a conceptual framework devised in the service of projects of description, prediction, and explanation won't efficiently serve human-centered functions, and that a conceptual framework tailored in the service of normative projects can't possibly carve nature at its joints.

In **Chapter 1: Emotion Theory in Ancient Greece and Rome**, Pia Campeggiani and David Konstan introduce us to the two main normative options articulated in Ancient Greece and Rome when it comes to how a sage should live: the Stoic's ideal of *apatheia*, or freedom from the emotions, and the Aristotelian ideal of *metriopatheia*, or having the appropriate emotions in the appropriate circumstances. Campeggiani and Konstan discuss in some detail the theories of the emotions developed by Plato, Aristotle, the Epicureans, and the Stoics, highlighting their groundbreaking contributions on emotion elicitation, phenomenology, and intentionality. They examine two influential assumptions about the nature of emotions introduced in Ancient Greece. One is that emotions stand in opposition to *activity* – as Aristotle put it, emotions are *pathê*, or things we are affected by, rather than things we do. The other assumption is that emotions stand in opposition to *reason* – Plato described them as wild horses to be controlled by the charioteer of reason. As the subsequent chapters make clear, versions of these two ideas, and in due course opposition to them, have shaped centuries of theoretical investigation on the emotions.

In **Chapter 2: Emotion Theory in Ancient and Classical India, 500 BCE–1200 CE**, Maria Heim discusses the conceptualization of emotions in the Indian tradition. Her primary focus is on an early Pali Buddhist canonical body of scripture called the Abhidhamma (with its commentaries) and the secular Sanskrit literature on aesthetics, both of which focused on the disaggregation of emotional experience in order to, respectively, therapeutically transform it and better understand the nature of art. Heim focuses on envy and jealousy as case studies to show how Sanskrit relies on a cluster of five terms – *amarsha*, *akshama*, *irshya*, *asuya*, and *matsara* – to draw subtle distinctions pertaining to the phenomenology of envy and jealousy. Heim also discusses strategies of emotion regulation proposed by religious thinkers within the Indian tradition, and in particular the practice of replacing one emotion with another (the cultivation of opposites). Heim's textual analysis shows how the study of emotions in classical India requires linguistic sensitivity to categories which are translatable into English, but demand extensive interpretive work because they do not perfectly coincide with the corresponding English categories.

Curie Virág emphasizes in **Chapter 3: Emotion Theory in Early and Medieval China, 500 BCE–1200 CE** that traditional Chinese thinkers were driven to theorize about emotions by their broader intellectual concerns in ethics, politics, psychology, and natural philosophy. She discusses theories of emotions developed in the early Confucian and Daoist traditions (by Confucius, Mencius, Laozi, and Zhuang Zhou), in the Han Empire (by Dong Zhongshu), and in the Period of Disunion at the end of the Han Dynasty (by Wang Bi and Xi Kang). Virág pays special attention to the historical development of the classical Chinese terms *qing* 情 (the term arguably closest to the English word "emotion"), heart-mind (*xin* 心), inborn nature (*xing* 性), and desire (*yu* 欲), emphasizing that the meanings of these terms have constantly been

renegotiated to reframe the inquiry into affective phenomena. Both Heim and Virág make a convincing case that emotions are understood within the Indian and Chinese traditions very differently from how they are understood within the Western tradition. This puts the methods of Western emotion research in stark relief, and hints at possible alternatives to it.

Chapter 4: Emotion Theory in the Middle Ages by Simo Knuuttila covers the medieval tradition in the West and its deep roots in Ancient Greek and Roman thought. Knuuttila documents how the Stoic doctrine of apatheia eventually came into conflict with the dictates of Christianity, which teaches us to fear God, to love one's neighbors, and to be angry at sin. The normative focus transitioned from leading the life of the sage to leading a life conducive to salvation, a process which culminated in Pope Gregory the Great's canonization of the seven deadly sins in the seventh century, many of which are emotions: anger, envy, sloth, gluttony, lust, pride, and avarice. Knuuttila discusses among others the theories of emotions developed by Augustine, Avicenna, Abelard, Aquinas, Scotus, and Ockham, emphasizing the relative disinterest of medieval thinkers in the phenomenology of the emotions compared to their interest in their evaluative and motivational dimensions. Knuuttila also tracks the shifting grounds of an essential taxonomic distinction in medieval emotion theory, namely the one between concupiscible passions and irascible passions, the former connected to the appetitive part of the soul and the latter connected to the spirited part of the soul (a distinction which originates in Plato).

As Sabrina Ebbersmeyer argues in **Chapter 5: Emotion Theory in the Renaissance**, in the Renaissance the tie of emotions with sin was partially severed – the emotions became valuable expressions of our earth-bound humanity. The humanists of the Renaissance not only accepted human nature as it is, but they also attempted to characterize human behavior without moral evaluation. When Plato's *Symposium* was rediscovered in the 15th century, interest in the emotion of love and its varieties became dominant, with love playing a central role in a virtuous life understood roughly along the lines of Aristotle's metriopatheia. Ebbersmeyer's chapter discusses several humanist accounts of emotions, including those by Petrarca, Bruni, Machiavelli, and Montaigne. Jointly, these contributions started undermining the view that emotions are enemies of reason, highlighting the many ways in which they can be beneficial to human flourishing. Work by Vives, Melanchton, Ficino, and Campanella also showed how a magical approach to healing went along with an interest in the physiology of emotions and in the role emotions could play in medical treatments'.

The 17th and 18th centuries continued to be a time of commingling of descriptive and normative concerns. Lisa Shapiro argues in **Chapter 6: Emotion Theory in the 17th Century** that the study of human anatomy and physiology that started in the Renaissance continued to thrive in the 17th century, leading to a heightened interest in the bodily underpinnings of the emotions. Shapiro focuses in particular on Burton's, Reynoldes's, and Descartes's contributions on the physiology of emotions, charting the progressive emergence of a mechanistic model of affective states. She also describes in some detail the influential theories of the emotions developed by Descartes, Spinoza, and Hobbes, explaining how they each developed some aspects of the medieval taxonomy of the emotions made popular by Thomas Aquinas, while retaining the view that emotions are fundamentally tied to actions. Shapiro concludes by exploring how new normative concerns shaped 17th-century theories of emotions, specifically concerns with the impact of emotions in politics and with their impact on free will.

As Eric Wilson tells us in **Chapter 7: Emotion Theory in the 18th Century**, the Cartesian mechanistic approach to emotions continued to be developed in the 18th century, but most philosophers of the period focused on emotions in the context of morality, politics, religion, and social interaction. Wilson discusses the theories of emotions articulated by Wolff,

Hutcheson, Hume, Kant, Rousseau, Mandeville, and Smith. He focuses on a variety of taxonomic distinctions popular in the 18th century, such as the distinction between selfish and public affections (Hutcheson), direct and indirect passions (Hume), and natural and cultural passions (Kant). Smith's and Hume's competing theories of sympathy are also covered in some detail. Finally, Wilson considers emerging 18th-century theories of emotion regulation, showing how the Platonic assumption that reason is the charioteer of the passions was called into question by the view that reason is the slave of the passions (Hume) and by the view that the passions are regulated by psychological mechanisms other than reason (Smith).

It is in the 19th century that descriptive questions about the roles played by emotions in the architecture of the mind start being explicitly distinguished from normative questions about the roles they ought to play in human life, although a great many accounts continued to be jointly influenced by both concerns. In **Chapter 8: Emotion Theory in the 19th Century at the Rise of Scientific Psychology**, Rainer Reisenzein discusses the groundbreaking theories offered by Darwin, James, Wundt, Meinong, and Freud, explaining how they laid the foundations for, respectively, evolutionary theories of emotions, bodily feeling theories of emotions, mental feeling theories of emotions, cognitive theories of emotions, and psychoanalytic theories of emotions. Reisenzein pays special attention to Darwin's theory of emotional expressions, to James's analysis of the elicitation of emotional experiences and of the origins of emotion mechanisms, to the tridimensional theory of emotional experience proposed by Wundt, to the cognitive basis of emotions analyzed in Meinong's theory of emotions, and to the role of unconscious affects in Freudian psychoanalysis. Reisenzein concludes by reconstructing the legacy of these five pioneers of emotion science. Collectively, they contributed to the emancipation of psychology from philosophy and continued to influence emotion theory as the new science of psychology changed its focus first from consciousness to behavior, and then from behavior to mental representations.

In **Chapter 9: Emotion Theory in the 19th- and 20th-Century Phenomenological Tradition**, Íngrid Vendrell Ferran introduces us to the phenomenological tradition of research on the emotions, with special focus on the theories developed between 1874 and 1950. She distinguishes between an initial phase of the tradition, starting with Brentano's descriptive psychology and Husserl's analysis of the intentional structure of consciousness; a realist phase associated with Scheler and Pfänder, who focused on the essence of emotions; a third phase associated with Heidegger's focus on the existential significance of emotions; and a fourth phase marked by Sartre and Merleau-Ponty's theories of emotions as embodied engagements. As Vendrell Ferran notes, phenomenologists viewed the intentionality of emotions as intrinsically affective and irreducible to the intentionality of any other mental states like perceptions or beliefs. Her chapter shows how approaches to affective intentionality can provide novel insights into current debates on the object-directness of emotions, on their cognitive bases, and on the relation between emotions and values.

The historical portion of the volume ends with Anna Kennedy and Keith Oatley's **Chapter 10: Emotion Theory in the First Half of the 20th Century**, which highlights eight theoretical approaches which thrived during the time period: introspectivist approaches, behaviorist approaches, evolutionary approaches, neurophysiological approaches, therapeutic approaches, aesthetic approaches, skeptical approaches, and everyday approaches to emotions. Kennedy and Oatley discuss seminal contributions to emotion theory by Titchener, Dunlap, Watson, Skinner, Shand, McDougall, Tolman, Hull, Cannon, Selye, Papez, MacLean, Beebe-Center, and Duffy. Their analysis reveals that the first half of the 20th century was an exceptionally rich portion of the history of emotion theory, marked by prescient, but often forgotten, contributions to the field. Their overview concludes in the mid-1950s, which is roughly the dividing line between Parts I and II of this volume.

Part II: Contemporary Emotion Theories

Two laments come to the lips of contemporary emotion theorists on a regular basis, mine included. The first is that their favored theory is misunderstood by its opponents. The second is that it is hard to understand other people's theories – both with respect to *what* they say and *why* they say it. This state of affairs is detrimental to progress, because it prevents theorists from getting clear on what exactly they disagree about.

The aim of Part II of this volume is to offer a remedy, in the hopes of integrating theoretical viewpoints where possible and understanding where the real dividing lines in contemporary emotion theory lie. For the purposes of this exercise, emotion theories have been organized into different *research programs*. A research program is an overarching theoretical approach within which specific theories can be articulated (Lakatos 1970). It consists of two main elements: a set of non-negotiable "hard-core" assumptions, plus a range of "auxiliary hypotheses" designed to protect the hard-core assumptions from refutation.

For example, in Newtonian physics, the law of gravitation and the three laws of motion belonged to the hard core – you could not call yourself a Newtonian and give up on them. Many other auxiliary hypotheses, however, were modifiable and did get modified over time when empirical data threatened hard-core laws (e.g. hypotheses about the number of planets in the solar system, the way light propagates, and the nature of atmospheric refraction).

The first four chapters of Part II try to distinguish the main research programs on emotions in philosophy, psychology, neuroscience, and sociology. The reason for focusing on these four domains is that other disciplines tend to rely on the accounts formulated in philosophy, psychology, neuroscience, and/or sociology as they embark in their own investigations, without articulating proprietary theories of what emotions are. The point of reconstructing research programs is to figure out what the main theoretical options are in the contemporary debate, distinguishing within each family of theories between non-negotiable commitments and revisable assumptions that can be changed to deal with emerging challenges.

In **Chapter 11: An Overview of Contemporary Theories of Emotion in Philosophy**, Michael Brady introduces us to the main research programs in the philosophy of emotions. He begins by asking what the aims of a philosophical analysis of emotions may be, distinguishing between naturalized approaches which take their lead from empirical investigations and a priori approaches which seek to systematize ordinary platitudes about emotions. Brady then considers judgmentalism, arguably the most popular research program in philosophy throughout the 1990s, covering Solomon's, Nussbaum's, and Marks's strands of judgmentalism. His second target is perceptualism, the most influential contemporary heir to judgmentalism, which assimilates emotions with types of perceptions, considering formulations by Greenspan, Roberts, Tappolet, Prinz, and de Sousa, and then criticisms by Brady and Helm. Brady concludes by reviewing recent motivational theories in philosophy by Deonna and Teroni and by Scarantino, canvassing their possible costs and benefits.

In **Chapter 12: An Overview of Contemporary Theories of Emotion in Psychology**, Agnes Moors reconstructs four phases in the development of a psychological theory of emotions: the articulation of a working definition that relies on superficial properties of emotions, the transition to a scientific explanation, the empirical testing of the explanation, and eventually the development of a scientific definition. She distinguishes five primary research programs in psychology, each coming in different flavors: the evolutionary program, the network-based program, the evaluation-based program, the psychological constructionist program, and the social constructionist program. Moors then discusses basic tenets, supporting evidence, and main challenges for several influential psychological theories of emotions, including Ekman's

basic emotion theory, Lange's network theory, Scherer's appraisal theory, Moors's goal-directed theory, Barrett's and Russell's versions of the psychological constructionist program, and Mesquita's and Parkinson's versions of the social constructionist program.

Chapter 13: An Overview of Contemporary Theories of Emotion in Neuroscience by Stephan Hamann tackles research programs on emotions in neuroscience. Hamann traces the early history of affective neuroscience, focusing on trailblazing insights by Cannon, Bard, MacLean, and Papez and on the emergence of the limbic system idea. He then distinguishes between two subfields of affective neuroscience, one devoted to testing psychological theories of emotions, and the other devoted to articulating and testing emotion theories developed by neuroscientists themselves. Five prominent neuroscientific theories of emotion are discussed in the chapter: LeDoux's survival circuits theory, Panksepp's affective neuroscience theory, Damasio's theory of emotions as somatic markers, Barrett's constructivist theory, and Adolphs's functionalist emotion framework. Hamann concludes with a table comparing and contrasting neuroscientific theories in terms of several dimensions, including the main theoretical constructs they use, their assumptions about neural localization, and the type of empirical evidence they rely on.

In **Chapter 14: An Overview of Contemporary Theories of Emotion in Sociology**, Kathryn Lively introduces us to research programs on emotions in sociology. She points out that, although some of the founders of sociology like Marx, Weber, and Durkheim prominently discussed emotions in their work, the emergence of emotions as distinctive objects of sociological investigation is a more recent affair. She traces it back to the late 1970s, when Hochschild offered a field-defining analysis of the emotional labor involved in managing one's own emotions in professional contexts (e.g. flight attendants are expected to convey positive emotions to customers). Lively distinguishes between cultural approaches to emotion management formulated in the wake of Hochschild's analysis, Kemper's and Bales's micro structural approaches to power and status dynamics in social interactions and groups, House's social structure and personality approach, Pearlin's stress process model, Heise's affect control theory, and Collins's interaction ritual theory.

Chapters 15, 16, 17, and 18 mark a transition from the high-altitude perspective of the first four overview chapters to a granular analysis of the four currently most influential research programs in affective science: basic emotion theory, appraisal theory, psychological constructionism, and social constructionism. I invited some of the leading proponents of these four approaches to explain how each research program understands emotions, what evidence supports the research program, and what are the main dividing lines and outstanding challenges within each research program.

In **Chapter 15: Basic and Discrete Emotion Theories**, Michelle Shiota introduces us to the contemporary literature on basic and discrete emotion theories, tracing their inspiration back to Darwin's work on emotional expressions. Shiota distinguishes a caricatured version of basic emotion theory she thinks is often incorrectly identified with the research program, from a more nuanced (and partially shifting) version articulated by Ekman over the decades. Shiota examines the existing empirical evidence on autonomic physiology, facial expressions, neural data, and coherence among components, noting where it supports Ekman's theory, where it is inconsistent with it, and where more studies are needed. She then discusses a group of "discrete emotion" theories that have emerged in recent years to deal with the critiques traditional basic emotion theory has received. Discrete theories preserve Ekman's assumption that emotions are evolutionary adaptations, but allow for a greater degree of variability in their physical manifestations. Shiota concludes by providing examples of the empirical evidence on context-dependent expressive and physiological changes that speaks in favor of discrete emotion theories.

In **Chapter 16: Appraisal Theories of Emotions**, Phoebe Ellsworth introduces the research program of appraisal theory, whose core assumption is that emotions are caused or constituted by evaluations of the meaning of the situation and its implications for well-being. Ellsworth discusses early pioneers of the research program like Arnold and Lazarus, and the impact of the now infamous epinephrine experiment run by Schachter and Singer in 1962, which appeared to give cognition a starring role in turning undifferentiated states of arousal into emotions. Several varieties of appraisal theory are examined in the chapter, including versions by Scherer, Ellsworth, Roseman, Frijda, Smith, Clore, and Collins. Ellsworth's discussion focuses in particular on seven of the most common dimensions of stimulus evaluation: novelty, intrinsic valence, certainty or predictability, goal conduciveness, agency, control, and compatibility with personal or social standards. Ellsworth concludes by distinguishing categorical vs. dimensional and causal vs. non-causal varieties of appraisal theory.

In **Chapter 17: Constructionist Theories of Emotions in Psychology and Neuroscience**, Lisa Feldman Barrett and Tsiona Lida discuss constructionist theories of emotions in psychology and neuroscience. One of the take-home messages of their chapter is that the label *psychological constructionism* – amply used throughout these volumes and in the literature – may be too narrow to capture the totality of the research program developed by Russell, Barrett, and others in opposition to classical theories of emotions. Barrett and Lida distinguish between two general approaches to the study of emotions: typology and constructionism. Typology holds that there are biologically and psychologically distinct emotion types whose physical signatures are to be studied empirically – this is the approach characteristic of basic emotion theories and (some) appraisal theories. Constructionism holds that emotions emerge dynamically from more basic ingredients, are extremely variable at the physical level, and are unified as members of the same category only by an act of "meaning making". Barrett and Lida then turn to a consideration of the empirical evidence for constructionism, which ranges from the predictive coding literature in neuroscience to the role of cultural inheritance in emotion knowledge. Barrett and Lida conclude by dispelling what they consider pervasive misunderstandings about the constructionist research program.

In **Chapter 18: Social Constructionist Theories of Emotions**, Batja Mesquita and Brian Parkinson introduce us to the research program of *social* constructionism. They begin by articulating the core assumptions of the program, which include the dependence of emotions on social cognition, the cultural specificity of emotions, their lack of passivity, and their social functionality. Mesquita and Parkinson examine several social constructionist theories, including Armon-Jones's proposal that emotions depend on interpretations shaped by socially constructed concepts, and Averill's proposal that emotions are transitory social roles. They then canvass various sources of evidence for the core assumptions of social constructionism, focusing especially on cultural differences in appraisal, on linguistic variability, on audience effects, on social functionality, on emotion categorization, on dynamic construction, and on the nature of emotional development in children. They conclude by discussing some of the ways sociocultural and evolutionary explanations of emotions could be integrated.

Chapters 19 and 20 focus on *research traditions* in emotion theory, which are larger-scale units of theoretical investigation than *research programs*. Whereas members of a research program must agree on the hard-core assumptions that define the program, members of a research tradition can belong to different research programs. What makes them members of the same tradition is that they endorse a central orienting assumption which directs and constrains the way emotions are conceptualized and studied. The two assumptions we focus on are that emotions are cognitions of some kind (cognitivist tradition) and that emotions are motivations of some kind (motivational tradition). Crucially, both assumptions span across

the divide between the philosophy of emotions and affective science, and so the two research traditions being considered here are interdisciplinary. Naturally, the diversity of viewpoints within each research tradition is much greater than the diversity of viewpoints within each research program. There are many ways to understand what cognition and motivation are, many ways to understand how they contribute – individually and in combination – to the instantiation of emotions, and many ways to understand what else is essential to emotions besides cognition or motivation. It is nevertheless fruitful to understand how research programs commonly considered to be competitors may share important commonalities at higher level of analysis, possibly as a preliminary step towards synthesizing them into a hybrid successor program.

In **Chapter 19: Cognitivist Theories of Emotions in Philosophy and Affective Science**, Bennett Helm discusses the cognitivist tradition in philosophy and affective science, defined by the assumption that understanding emotions requires making a central appeal to cognitions as their causes or constituents. Helm distinguishes two concepts of cognition: cognition as an information processing state (the notion prevalent in psychology), and cognition as a state with a mind-to-world direction of fit (the notion prevalent in philosophy). He then proceeds to consider how different cognitivist proposals fare with respect to a key set of theoretical desiderata, including their ability to account for object-directedness, rationality, differentiation, distinctness, phenomenology, causal roles, and importance. Helm's discussion covers a number of diverse research programs across disciplines, including belief-desire accounts (e.g. Marks, Reisenzein), judgmentalist accounts (e.g. Solomon), appraisal accounts (e.g. Lazarus), psychological constructionist accounts (e.g. Barrett), Neo-Jamesian accounts (e.g. Prinz), and perceptualist accounts (e.g. Tappolet). Helm concludes by unveiling some of the important insights shared by the diverse members of the cognitivist tradition.

In **Chapter 20: Motivational Theories of Emotions in Philosophy and Affective Science**, Andrea Scarantino discusses the motivational tradition, which is defined by the assumption that emotions are, or at least essentially comprise, motivations to act. Scarantino begins by reconstructing the emergence of the concept of motivation in early 20th-century psychology, suggesting that it was introduced to explain goal-driven behaviors that cannot be understood through chains of reflexes. Scarantino then summarizes pioneering motivational proposals developed in the 1960s by Leeper, Tolman, Simon, and Bindra, before turning his focus to seven contemporary accounts which understand the connection between emotion and motivation in different ways. In particular, he discusses the motivational sides of basic emotion theory and appraisal theory, comparing and contrasting Ekman with Plutchik in the basic emotions camp and Scherer with Roseman in the appraisal camp. Scarantino then discusses Frijda's seminal contributions on the nature of modes of action readiness and control precedence, including Frijda and Parrott's recent theory of ur-emotions. The chapter ends with a comparison between two motivation-focused proposals in the philosophy of emotions, Deonna and Teroni's attitudinal theory and Scarantino's motivational theory.

Part III: The Elements of Emotion Theory

Parts I and II jointly make an incontestable, if sobering, case that the community of emotion researchers has yet to coalesce around a shared understanding of what emotions are. At the same time, Parts I and II show that there has long been strong consensus on the main elements out of which emotions are composed. The notion of element at work here is very broad – if we think of emotions as recipes, the elements are all the ingredients or components which can feature in recipes. Some elements have been acknowledged as central since Ancient Greece,

most significantly experiences/feelings/qualia (often valenced, often bodily, often activating), action tendencies/desires/impulses, and appraisals/evaluations/cognitions. Other elements have become prominent more recently, including expressions/displays/manifestations in the 17th century with Descartes and Le Brun, peripheral nervous system changes in the 17th century with Descartes and Burton, and central nervous system changes in the 19th century with James and Hughlings Jackson.

The objective of Part III is to shift the focus of theoretical analysis from the emotions to the elements themselves. The elements investigated are ontology, physiology, brain circuity, expressions, actions, conscious experience, valence, and arousal (appraisal is another element of emotions, but since it is the primary focus of appraisal theory, it is discussed at length by Ellsworth in Chapter 16, Part II). The first objective in breaking down emotions into their core ingredients is to take stock of what we have learned about the main ways to understand each ingredient, about the subtypes of each ingredient, about their mechanisms of operations and about their biological and social functions. The second objective is to determine whether learning more about each individual emotion ingredient can help us choose between competing theories of emotions, or at least provide constraints that any viable theory of emotions must satisfy.

Chapter 21: Are Emotions Events, Processes, States, or Dispositions? by Hichem Naar asks a crucial but often neglected question any emotion theorist needs to ask, namely: What is the ontological status of emotions? In other words, what kinds of things are emotions at the most abstract level? Naar argues that there are two main ontological options, each with costs and benefits: emotions can be things that occur (i.e. occurrents) or they can be things that persist (i.e. continuants). Each option comes in two varieties. If we think of emotions as *occurrents*, we have to decide whether they are *events* that occur *at a time* or *processes* that unfold *over time*. If we think of emotions as *continuants*, we have to decide whether they are *states* or *dispositions*. After having clarified these four primary options, Naar considers some of the implicit ontological assumptions of basic emotion theory, appraisal theory, and psychological constructionism. He argues that affect programs have been understood by basic emotion theorists as occurrents, continuants, or a combination of the two; that appraisals have been understood by appraisal theorists as occurrent processes; and that emotional experiences have been understood by psychological constructionists as events, exploring some of the implications of these ontological claims.

In **Chapter 22: Is Emotion Physiology More Compatible with Discrete, Dimensional, or Appraisal Accounts?**, Bruce Friedman and Julian Thayer take stock of a debate on emotional physiology that James started 130 years ago when he suggested it was possible for each emotion to have a unique bodily reverberation. Friedman and Thayer begin by clarifying the structure of the autonomic nervous system (ANS) and of the cardiovascular, respiratory, and electrodermal measures typically used by psychophysiologists to take stock of emotions. They then take us back to the James-Lange vs. Cannon-Bard dispute on visceral emotional changes as a springboard for examining the degree to which physiological data support, respectively, basic emotion theory, dimensional approaches to emotions in terms of valence and arousal, and appraisal theories. Their conclusion is that every theory can find *some* support in the physiological data, but no theory is compatible with all the data. The final section of their chapter illustrates how a dynamic systems theory approach may offer an integration of the insights of competing emotion theories and lead to novel empirical hypotheses which cannot be formulated within any of the traditional emotion theories.

In **Chapter 23: Can Brain Data Be Used to Arbitrate Among Emotion Theories?**, Philip Kragel, David Sander, and Kevin LaBar examine whether neuroscientific data can used to arbitrate between alternative emotion theories. They first explain what types of neural data

are available, distinguishing between electrocorticography (ECoG), electroencephalography (EEG), functional magnetic resonance imaging (fMRI), functional near-infrared spectroscopy (fNIRS) and lesion studies. Their primary focus in the chapter is on fMRI evidence, in an attempt to demonstrate how univariate and multivariate neuroimaging approaches have contributed to the study of valence and arousal (dimensional theories), to the study of cognitive appraisals (appraisal theories), and to the study of discrete emotions (basic emotion theories). Their conclusion is that neural data alone cannot settle which theory of emotion is to be preferred, in part because not enough efforts have been made to increase the granularity of neural predictions and to overcome confirmation bias.

In **Chapter 24: What Do Nonverbal Expressions Tell Us About Emotion?**, Disa Sauter and Jim Russell introduce us to the literature on the nonverbal expression of emotions. They first explain how expressive bodily movements are conceptualized by competing research programs. For basic emotion theorists, expressions are effects of the activation of affect programs; for appraisal theorists, emotional expressions are effects of appraisals; for psychological constructionists, emotional expressions are effects of domain-general processes not specific to emotions. After having canvassed the evidence in support of each of these three approaches to emotional expressions, Sauter and Russell conclude that the evidence is inconclusive and that a winner cannot be declared just yet. They sketch the outlines of a possible integrative path that may help move the field forward, focusing on the need to understand emotional expressions as lying on a continuum between biological preparedness and cultural learning. Their view is ultimately that emotional expressions do not constitute a unified subject matter for emotion theory, because they differ in terms of their origins, current functions, and mechanisms of operation.

In **Chapter 25: Which Emotional Behaviors Are Actions?**, Jean Moritz Müller and Hong Yu Wong ask where we should draw the line between behaviors that happen to the emoter (e.g. crying) and behaviors that the emoter intentionally performs (e.g. apologizing). They propose that a behavior counts as an *action* just in case it is voluntary, goal-oriented, consciously performed, and done for reasons. This model of action is enshrined in philosophy by the so-called Humean theory of action, according to which actions are brought about by desire and belief pairs which rationalize them. Müller and Yu Wong argue that not all behaviors prompted by emotions are actions in this technical sense. For example, *basic bodily expressions* like frowning out of anger, and *reflex-like emotional reactions* like recoiling out of fear of a suddenly looming object are neither voluntary nor done for reasons. The chapter focuses on a class of behaviors whose agency is highly contentious, namely what they call *complex expressive behaviors* such as jumping up and down out of joy or affectionately ruffling another's hair out of love. Hursthouse, who described these behaviors as *arational actions*, famously argued that the Humean theory of action cannot explain them, because they are neither instrumental nor caused by a belief-and-desire pair. Müller and Yu Wong consider several competing interpretations of complex expressive behaviors which try to show either that they fit the Humean model with some modifications (e.g. Smith), or that they do not fit the Humean model but are nevertheless instrumental (e.g. Scarantino and Nielsen), or that they do not fit the Humean model and have a non-instrumental point (e.g. Bennett, Helm).

Richard Dub in **Chapter 26: Emotional Experience: What Is It and What Is It For?** focuses on the nature and functions of emotional experience. Dub begins by introducing working definitions of state consciousness, phenomenal consciousness, access consciousness, and emotional phenomenology. He then focuses on three broad theories of emotional experience, discussing their respective advantages and disadvantages. *Somatic theories* ground emotional

experience in bodily changes, and they exist in Jamesian and non-Jamesian varieties. *Cognitive/perceptual theories* take emotional experience to be modeled either after the phenomenology of cognitive states like thinking or judging, or after the phenomenology of perceptual states like seeing or hearing. *Central theories* propose that emotional phenomenology is irreducible to any other kind of sensory, cognitive, or perceptual phenomenology. Dub then turns to the question of what the functions of emotional consciousness might be, discussing its possible adaptive roles in facilitating flexible action, globally broadcasting content, storing memories, transferring information, facilitating the understanding of values, and solving control dilemmas. Dub concludes by examining whether we should foreclose the possibility of unconscious emotions on conceptual grounds, and, if not, what evidence there might be for unconscious emotions.

In **Chapter 27: How Should We Understand Valence, Arousal, and their Relation?**, Giovanna Colombetti and Peter Kuppens discuss the central constructs of valence and arousal, each of which can refer to multiple notions not always clearly distinguished from one another. They discuss three main concepts of valence: *experienced valence* (how pleasant or unpleasant an emotion feels), *behavior valence* (the direction of an organism's emotional behavior toward or away from something) and *appraisal valence* (the positive or negative character of emotion-eliciting appraisals). They then consider the notion of arousal, distinguishing between two main concepts of arousal: *experienced arousal* (feeling activated) and *physiological arousal* (sympathetic arousal or brain arousal). Colombetti and Kuppens argue that it is a mistake to conflate these different notions of valence and arousal, that emotions often have mixed valence in several senses of valence, and that the notions of experienced and physiological arousal tend to be too narrowly defined in the literature. Finally, they examine what the relations between valence and arousal might be under various ways of understanding each term, considering the evidence for their independence, for a linear relation between them, for a symmetrical V-shaped relation, and for an inverted V-shape relation, concluding that the jury on these options is still out, and more evidence is needed.

References

Griffiths, P. E. (2004). Emotions as Natural and Normative Kinds. *Philosophy of Science*, *71*(5), 901–911.

Hacking, I. (1999). Making Up People. In M. Biagioli (Ed.), *The Science Studies Reader* (pp. 161–171). New York: Routledge

Lakatos, I. (1970). Falsification and the Methodology of Scientific Research Programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the Growth of Knowledge* (pp. 91–195). Cambridge, UK: Cambridge University Press.