Embodied Emotions

In this book, Rebekka Hufendiek explores emotions as embodied, action-oriented representations, providing a noncognitivist theory of emotions that accounts for their normative dimensions. *Embodied Emotions* focuses not only on the bodily reactions involved in emotions but also on the environment within which emotions are embedded and on the social character of this environment, its ontological constitution, and the way it scaffolds both the development of particular emotion types and the unfolding of individual emotional episodes. In addition, it provides a critical review and appraisal of current empirical studies, mainly in psychophysiology and developmental psychology, which are relevant to discussions about whether emotions are embodied as well as socially embedded. The theory that Hufendiek puts forward denies the distinction between basic and higher cognitive emotions: all emotions are embodied, action-oriented representations. This approach can account for the complex normative structure of emotions, and shares the advantages of cognitivist accounts of emotions without sharing their problems. *Embodied Emotions* makes an original contribution to the ongoing debates on the normative aspects of emotions and will be of interest to philosophers working on emotions, embodied cognition and situated cognition, as well as neuroscientists or psychologists who study emotions and are interested in placing their own work within a broader theoretical framework.

Rebekka Hufendiek is a postdoctoral candidate at the University of Basel. Her research focuses on philosophy of mind and psychology with a particular interest in embodied cognition, emotion theories, and naturalism. She has written several papers and reviews on embodiment, emotions, the modularity of mind, and the vices and virtues of naturalist approaches to the mind.

Routledge Studies in Contemporary Philosophy

Embodied Emotions

A Naturalist Approach to a Normative
Phenomenon

Rebekka Hufendiek

For George Thimble

Contents

x   *Contents*

Preface

This book is about emotions. Emotions are sometimes heated and sometimes clumsy, but they are always skillful, embodied responses to the situations in which we find ourselves. They shake our bodies and motivate us to action. Instead of being rationally processed strategies for what to do, they offer rough and ready solutions to problems we are faced with. Emotions are complex whole-organism responses that are shaped by evolution and by culture alike. But although they are complex and skillful, they are often rather heavy-handed responses that conflict with our rational goals. That does not mean that emotions lack a representational or normative dimension. It rather means that the norms that our emotions respond to need not be the norms that we rationally endorse.

This book originated from my dissertation that I submitted and defended at Humboldt University in 2012, and it is the result of a long process of tinkering. It has been in the works for many years and I am grateful to many people for feedback, help, and support. I would first of all like to thank my supervisors Dominik Perler and Markus Wild for all the advice and encouragement they have offered over the years and—more than anything—for their patience.

While working on the dissertation I was a member of the Collegium for the Advanced Study of Picture Act and Embodiment at the Humboldt University of Berlin. The collegium offered a very inspiring environment and provided many opportunities to discuss issues around embodied cognition and emotion theory with colleagues and guests. I am grateful to Horst Bredekamp and John Michael Krois, who founded the Collegium. John Krois passed away unexpectedly in 2010. He is sorely missed as a supervisor and friend who originally introduced me to the debate on embodied cognition.

While turning the dissertation into a book I was a member of a project on "Biosemantics and Normative Pragmatism" at the University of Fribourg and Basel, funded by the Swiss National Science Foundation and led by Markus Wild. The research group provided room for discussing various aspects of naturalist semantics and normativity that were highly inspiring for the present work.

While I was working on the book I had the opportunity to discuss parts and chapters of it with many people. I am especially grateful to Saray Ayala, Lee Chichester, Anna Ciaunica, Jason Clark, Joshua Crabill, Shaun Gallagher, Dan Hutto, Thomas Jacobi, Lena Kaestner, Michael O'Leary, Jesse Prinz, Andrea Scarantino, Jelscha Schmid, Christine Sievers, Pietro Snider, Achim Stephan (and the Animal Emotionale Project in Osnabrück), Fabrice Teroni (and the Thumos-group at the Center for Affective Sciences in Geneva), Patrizia Unger, Sven Walter, Anna Welpinghus, and an anonymous referee.

Special thanks for reading, rereading, and re-rereading this work over the years and for permanent feedback and support in the most urgent moments goes to Joerg Fingerhut, Stephan Schmid, and Julia Staffel. This book would not be what it is without you.

Last but not least I would like to thank friends who are family and family who are friends. You know who you are and how much I owe to you.

Introduction

Consider a typical example of an emotion such as pride. What are the essential elements of pride? What does it mean to be proud? Pride is a pleasant feeling, one could reason, and there is always something we are proud *of*, namely, ourselves or closely related people. Things get more complicated if we unpack the cognitive dimension of pride a little further. Pride, it seems, is not simply directed at the self. Instead, pride always concerns properties that are "praiseworthy," and these properties must be ascribed to the self. This suggests that pride cannot be described as being a simple feeling or a reflex-like behavior. Rather, pride, it seems, should be understood as a complex cognitive evaluation that essentially involves conceptual reasoning. In this vein, Donald Davidson (1976)[1] suggests that pride is the judgment, "I am praiseworthy!" drawn from two premises, where the first premise is a judgment that everyone who exemplifies a certain property is praiseworthy and the second is a belief that one exemplifies that property oneself (or that somebody closely related or identified with exemplifies that property).

With this description, a cognitivist account such as Davidson's captures the semantic, evaluative, and rational dimension of pride. The semantic dimension of pride consists in our calling someone's pride appropriate when the two premises are fulfilled. Of course, we can err on whether we really have the praiseworthy property in question. Yet the mere fact that we can reasonably discuss whether pride is appropriate or inappropriate in such a manner suggests that pride has semantic adequacy conditions. The evaluative dimension of pride further consists in the ascription of a value to oneself. Pride would not be pride if a proud person would not take herself (or a closely related person or an object/institution she identifies with) as having achieved something that is of value. With regard to these norms we can certainly disagree about which properties are praiseworthy and which are not. But again this only shows that an emotion such as pride is in an interesting sense about things of value. Finally the rational dimension of pride consists in the possible relations in which pride can reasonably stand to other emotions and to other mental states. We can for example reasonably expect somebody who is proud of the virtuous behavior of her sister to be disappointed and maybe even ashamed if the sister turns into a corrupt fraud who appears in

the press every once in a while for having bribed somebody. We cannot reasonably expect her to feel joy or any positive emotion about an event that she evaluates as negative. Most likely she will be surprised if this change in manners comes unexpectedly, and angry if she blames her sister for voluntarily disrespecting the values she has been taught to respect, and depending on the degree to which she identifies with her sister, her pride might turn into shame.

A cognitivist account such as Davidson's can capture this normative dimension of an emotion such as pride. However, it raises several questions as well. If pride is essentially the judgment that one is praiseworthy, how can we account for the obvious differences between emotions and judgments? What role does the pleasant feeling or the erect body posture that is associated with pride play? If pride is the result of cognitive evaluation, why do we often feel proud for reasons that we would judge to be stupid under closer scrutiny? And, if pride presupposes such a demanding logical inference, how can it be ascribed to infants? After all, behavioral and expressive elements that look much like pride can be observed in infants long before they master conceptual reasoning. This is captured in a quote from an interview with the mother of an eight-month-old boy:

> He's rather pleased with himself now that he can crawl and pull himself up on the furniture and he'll crawl across the room and pull himself up on the toy box . . . And then turn round and bang on the toy box and turn round and look at you and give you a big smile as if to say look at me. Look at what I've done.

<div align="right">(Reddy 2008, 138)</div>

A steadily growing amount of evidence from developmental psychology, ethnology, and behavioral studies in animals suggests that expressive patterns and bodily postures associated with emotions such as pride, jealousy, embarrassment, and shame might be present in apes, infants, and across cultures.[2] This casts into doubt cognitivist accounts such as Davidson's, insofar as infants and animals are obviously not able to conduct the inferential reasoning that Davidson takes to be constitutive of a pride reaction.

Not only theories about the nature of pride but also theories of emotion in general can be coarsely divided into two camps: cognitivist and noncognitivist accounts. Cognitivist accounts focus on the cognitive abilities that seemingly underlie emotions such as guilt, shame, and pride. Emotions, according to these authors, are intentional and normatively assessable, which requires us to see them as involving complex representations or conceptual reasoning. On the other hand there are noncognitivist accounts that focus on the early occurrence of emotions such as fear, anger, and joy, their presence in other animals, and their bodily foundations in panculturally present facial reactions, bodily postures, visceral responses, and neural processes. These accounts tend to give biological explanations of emotions that highlight the role of feelings, the motivating potential of emotions, and their

close connection to reflex-like behavior. What is notoriously lacking in these noncognitivist accounts is a comprehensive explanation of the normative dimension of emotions. This gives these accounts an unpleasant reductionist flavor, in the sense that they end up with poor and inadequate descriptions of the phenomenon.

What is it about emotions that makes them such a tough case, which resists an integrative explanation that brings the insights of cognitivist and noncognitivist approaches together? Emotions are notoriously difficult to categorize, and they seem to cross borders between categories that philosophers traditionally have wanted to separate, like body and mind, nature and culture, rationality and irrationality. They are usually triggered automatically and involve bodily arousal, but at the same time they seem to be cognitively demanding operations that are concerned with our well-being and involved in moral reasoning. Philosophers have compared emotions to judgments, to account for the fact that emotions can be intentional, cognitively complex, and normatively assessable. Yet emotions, as they occur in infants, animals, and many situations in human adult life, seem to be much less articulated and cognitively demanding than judgments. Philosophers have compared emotions to perceptions as well, to account for the fact that emotions are present in infants and animals and to explain their cognitive impenetrability. But emotions seem to be evaluative and motivating in a way that perceptions are not. Furthermore, emotions have traditionally been described as having their origins in evolution. Darwin observed that apes, horses, and even bees show behavioral patterns that look much like human anger expressions. But there are also approaches—mainly from ethnology and the social sciences—that describe emotions as social constructs that radically differ from culture to culture and offer a rich number of examples of highly culture-specific emotions.

In trying to explain emotions, philosophers and scientists have frequently met the same dilemma again and again: in order to capture the phenomenon of pride, for example, in an adult adequately, and in order to unpack and make explicit the kind of representation or evaluation that seems to be at play when somebody is proud of her well-maintained house, it has often been assumed that emotions are complex representations that entail conceptual content themselves. But such an approach notoriously overintellectualizes the phenomenon, such that the phenomenal, motivating, and bodily elements tend to be forgotten. Noncognitivist approaches that focus on emotions as behavioral reactions or feelings, on the other hand, tend to reduce the phenomenon to a simple input–output reaction, leaving its intelligence and normative assessability unexplained. The dilemma for emotion theorists therefore is that depending on whether one chooses a cognitivist or a noncognitivist approach, the resulting theories are either overintellectualizations or inadequate reductions of the phenomenon to be explained.

This dilemma, on closer examination, is the result of two difficulties. One obvious difficulty in emotion theory is to develop an approach that tells a

unified story about the relation of mind and body in emotions. Such a story would have to account for the fact that emotions are obviously intelligent ways of interacting with complicated social scenarios, while respecting the fact that emotions involve a bodily dimension that is more than just a coincidental by-product of a certain type of mental representation. A second obvious difficulty for a theory of emotions is to account for the similarly exigent question of how to capture the relationship between nature and nurture. Emotions obviously involve certain features that have their origins far back in the history of evolution, but they are also shaped by culture and social context to a high degree. Put very generally, accounts that highlight the role of mind and culture tend to overintellectualize while accounts that highlight the role of body and evolution tend to be inadequately reductionist.

The problems that arise when trying to come up with an integrative account that explains the bodily, cognitive, natural, and cultural sides of emotions has led some authors to the conclusion that emotions should simply not be described as belonging to one single category. Paul Griffiths prominently developed the so-called disunity thesis arguing that folk psychology collects various psychological phenomena under the label "emotions," which should be held apart by the sciences because they do not form a natural kind (Griffiths 1997). Griffiths accordingly distinguishes between evolutionarily acquired basic emotions that show a certain bodily and behavioral profile and higher cognitive emotions that differ among cultures and presuppose certain cognitive abilities.

It is certainly true that there are prototypical examples of emotions that appear to be more "hardwired," like fear as a response to loud noises or disgust as a response to the smell of rotten food, and other examples of culture-specific and cognitively demanding emotions, like feeling guilty for not going to work or being proud of one's country. However, current empirical evidence from various disciplines does not speak in favor of the strict distinction that Griffiths suggests. What has been said about pride so far already gives us reason to be suspicious whether pride could simply be described as a "higher cognitive emotion," since recent evidence suggests that pride occurs early in infancy and is associated with a bodily posture that occurs across cultures. I therefore think of the differences between different emotion-types in a developmental continuum and not in terms of two distinct classes. Furthermore, I highlight that the complexity of a reaction does not imply that it must be *cognitively* complex. Bodily reactions that have their roots in evolution can realize very complex forms of intelligent behavior without being cognitively complex or involving complex mental representations. Furthermore, that an emotion has its roots in evolution does not imply that it is not shaped by culture and social context to a high degree.

Put roughly, the theory I develop takes emotions to be embodied, action-oriented representations that are embedded in a social context. Instead of reducing emotions to feelings or behavioral reactions, I show

how the complex representational structure that cognitivist accounts ascribe to emotions can be translated into the complex interaction between bodily skills and the structured social environment. The normative dimension of emotions cannot be explained with reference to what is in the head alone, but only when combined with reference to the complex interplay between bodily reactions and the social environment. This approach avoids both overintellectualization and inappropriate reductionism.

The present approach is fundamentally naturalist in its aim insofar as it takes mindfulness to be something that emerged incrementally in an evolutionary process, starting with simple organisms that had to successfully interact with their environment to survive. The reason that we think of emotions as a place where mind and body closely interact is precisely because emotions evolved as more complex and plastic mechanisms from simple reflexes and homeostatic reactions. Emotional episodes range from the fear response of a baby that hears an unexpected loud noise to an adult getting more and more nervous while thinking about an upcoming exam. I argue that an account that takes emotions to be embodied, action-oriented, and embedded best makes sense of the huge variety of emotional episodes without inadequately reducing the phenomenon and without abandoning a naturalist framework. In what follows, I introduce biological naturalism and the framework of embodied cognition on which my approach relies.

1. NATURALISM

The term *naturalism* is used in a variety of ways; it is a red flag to some and a *sine qua non* to others. I therefore explain in detail which version of naturalism I commit myself to and why I think that most standard objections and prejudices against naturalism do not apply to this version. The most general naturalist claim is that a theory requires that philosophical explanations be consistent with scientific explanations of the phenomenon at issue. This general claim entails an ontological thesis, that the world comprises nothing but material objects constrained by natural law, and a related methodological claim, that since all facts are natural facts, the methods by which we investigate these facts must be suitable to such an investigation (Papineau 1993, Ladyman, Ross 2009). In this broad sense, most naturalists would agree on the explanatory claim as well as its ontological and methodological implications. Yet if these two general claims are explained in more detail, disagreement arises and naturalists split into different camps.

Naturalist approaches to the mind are often thought of as being reductionist in a strong sense. Strong reductionists claim that the relation between natural sciences and higher-level domains is one of deduction. Facts about a living organism can, in principle, be deduced from physical knowledge about that organism; facts about mental states can be deduced from neuroscientific data. That different explanatory levels can, in principle, be reduced

to each other, according to strong reductionism, has an ontological explanation: the biological world contains nothing that is not physical and the mind entails nothing that is not realized by neural processing.[3] Examples of such reductionist views are early mind-brain identity theories. Mind-brain identity theories claim that every mental state or property is identical to a physical state or property (see, e.g., Place 1956, Smart 1959).

Naturalism is not bounded by such strong versions of reductionism, though. On the contrary, there are good reasons to deny reductionist claims. The ontological claim that there is only natural stuff and natural laws in the world does not entail the claim that there is only one level of description to which all (mental) phenomena can be reduced and by which all (mental) phenomena can adequately be described. A naturalist picture can entail the view that there are lower and higher orders of organization that should not be reduced to each other and need different methods to be described adequately. Such a claim can remain on the explanatory level and simply argue that it makes no sense to reduce, for example, biological explanations to physical ones for reasons of complexity, yet it can also be argued that biological laws, or regularities, are real and differ in the way they structure the world from physical laws. The two views come up with different arguments against strong reductionism.

The claim that the natural world might consist of several "layers" that demand different scientific methods to be adequately explained raises many questions, though. There are plenty of models from various disciplines suggesting different levels of explanation. It could be reasonable to assume a computational level in addition to a task-description level and an implementation level, as notably suggested by David Marr (1982) for psychology. Biological approaches can reasonably distinguish between an implementation, or anatomical, level and a task-description level as well, yet it has been suggested that we should assume a further abstract level to explain evolutionary outcomes by the generalizations of population genetics and evolutionary game theory (Griffiths 1997, 217; see also Craver 2007). I do not discuss in detail here what an integrative model of explanatory levels for biological and psychological approaches to emotions should look like. In what follows I argue that emotions are embodied action-oriented representations shaped by biological and social processes. To account for these processes, biological traits and regularities, like social objects and dynamics, will have to be described in their own right. To neatly describe the way in which emotions are embodied, the physiological level of implementation must be considered. Making sense of the claim that emotions are action-oriented requires a look at the task-description level and the ascription of biological or social functions to particular emotions. From a pragmatic point of view cognitive science needs methodological pluralism, different levels of description, and ways to describe how these levels can interact as closely as they do. I tend to think that strong reductionist approaches are not helpful for such a project, but I do not argue for this view here. Instead, I present three objections

that have been brought forward against strong reductionist approaches and argue why they do not apply to the naturalist view that I develop with regard to emotions throughout this book. I thereby leave open the question of whether other versions of naturalism could account for these objections as well.

Strong reductionism is usually a form of physicalism. Physicalism claims that the world contains only entities and regularities that can be explained by physics, since everything is physical, supervenes on the physical, or is necessitated by the physical. Physicalists tend to think of the world as a four-dimensional entity with physical properties and relations instantiated in various points or regions within it. Physicalism has been paradigmatic in the philosophy of mind from the 1950s on. The early mind-brain identity theories discussed above, for example, usually embrace physicalism. Common objections against physicalism include (1) physicalism is restricted to explanations that are in accordance with natural laws, while deliberate reasoning and spontaneous action seem to obey different rules; (2) physicalism is committed to the claim that everything can be reduced to the implementation level and cannot account for multiple realizability and functions; and (3) physicalism is restricted to the description of facts. Mental processing, however, fundamentally includes a normative dimension. Representing things as true or false, appropriate or inappropriate, good or bad, pretty or ugly presupposes norms with regard to which something is true or false, and so on. These objections show that physicalism produces a gap between the natural and the mental world in the first place that frequently leads to reductionist views with regard to the logical, functional, or normative dimension of the mental. All three objections can also be turned against a physicalist theory and be used in favor of a biological approach to emotions since (1) taking emotions to include incrementally evolved reactions implies seeing them as governed by biological regularities rather than by physical laws, (2) emotions evolved differently in different species such that their multiple realizability needs to be explained on a functional level, and (3) emotions do have a normative dimension, since they respond to things that are good or bad for the organism's well-being. While physics cannot account for such normative notions, they are part and parcel of biological explanations.

With regard to the first point, the problem with physicalism arises from a strict understanding of what natural laws are. Natural laws (according to positivist and empiricist definitions) are empirical generalizations that are to be understood as strict, that is, without exceptions. Applying such a strict view of natural laws to the mental entails many problems. The main problem I want to highlight here is that many other disciplines from the cognitive sciences cannot reasonably use such a strict conception of natural laws. Evolutionary biology certainly plays an important role in the explanation of the mental. It is the only scientific discipline that can give reasonable explanations of *why* certain traits evolved and is therefore essential when it comes to explanations on the task-description or design level. But evolutionary

biology operates with parameters, such as selection pressure, that cannot be described as laws in a strict sense but rather as "law-like regularities" since evolution itself is a contingent historical process. Evolutionary theory is a historical science. The outcomes of evolutionary theory are steadily affected by historical accidents. Furthermore, the historicity of evolution implies that adaptive forces are not sufficient to predict what will occur or explain what has occurred (Griffiths 1997, 71f.). The parameters of evolution are not fundamental constants, as are the parameters of physics, since they lack projectability. When a physical parameter is measured accurately once, it need not to be measured again. The most important parameter values in evolution, such as strengths of selection, mutation rate, and migration rate, are not at all constant. Even when measured adequately once at a certain place and time, they need to be constantly measured again during ongoing research. Therefore, with regard to evolution, one should speak of regularities rather than strict laws (Brandon 1997).

The second objection is that physicalism is restricted to the implementation level and cannot account for multiple realizability, that is, the fact that types of mental states can be realized in different ways in different organisms and even at different times in the same individual. Physicalism is committed to the reductionist position that claims and principles from other sciences can, in the end, be reduced to claims and principles from physics. Yet what I suggested earlier is that research on multiply realized mental phenomena works best when operating on a task-description level. The stable patterns that can be observed on this level are produced by regularities that cannot be captured adequately on a physical level of description. Again, the biological naturalism assumed here supposes that the task-description level is not reducible to the implementation level in the sense that if you would know all the facts about a certain situation as it can be described on the physical level you would not be able to make predictions about entities that need to be described on the biological level like living organisms. The selection processes that are responsible for the reproduction of certain traits that fulfill certain functions cannot be captured on the physical level; they must be taken as irreducible aspects of reality, at least as long as we have no idea how to even capture them in physical terms. The functional level and the implementation level are not independent, though. Evolutionary processes are best described by paying attention to the details of concrete anatomical realizations of certain traits, while at the same time taking into account that categories of certain traits reproduce because they happen to fulfill a certain function for the organism. I argue for an account that integrates the implementation and the task-description levels in Chapter 2.

The third objection, that physicalism cannot account for the normative dimension of the mental, is, once more, a worry that is not justified with regard to bio-naturalism. What is usually meant by the normative dimension of the mental in a broad sense is that (1) mental processes meet certain norms, such as being true or appropriate, and (2) reasoning is normative in

the sense that we take certain things to be true, good, or pretty and others not, where many would argue that being true, good, or pretty are not features of the things we represent themselves. The world of the mental seems to add something here, which cannot be spelled out in scientific terms. Now, depending on what we take norms to be, it might not be true that no scientific discipline has anything interesting to say about the normative dimension of the mental. Evolutionary biology, again, is a scientific discipline that presupposes that well-being is a basic value for an organism. Related to this assumption are further assumptions, such as the claim that the organism is organized in a functional way where the different parts of the organism maintain the organism's well-being. Furthermore, the environment of the organism entails certain things that are of value for the organism in the sense that they can help to maintain its well-being. A broad notion of a norm claims a norm to be a standard from which actual facts can depart. Such a broad notion is not equivalent to statistical norms or simply an average. Biological standards always entail a claim about how something should be with regard to the well-being of the organism.[4] But the notion is not restricted to prescriptions or social rules that tell members of a social community what they ought to *do*. Instead, it also includes claims about how things ought to *be* with regard to the organism's well-being. If we hold such a broad notion of normativity then biology can add much to an understanding of the normative dimension of the mind. With regard to the normative dimension of emotions, I develop an account in Chapters 4 and 5.

To account for the normative dimension of emotions, I rely on several claims found in biosemantics and ecological psychology. Ecological psychology explicitly claims that values can be derived from biology and uses this claim for the theory of affordance perception. In laying out the principles of his "ecological approach to perception," James Gibson has argued that the environment of an organism cannot reasonably be described as an ensemble of objective facts independently of the organism and its relation to the environment. A description of the organism's environment is a description of the relationship between the organism and the things it is surrounded by. For example, it doesn't matter what size the stone in front of an animal is, it matters whether the animal can climb it or has to go around it; the precise chemical constitution of the berry in front of the animal does not matter, but whether it is nutritious or poisonous for the animal does. In this sense an ecological description of the way an organism is situated in its environment always includes a reference to external values; "physics may be value-free, but ecology is not" (Gibson 1986, 140). What follows from Gibson's claim is that an ontological approach, which sketches the environment in relation to particular organisms, has to make sense of the idea that things in this environment are of value for the organism. It is the value of something for an organism that explains its action-orientedness toward it. Ruth Millikan's biosemantics suggest that the adequacy or truthfulness of an emotion can be explained with reference to the function of the representational system in

question. It is a main aim of this work to develop an explanation of the normative structure of emotions in a naturalist context. I suggest that emotions are subject to semantic, rational, and social norms and that these norms can be explained by taking emotions to be embodied action-oriented representations and develop a Gibsonian ontology to account for the affordances that emotions represent.[5]

The preceding discussion makes clear that naturalism is not limited to physicalism and that there are good reasons to assume that biological traits develop and function in accordance with regularities that are not identical to physical laws. There is, however, a further worry with regard to naturalist accounts that applies to biological naturalism as well. Many people think that naturalism tends to reduce or minimize the role of culture or social influences for mental phenomena, in general, and for emotions, in particular. This is certainly true for many biologically inspired approaches, but not for the one defended here. I argue that emotions involve incrementally evolved homoeostatic reactions of an organism that are set up to regulate well-being. This implies that a huge portion of the emotional equipment is inherited and that emotions are more than just social constructions.[6] But this is not the whole story. First, genetic inheritance alone is never a sufficient explanation for a certain behavioral disposition; the genotype–phenotype relationship is complex and genes are parts of networks, which influence and interact with each other and with features of the environment. The relations between genes and behavior are not one-to-one, not even one-to-many—they are many-to-many. There is not one gene responsible for one type of behavior such as "aggressive behavior" or "cooperative behavior" (Greenspan 2001; see also Churchland 2011 for discussion). Furthermore, emotions are far more complex and plastic than basic regulation mechanisms such as simple reflexes or the regulation of metabolism.[7] If you think of emotions as reactions that involve an input and an output side, you find hardwired elements on each side but also room for the influence of culture and experience. Fear, for example, can be elicited by unexpected noises in newborns. Unexpected loud noises seem to be preconditioned stimuli that we are hardwired to react to. Yet emotions can be triggered by an infinite variety of stimuli, and the overwhelming majority of these stimuli are learned and thereby highly influenced by culture and personal experiences. When it comes to the output side, that is, to the variety of behaviors that can be triggered by an emotional reaction, there is strong empirical evidence for the pancultural occurrence of certain facial expressions, for example, in fear, sadness, and anger (Ekman 1971, 2003), and for particular body postures associated with pride and shame (Tracy and Robbins 2004, Clark 2010). Yet, again, this does not mean that the bodily expression of emotions, such as fear and shame, is not shaped by culture to a high degree. There are display rules that govern the modulation of emotional expression. In Western societies the adequacy of emotion expression is strongly related to gender roles. More than 90% of fear faces in Hollywood movies belong to female actors. Anger expressions

are often conceived of being more appropriate when shown by men, because of the higher power status ascribed to them, while women are supposed to smile more often in social conflict situations and thereby send signs of appeasement (Henley 1995, Averill 1997). Such regulation of expressions can vary among cultures. The culturally variable regulation of emotional output is not restricted to behavior, however. We can, for example, also consciously regulate respiration patterns and thereby immediately influence the autonomic system. People can learn to hyperventilate by breathing deeply in quick succession and thereby voluntarily produce panic or—the other way around—can learn to overcome states of panic by voluntarily controlling their breath. The goal of such training is to automatize breathing patterns in such a way that panic states occur less frequently. Most interesting in the breathing example is that respiration patterns can also be transmitted through cultural practices without the learning individuals being conscious of their embodying a cultural habit with an impact on emotional life (Lyon 1999, Laird 2007).

It is a central aim of this work to offer a theoretical framework that highlights and explains how emotions are shaped through social influence on the pattern of bodily reactions that constitute particular emotional episodes. To think of emotions as being exclusively either social constructs or biologically hardwired is obviously absurd and a position rarely found in current research. Emotions unfold in infants in the interaction with their social environment.[8] We can think of the infant as a skillful agent exploring its environment and of the environment as being inherited along with the genes. To inherit an environment does not mean to be thrown into the same portion of the world as our ancestors. Rather, it means that parents create very specific conditions for their infants that are the product of cultural heritage and rational deliberation. Infants can learn to interact successfully with such environments before they are able to understand them conceptually. This casts doubt on Davidson's interpretation of pride and comparable cognitivist approaches to emotions, as well as on classical conceptions of "higher cognitive emotions" (Griffiths 1997, Lewis 2014).

So far I have argued that naturalism is not equivalent to reductionism or physicalism and that an adequate approach to emotions should describe them as fundamentally biological and social phenomena. The following chapters unpack what this means in detail by bringing together empirical studies from various disciplines as well as conceptual and phenomenological considerations.

Two general rules should be kept in mind for this project: (1) Naturalist theories of the mind gain their plausibility not only from the fact that they rely on empirical data but also from their ability to integrate these data into a unifying theory that explains them in a broader framework of understanding the human mind or human nature. (2) Although such an integrative account needs to incorporate phenomenological insights and conceptual considerations with observations about the design of an organism

and details from the implementation level, and although there is no general rule for how to do so, there is a difference between accounts that start their explanation from a naturalist stance and theories that are primarily dedicated to conceptual analysis or phenomenological considerations.

2. EMBODIED COGNITION

Embodied or situated cognition has become a central research paradigm in current cognitive sciences. There is a growing commitment to the idea that the mind and human intelligence must be understood in the context of their relation to the body and interactions with the world (Wilson 2002, Robbins and Ayede 2009, Shapiro 2014). Human intelligence evolved from ancestors whose neural resources were largely devoted to on-line interaction with the world. "Intelligence" not only refers to conceptual or inferential reasoning but also to all the possible ways to interact successfully with the environment that are not simply reflex-like but rather entail learned or spontaneous reactions to all kinds of problems that we might be facing. Many of the groundbreaking ideas in this arena stem from a shift in light of new research on infant development (Thelen and Smith 1994), animal behavior, and artificial intelligence (AI) (Brooks 1991), when it became obvious that modeling the mind as something constituted by abstract symbolic processing alone leads to problems on the simplest levels of interaction with the real world and that biology has rather different solutions than early research on AI:

> The classical AI planning system can sit back and take its time, eventually yielding a symbolically couched description of a plausible course of action. The embodied planning agent must take action fast—before the action of another agent claims its life . . . The cockroach has a kind of common sense that the best current artificial systems lack . . . At root, our minds too are organs exquisitely geared to the production of action, laid out in local space and real time.
>
> (Clark 1997, 7f.)

Connected with the paradigm shift sketched here are further claims, such as the rejection of the classical notions of mental representation, symbolic processing, and the modularity of the mind. We actively access the world in perception via sensorimotor skills, not inner representational capacities (O'Regan and Noë 2001); complex inner representations are not needed to guide behavior since "the world is its own best model" (Brooks 1991); real brains do not seem to use linguaform, inferentially and syntactically structured information when spontaneously reacting to the outside world (Dreyfus 1991, Clark 1997, Chemero 2009); and the mind as a whole should not be seen as a system, where perception gains input, which is then cognitively processed and finally results in a behavioral output. This classical "sandwich

model of the mind" is denied by embodied approaches, and it is suggested instead that perception and action are closely intertwined and constitute several forms of direct on-line access to the world (Hurley 1998, 2001). These considerations about the situatedness of the mind constitute the background to my approach. Lawrence Shapiro (2010) distinguishes three main tasks for embodied or situated approaches: (1) they should explain whether certain cognitive processes are constituted by bodily or environmental processes and not just caused by them; (2) they should make clear whether and how an embodied approach can replace an explanation of a certain cognitive process in the traditional terms of complex inner representations by instead using an explanation of embodied skillful interaction with the environment; and (3) they should point out the extent to which our abstract conceptual reasoning is based on embodied schemas.

Throughout this book, I argue that emotions are embodied, action-oriented representations that constitute a nonconceptual form of skillful engagement with the social world we are embedded in. In the following section, I point out what I mean by the claims that emotions are "embodied," "action-oriented," and "embedded." I also clarify what constitution, conceptualization, and replacement mean in the context of emotion research and why I take replacement to be the most central and most neglected task for an embodied approach to emotions.

a. Embodied Emotions

The claim that emotions are embodied, roughly speaking, is that emotions involve bodily reactions and that these bodily reactions realize, or constitute, a kind of intelligent behavior, or interaction, with the environment. Talk about "embodied emotions" has become popular, yet what people seem to mean when bringing these two notions together seldom amounts to the above claim that emotions constitute an intelligent access to the world. The notion of "embodiment" at issue here clearly means more than just some kind of interdependency between brain and body. Cognitive processes not only depend on bodily processing in the trivial sense that the brain needs to be nourished by the rest of the body; such wisdom is already captured in the Latin saying "mens sana in corpore sano" and certainly would not do as the foundation of a new paradigm in cognitive science. Furthermore, it is not sufficient to claim that bodily reactions somehow accompany emotions. The talk about "embodied emotions," however, is often restricted, in a misleading way, to the claim that emotions are accompanied, or followed, by or expressed through bodily reactions. When talking about embodied emotions people often mean not much more than that being sad *causes* a certain kind of bodily posture or facial expression or inner arousal.[9] Paula Niedenthal and colleagues, for example, write, "By embodiment we mean the bodily states that arise (e.g., postures, facial expressions, and uses of the voice [such as prosody]) during the perception of an emotional stimulus

and the latter use of emotional information (in the absence of the emotional stimulus)" (Niedenthal et al. 2005, 23). Niedenthal's work is highly interesting with regard to the question whether emotional concepts have an empirical basis in emotional feelings with regard to the question of conceptualization. Yet for the claim that bodily processing is involved in the constitution of cognitive processing, or enables a certain kind of intelligent interaction with the world, it does not suffice to say that certain postures or expressions occur "during the perception of an emotional stimulus," that is, that they *occur* together with a certain stimulus. The claim needs to be that the bodily posture, facial expression, and internal arousal accompanying sadness are not only an output or a coincidental by-product of the emotion but functional elements that constitute intelligent access to the world.[10]

The view I defend claims that different emotion types include different patterns of bodily reactions that evolved because they prepare the organism for certain kinds of action (Chapter 2). These bodily reactions are directly involved in the constitution of action-oriented representations. These representations establish a type of nonconceptual access to the world. They enable us to interact with our environment in an intelligent way. Through them we perceive a situation as a danger that should be avoided or as a rule violation that we should compensate for (Chapter 5). Bodily reactions thereby play a constitutive, and not merely a causal, role in emotional processing. More importantly, they play a role in a theory that aims to replace standard cognitivist approaches that think of emotional processing as essentially involving symbolic representations and conceptual knowledge.

b. Action-Oriented Emotions

Emotions are embodied action-oriented representations, that is, representations that not only represent something as a descriptive fact, but have a directive component at the same time. This is the kind of representation format that Millikan (1993) calls "pushmi-pullyu"[11] and Andy Clark (1997), action-oriented. I also take it that the fact that emotions are motivating and have a directive content is due to their being embodied. The bodily reactions involved in emotions are not randomly occurring arousal. They are well-adapted and skillful reactions that prepare the organism for action and *thereby* also realize an embodied action-oriented representation.

While my work fundamentally rests on the assumption that sensorimotor abilities play a constitutive role for the way we intelligently access the world, it departs from current enactivist approaches that make a similar claim in several respects. Enactivism nowadays comes in several varieties and there are several controversial issues even among enactivists. The version of enactivism brought forward by Kevin O'Regan and Alva Noë (2001) can be labeled as *sensorimotor enactivism* (SE). Sensorimotor enactivists rely heavily on the notion of skill, which can be used interchangeably with know-how. A skill is a kind of practical nonconceptual knowledge. Skills

are abilities but not all abilities are skills. While breathing and digesting are abilities that need not be learned, playing the violin and cooking chili are skillful abilities that presuppose a training process.

Those who embrace *radical enactivism* (RE) object that calling perception a skillful ability is an overintellectualization of perception. People do not actively train to perceive in the way they train to play the violin. It might therefore seem more reasonable to describe perception as an enactive ability with a biological purpose (Hutto 2006, Hutto and Myin 2013). I think that this critique of the notion of skill partly rests on an overintellectualization of the notion of skill itself and partly overlooks the substance that the focus on trained skills adds to the enactive view. As I argue in detail in Chapter 5, training does not need to imply any kind of conceptual or representational guidance but can be understood as a process of adaptation. In this sense I suggest that perception and emotion (the faculties, not single mental states or processes), can be seen as evolutionarily set up representational abilities that become more and more skillful during development.

Another enactivist tradition can be labeled as *autopoietic enactivism* (AE). AE broadens the notion of cognition to include all kinds of coordinated interactions that organisms perform in their environment. Such structured interactions are called *sense-making*, which, according to AE, is the mark of the mental and an activity that all living organisms exhibit as autonomous and adaptive systems. Autonomous systems are those that are inherently purposeful, in that they generate ends or purposes within themselves in order to maintain themselves (Varela, Thompson, and Rosch 1991, Thompson 2007, Colombetti 2014, Di Paolo and Thompson 2014).

While the present approach subscribes to the general enactivist claim that emotions are not just passive feelings but skillful and action-oriented, I disagree with the antirepresentationalist attitude found in most accounts of SE, RE, and AE alike. I also depart from AE when it comes to the relation of the active sense-making organism and the structured world it is surrounded by. Action-oriented approaches need to be supplemented with the claim that organisms are embedded into a structured environment, and to develop this claim and its ontological implications further than enactivist authors so far have done. I argue in Chapter 3 that enactivist accounts have a tendency toward what I call "bioconstructivism," which does not fit well with the externalist assumptions that embodied accounts appear to be committed to. Third, I develop a notion of extended functionalism that relies on Clark's approach as well as on Millikan's notion of proper functions. Functionalism is not fully compatible with enactivism when it comes to the question of what it means to say that a system is embodied (see, e.g., Clark 2008a).

c. Embedded Emotions

A lot of ink has been spilled trying to answer the question whether environmental elements can play a truly constitutive role with regard to emotions

or just causally interact with them (see, e.g., Krueger and Colombetti [forth-coming] for an overview). The discussion is largely focused on single exter-nal elements and the triggering or regulating role they play in emotional processing. What tends to be forgotten in this discussion is the question of what kind of intelligent access to the world is constituted through interac-tion with the environment. In this work I largely leave questions of constitu-tion aside and focus on what Shapiro calls "replacement" (of complex inner representations with embodied skillful interaction with the environment) instead. What is needed to replace the assumption that emotions involve cognitive appraisals, I argue, is (a) the commitment to embodiment, that is, the claim that bodily reactions can be adapted to the environment and realize intelligent behaviors; (b) the action-oriented claim that the ability to be emotional is a skillful ability; and (c) the commitment to the embedding thesis. The claim that emotions are embedded implies the need to develop an account of the ontological structure of the biological and social world into which emotions are embedded.

Once such an ontological structure is in place I argue that even apparently cognitively complex emotions, such as guilt and pride, can be acquired early and do not presuppose complex cognitive representations. A developed social ontology allows for the claim that emotions' intentional objects are normative yet external and can be represented by embodied action-oriented representations.

The main claim I establish against traditional cognitivist approaches to emotion is that emotions do not necessarily involve judgments, symbolic representations, or anything with conceptual content; rather, they can be described as embodied, action-oriented representations that can be ascribed to infants and animals as well as adults. But it is often neglected in current debates on embodied and embedded emotions that cognitivist approaches can account for a lot of features that emotions have which embodied accounts have problems explaining. Emotions appear to have intentional objects, they stand in certain rational relations to each other and to other mental states, and they are concerned with social rules and norms. Cognitiv-ists can explain all these features with reference to complex inner appraisals.

Most accounts of embodied and/or embedded emotions are implicitly committed to certain forms of externalism that I call *diachronic environ-mental externalism* (DEE) and *synchronic environmental externalism* (SEE). DEE is the claim that the environment takes on an active structuring role in the evolution of emotions. Biologically inspired versions of naturalism, and particularly the teleosemantic approach to which Jesse Prinz is com-mitted, claim that organisms develop representational powers or the ability to enact meaning in direct interaction with the world and in response to certain adaptive pressures exerted by their environments. For example, vari-ous organisms have developed eyes and the ability to use visual information *because* the ability to use visual data is of immense value for guiding the

organism's behavior. This view depends on the claim that there are situations in the external world that our ancestors repeatedly faced, thus making it beneficial for these organisms to develop mechanisms in response. Maybe things in the world are not intrinsically colored, but the general assumption is that things in the world must have color-corresponding features of some sort that explain the discriminative function of our color-detection mechanisms. Similarly, with emotions, if fear, for example, is set up to represent dangerous situations, then there must be things in the world that were dangerous for the organism in question when the reaction was set up, or at least there must have been situations of a certain type in which it was functional to develop a fear-mechanism to represent them.

Embodied and/or enactive approaches also rely on SEE. Brooks's famous quote that "the world is its own best model" (Brooks 1991) captures nicely what SEE is all about. What Brooks means, in a narrow sense, is that once we focus on direct forms of intelligent interactions with the outside world, we can refute the concept of internal representations as being something in between the acting organism and the world altogether. While among authors in the embodiment field there is a large controversy concerning whether representations should be eliminated altogether or if the term should simply be modified (see, e.g., Clark 1997, 2014, Gallagher 2008b, Chemero 2009, De Bruin and Kästner 2012, and Chapter 3), there is broad consensus about SEE—that is, the claim that no complex, world-representing inner machinery is needed to successfully interact with the world. The world is out there, it is structured, and organisms are well adapted to it since they have many ways of directly accessing information to guide action without first forming complex inner models. This is a shift in the explanatory burden, from the head to the world. The focus on the structure of the environment, and the organism's bodily skills that interact with it, allow us to explain and simulate intelligent behavior without necessarily assuming the involvement of complex cognitive processing. The claim that "the world is simply out there" is an ontological claim that plays a crucial role in the explanation of how emotions can be embodied but meaningful:

> Radical embodied cognitive science requires a new ontology, one that is at odds with today's physicalist, reductionist consensus that says the world just is the physical world, full stop. Without a coherent understanding of what the world is like, such that it can contain meanings and is not merely physical, direct perception is simply indefensible.
>
> (Chemero 2009, 136)

The claim that we can directly perceive complex features of the world only makes sense if we describe the world as a structured environment to which an organism is adapted. Part of the claim that needs to be developed is the notion of a biologically inspired naturalism. The ontological

enrichment that biologically inspired approaches bring about allows us to see the environment as structured in a way that is directly perceived by the organism. With regard to emotions, however, the new ontology stretches from the biological into the social domain. Emotions not only represent situations that are of importance for the survival of the organism in the wild—such as dangerous predators and indigestible food—they also mainly represent urgent features of social scenarios, such as the loss of attention from a closely related person or a rule violation committed by oneself or others. DEE and SEE, with regard to emotions, therefore demand a certain kind of normative realism that explains the reality of the social context as analogous to a structured ecological niche; the social world developed historically, exists objectively, and the organism inherits the social environment in the sense that it is well adapted to it. Infants and animals can represent rule violations and the loss of attention because these are features that are of central relevance for the organism and have a constant causal impact on it.

3. OUTLINE

In Chapter 1, traditional cognitivist approaches to emotions are discussed. These accounts conceive of emotions as complex cognitive, evaluative states, usually as some kind of judgment. Relying on the insights of cognitivist approaches, I develop what I call "the normative structure of emotions." Emotions can be described as (1) being subject to *semantic norms*—they are usually described as having a representational content, (2) being subject to *rational norms*—they can be described as standing in rational relation to other emotions and other mental states, and (3) being subject to *social norms*—while all emotions have an evaluative character, some emotions, such as guilt and shame, are concerned with social rules and norms. I argue that while traditional cognitivist approaches have a natural explanation for the normative structure of emotions, noncognitivist and embodied approaches fail to explain the normative structure of emotions. The *normative challenge* is to account for this normative structure in the context of a noncognitivist approach.

In Chapter 2, I introduce the famous Jamesian claim that emotions essentially entail patterns of bodily arousal and the objections that cognitivists have raised against that claim. I demonstrate how current approaches have refined the original thesis to meet cognitivist objections and review recent evidence from psychophysiology that suggests that there is good reason to defend an updated version of the Jamesian claim. I furthermore suggest that the Jamesian claim can be defended for all emotions and that there is no neat distinction between so-called basic and higher cognitive emotions. I argue that the Jamesian claim has to be understood in a Darwinian context: emotions entail homeostatic reactions that evolved incrementally and

prepare the organism for action. The conclusion of the chapter is that current embodied approaches explain the Jamesian claim well, specifically in a Darwinian framework.

While Chapter 2 describes the bodily reactions involved in emotions and gives a vivid impression of their central role in emotions, it doesn't address how embodied approaches could meet the normative challenge introduced in Chapter 1. Chapter 3 discusses noncognitivist solutions, in general, and embodied and enactive accounts, in particular. The aim is to show that these theories develop various interesting solutions to account for the intentionality of emotions on a noncognitivist level but that none of these approaches can account for the normative structure of emotions.

Chapters 4 and 5 subsequently develop an account of embedded, action-oriented emotions on the basis of what has been said about the embodiment of emotions and the normative challenge in prior chapters. The general aim of the book is to provide a theory of emotions as embodied, action-oriented representations that is externalist and naturalist in spirit. The special difficulty for such an approach is the normative challenge. How should one deal with the fact that emotions are subject to semantic and rational norms and that they are about things that are of value for us? Cognitivist theories propose that we assimilate emotions to judgments. In contrast, I suggest that an embodied account of emotions requires instead (1) a theory that allows homeostatic reactions to carry information that has the function of guiding an organism's behavior and (2) an ontology that makes sense of the claim that "the world is its own best model" in the case of emotions, that is, an ontology that takes organisms to be embedded in a structured environment in which certain things are of value for us and should be approached, while others should be avoided. Chapter 4 is, then, mainly dedicated to developing those ontological claims and to explaining how the description of a structured environment can come to replace assumptions about complex inner cognitive appraisals.

In Chapter 5, I propose that emotions are action-oriented representations that refer to certain kinds of affordances that we encounter in our environment. These action-oriented representations constitute a practical knowledge of the social world we are embedded in. Emotions are constituted by embodied homeostatic reactions, which become skillful in a process of interaction with the social world from early on. Infants learn to represent core relational themes of high complexity through simple bodily reaction patterns. I call emotions "clumsy skills" because the goals they motivate us to act on are often at odds with the rational goals we have. Emotions are still governed by norms, but the norms emotions are guided by can differ from those we are rationally convinced we should follow. The normative challenge can be met if we replace the judgmental structure that cognitivists assume, with a description of the interplay between a complex social world and an embodied organism that is situated within this world and ready to interact with it.

NOTES

1 Davidson refers to Hume (1739), arguing that Hume's theory of causally associated ideas cannot make sense of the conceptual relations at work in pride and that Hume would have done better in describing pride as a judgment.

2 See e.g. Hart and Carrington 2004, Fessler 2004, Tracy and Robbins 2004, 2007a, 2007b, Draghi-Lorenz, Reddy and Morris 2005, Reddy 2008, Clark 2010. There is even a tendency in research on primates to ascribe moral emotions to them (see, e.g., Rowlands 2012).

3 With regard to the mind-body problem, it might be the ultimate goal of strong reductionist theories to use physics as a base theory but only via intermediate theories and in the long term. To refer to physics for explaining mental phenomena in present debates is, for obvious reasons, not an attractive possibility. The proposed reduction is therefore one from other disciplines in the cognitive sciences to neuroscience.

4 To argue for this point Millikan (1993) gives the example of the sperm's function of fertilizing an egg. The function of something and what it does on average couldn't be more different.

5 I am aware of the fact that it is not "Gibsonian" at all to talk about affordances as things that can come to be represented. I will account for this in the further development of the view of emotions as embodied action-oriented representations (mainly in Chapter 3 and Chapter 5).

6 I do not argue in detail for the claim that emotions are more than simply social constructions without any roots in evolution. The first chapter offers a huge amount of empirical data that could not be reasonably explained by a radical social constructivist account. For arguments against the claim that emotions are just learned social roles or socially generated scripts, see, for example, Griffiths (1997) or Prinz (2004).

7 Thinking of emotions as "sandwiched" between simple reflexes and homeostatic reactions, on the one hand, and higher cognitive processing, such as deliberate conceptual reasoning, on the other, is fairly common in the various disciplines of the cognitive sciences (see, e.g., Damasio 1999, 53f.; Zajonc 1980).

8 For a recent rejection of both positions, see Welpinghus (2015).

9 This objection has also been raised by Stephan, Walter, and Wilutzky (2013).

10 For a more detailed discussion of how the notion of embodied emotions is usually used in current debates and why most of these approaches do not fit the central claim of the embodied cognition paradigm see Stephan, Walter, and Wilutzky (2013).

11 See also Scarantino (2014) for a current approach applying pushmi-pullyu representations to emotions.

1 Cognitivism and the Normative Dimension of Emotions

1. INTRODUCTION

Philosophers and psychologists from the 1960s on started to criticize scientific approaches to emotions. The general aim of this critique has been to fundamentally change the scientific view of emotions as pure states of physiological arousal or observable behavior that can be fully explained with regard to the causal processing they involve, or as mere bodily feelings without any kind of meaning. Apart from the rejection of scientific reductionism, the claim that emotions involve cognitively complex evaluations has often been defended in order to give a more adequate description of how emotions can play a role in our social interactions, the gaining of knowledge, or moral reasoning. The writings of the philosopher Robert Solomon are most telling in this regard:

> Against those romantics and contemporary bourgeois therapists who would argue that emotions simply *are* and must be accepted without judgment, it appears that emotions themselves are already judgments. And against several generations of moral philosophers who would distinguish between morality based upon a principle and morality based upon emotion or "sentiment," it appears that every "sentiment," every emotion is already a matter of judgment, often moral judgment.
>
> <div align="right">(Solomon 1973/2003, 8)</div>

With such claims that every emotion is a matter of (moral) judgment, cognitivists highlight that emotions have an evaluative or normative dimension. This normative dimension, they submit, is the main reason why emotions cannot be understood as mere bodily feelings or behavioral reactions. I take it to be a main problem of current noncognitivist and, particularly, embodied approaches to emotions that they cannot account for this normative dimension. Yet a theory that cannot account for the normative dimension of emotions falls back beyond the insights that cognitivists already pointed out, beginning in the 1960s. It is therefore the aim of this chapter to reconstruct the normative structure of emotions, and it will be the aim of the chapters that follow to account for it in an embodied framework.

In the broadest sense, the phenomenon to be explained is that emotions are *normatively assessable*; they can be appropriate or inappropriate. We talk about "justified anger," "unwarranted jealousy," or "disproportionate fear." What is controversial is the theoretical explanation of this phenomenon. To what kind of norms are we referring, when we say these things? Anthony Kenny offers the following suggestion:

> In fact, each of the emotions is appropriate—logically, and not just morally appropriate—only to certain restricted objects. One cannot be afraid of just anything, nor happy about anything whatsoever . . . if a man says that he feels remorse for the actions of someone quite unconnected to him, or is envious of his own vices, we are at a loss to understand him.
>
> (Kenny 1963, 192)

According to Kenny, emotions appear to have a dimension of moral appropriateness and (maybe more important or fundamental) of logical or rather rational appropriateness. What Kenny has in mind is nicely illustrated in the following quote from *The Passions of the Soul* where Descartes gives a definition of envy:

> Envy . . . is a kind of sadness mingled with hatred, which results from our seeing good coming to those we think unworthy of it. Such a thought can be justified only in the case of goods of fortune . . . But sometimes fortune gives advantages to someone who is really unworthy of them. Then envy stirs us up only because having a natural love of justice, we are vexed that it is not upheld in the distribution of these goods.
>
> (Descartes 1649/1988, §182–183)

This definition appears to cover both the rational and moral appropriateness of emotions: it is rationally appropriate to call an emotion "envy" if the envious person entertains a judgment about somebody else getting a good she is not worthy of. While in principle any person getting any kind of good can become an object of envy, I cannot possibly envy myself, and it is also not possible to envy somebody for something that I do not consider to be a good. It seems that there are certain rational restrictions on what can become an object of envy. Furthermore, if the person I envy is "really unworthy" of the goods she received, then my envy is also morally appropriate, which implies that there are moral standards that can be applied to emotions to tell the warranted from the unwarranted ones.

The normative dimension of emotions is a complex phenomenon that involves claims about logical and moral norms alike. In what follows, I suggest that an accurate analysis of what is at issue in the debate about the normative dimension of emotions involves three different claims regarding (1) semantic norms, (2) rational norms, and (3) social norms. All these claims

do their share in explaining the phenomenon of normative assessability. One need not accept all these claims; in fact there are current approaches to emotions that do not accept any of these claims. But accounts that deny that emotions have semantic, rational, or social norms need to come up with an alternative explanation of the normative assessability of emotions to avoid inadequate reductionism. There are of course good reasons not to be a cognitivist. I discuss the most common objections against cognitivist theories later. Yet I take the attention to the normative dimension of emotions to be the greatest strength of cognitivism, and consequently the biggest challenge to current noncognitivist theories is to come up with an account of the normative structure of emotions in noncognitivist terms.

2. SEMANTIC NORMS

Probably the most prominent objection against William James and the "feeling theory" of emotions is that the approach to emotions as feelings of bodily arousal denies the *directedness* of emotions toward the world, or their *intentionality*. William James (1884) developed the view that emotions are feelings of bodily arousal. Being afraid means to feel that one's heart is beating faster, adrenalin is released, and the whole body is trembling. The insights and blind spots of James's approach are discussed further in the next chapter. Suffice it to say, James and the "feeling theory" have been the favorite opponent of the cognitivist tradition, since it seems that James takes emotions to be bodily feelings without any cognitive dimension.[1]

As Solomon puts it, "[f]eelings do not have 'directions.' But I am angry 'about something.' The relationship between my being angry and what I am angry about is not the contingent relation between a feeling and an object" (Solomon 1973/2003, 4). What is generally true for intentional states of the mind is also true for emotions: the intentional object need not be an external object standing right before our eyes causally triggering the representation. Emotions can be caused by memories and imaginations or by thoughts about nonexistent entities such as a possible war, and they can have physiological causes, like too much coffee, while still being directed at an external situation. It is therefore important to distinguish between an emotion's cause and its object: while anger is always directed at an object or a state of affairs, it is not always the object that caused one's being angry. Anger can be caused by a complex chain of annoying events and simply erupt in a harmless situation, or it can be caused by insufficient sleep and too much coffee but still be directed at somebody's remark interpreted as a terrible insult.

The conclusion drawn from this is that emotions not only have intentional objects but that subjects, through emotions, grasp the objects in a particular way. The intentionality of emotions, their being directed at something, cannot be explained by the mere reference to the objects in the extension of that

emotion. Instead, emotional content needs to be cashed out with regard to the *intensionality* of emotions, the particular mode of presentation in which the object is given in experience. The shift from viewing emotions as mere physiological responses or bodily feelings to a cognitively adequate description of their content stems from the insight that emotions are intensional in the sense that they present their object in a certain mode of presentation. Martha Nussbaum in this vein points out that emotions are not "*about* their objects merely in the sense of being pointed at them and let go, the way an arrow is released towards its target . . . Their aboutness is more internal, and embodies a way of seeing" (Nussbaum 2001, 27). I will come back to the explanation of this "way of seeing" later.

In claiming that emotions have intentional objects that cannot be identified with the cause of an emotion and are given to us in a certain mode of presentation, cognitivists treat emotions as involving (or sometimes as being) a certain type of mental representation. Essential to all kinds of mental representations is that they can go wrong. We can have false beliefs and we can draw wrong inferences or entertain confused concepts. In order to tell perceptions from misperceptions and true beliefs from false, such representations must have correctness conditions. There must be an external standard of reference to which the distinction between true and false representations can be drawn.

Different theories differ widely in how they characterize the representations involved in emotions. Many cognitivists take emotions to include judgments, while others assume that emotions are constituted by beliefs or belief–desire pairs. In critiquing classical cognitivism many authors have compared emotions to perceptions, claiming that emotions have a certain kind of nonconceptual content. While all these accounts characterize emotions as representational, they make different assumptions about the correctness conditions of emotions. Beliefs and judgments entail propositional content and aim at truth, while desires aim to be satisfied. Thus, beliefs and judgments are said to have truth conditions while desires are said to have satisfaction conditions. People who ascribe nonconceptual content to perceptual representations usually claim that these representations are not about states of affairs but rather present objects to us. Therefore, perceptions cannot be said to have truth conditions, but rather adequacy conditions. Since we do not perceive that something is the case, perceptions cannot be wrong in the way beliefs about states of affairs can be. But we can perceive objects in ways that are adequate or inadequate in relation to how they really are.

After leaving the differences among representationalist theories aside, it should be clear that every theory that takes emotions to involve a kind of mental representation must assume that there are correctness conditions that apply to emotions. It must thereby assume as well the existence of semantic norms that apply to emotions. One way to understand the normative assessability of emotions is simply to say that emotions are mental

representations, and as such they can be adequate or inadequate in a certain situation: one might have good reasons to be angry or overreact in a harmless situation because of other stressful factors. It is certainly one option to explain our talk about "justified anger" or "appropriate jealousy" with reference to semantic norms. The claim that emotions are subject to such semantic norms explains how they can be meaningful and appropriate or inappropriate in a way that pure physical reactions are not. Of course, with the commitment to semantic norms arise several other questions about the nature of these norms, their constitution, and so on that each theory has to answer.

The particular "aboutness" of emotions, and thereby the kinds of norms we have to assume, can be further specified. In contrast to perceptions, emotions depend on former representations: first, we see an object, and then we get scared. First, we remember that we forgot a friend's birthday, and then we feel guilty. It has often been remarked that emotions highlight things that are of particular importance to the subject. Furthermore, each emotion type highlights a situation in a particular way. A common suggestion to capture the "particular way" in which emotions make their objects appear is to introduce for each emotion type a formal object. The claim is that for all emotions E, E represents the formal object of E (as instantiated by some material object O; Scarantino 2010). To assign a formal object to an action means to place restrictions on what may occur as the direct object of a verb describing the action. "Only colored things can be seen" is an example where "colored things" are the formal object of the act of seeing. "Only property can be stolen" is another example where "being somebody's property" is the formal object of stealing. In the case of emotions the claim is that only dangerous events or objects can be (appropriately) feared, so being dangerous is the formal object of fear. Every type of emotion has such a formal object: being an offense is the formal object of anger, being indigestible is the formal object of disgust and so on. The notion of the formal object offers a more precise definition of the "particular way" that each emotion represents its object. For example, it means that each fear reaction represents a particular event or thing as being dangerous, while each anger reaction represents something as being offensive. The correctness conditions of emotions therefore concern the formal object and the question of whether it is really present—whether, to continue the example, it is adequate or inadequate to represent a situation as dangerous or offensive.

Given that the formal object is an appropriate notion to capture the intensionality of emotions, what we need is an account that specifies the way each emotion represents its particular object. Such an account is offered by the psychologist Richard Lazarus. According to Lazarus, emotions are always concerned with relations between the individual and the environment. These relations do not merely depend on the way in which a particular object or situation strikes an individual; rather, the individual herself has certain beliefs, values, and dispositions that enable her to appraise certain situations in a

particular way. In accordance with Nussbaum, Kenny, and Solomon, Lazarus holds that emotions never simply represent objects or bare facts. An account that describes merely causally generated input–output patterns, therefore, does not offer a satisfying analysis of emotional content. The relational meaning emotions contain is not a simple set of properties, but rather, the result of an abstract evaluation concerning the relevance of the features of the environment to the person. According to Lazarus, emotions are judgments with relational meaning and they are established through a prior appraisal process:

> If we feel threatened, insulted, or benefited—these are, of course, appraisals—the relational meaning of each does not stem from either the person or the environment; there must be a conjunction of an environment with certain attributes and a person with certain attributes, which together produce the relational meaning.
>
> (Lazarus 1991, 90)

Representing something as a threat cannot be explained as a mere causal stimulus–response–reaction, since being a threat in the first place depends not only on certain attributes of the external object but also on attributes of the experiencing subject. What Lazarus apparently has in mind is that "being dangerous" or "being a benefit" are relational properties: An object O is only dangerous in relation to a subject S because certain features of O can cause harm to S. Lazarus furthermore assumes that the relational contents of emotions can only be grasped by a subject with a certain set of background beliefs, where the subject appraises the situation in a certain way in accordance with her background beliefs. Emotions represent what Lazarus (1991) calls "core relational themes." Core relational themes are the recurring types of relations between the individual and the environment that fundamentally concern one's well-being, such as "being in danger," which is the relation between an individual and the environment we represent in fear, or loss, which is the relation we represent in sadness.

Several authors, particularly in the field of embodied cognition, have adopted the idea that emotions are about core relational themes (e.g., Prinz 2004, Hutto 2012, Colombetti 2014). All of these authors agree that emotions do have intentional objects, yet many would deny that to entertain the emotions in question one needs to have the cognitive abilities that Lazarus's account demands. Lazarus's assumption that emotions entail various systematic appraisals in conjunction with a substantial amount of background knowledge makes it hard to explain why infants and animals can have emotions. Also, if emotions were judgments they would frequently contradict other judgments we entertain. Yet with the assumption of complex appraisal processes, cognitivists can explain the fact that emotions are not only subject to semantic norms but also to rational norms. The nature of these rational norms and the cognitivists' standard strategy to account for them shall be at issue in the next section.

3. RATIONAL NORMS

We have seen that the commitment to semantic norms allows for a partial explanation of the normative assessability of emotions. When we say, "It is really inappropriate to be afraid of that little spider, it is not dangerous at all," we are referring to semantic norms. Fear is inadequate because the formal object of fear is not instantiated by the particular spider in question. Many cognitivists further argue that emotional content can only be understood with regard to the rational connections among the emotion in question, other mental states, and other emotions. Such approaches hold that to fully explain the normative assessability of emotions we need more than just to point at the representational character of emotions. We also have to assume that these representations are the result of reasoning processes that stand in certain rational relations to the situation, other thoughts, and other emotions.

According to Anthony Kenny, who originally introduced the notion of the formal object into the debate over emotions, the formal objects of emotions differ from the objects of nonintentional acts because they are restricted not only by causal laws, but also by logical constraints: "If the emotions were internal impressions or behaviour patterns there would be no logical restrictions on the type of object which each emotion could have" (Kenny 1963, 191). But obviously, emotions are logically or rationally restricted: the formal object of an emotion restricts its intentional object by specifying the conditions under which having this emotion about this object is appropriate or reasonable, but not the conditions under which it is causally necessary. There are causal laws explaining why we can only visually discern things that are not smaller than 0.1 millimeters, but there is no such law explaining why people like us can only envy goods of others and not one's own goods or another's evils. When we say that we envy our neighbors for having such a pretty garden, this might be adequate or inadequate, but it does make sense. When we say the same thing about ourselves it doesn't make sense, because to envy somebody *means* to desire *someone else's* property. One's own property cannot reasonably be desired, since it already belongs to oneself.

The example of envy clarifies that an emotion's appropriateness depends on logical or rational norms rather than causal laws. Furthermore, there are rational connections between different emotions, and between emotions and other mental states. Think, for example, about the connections between backward- and forward-looking emotions. Bennett Helm points out that it seems "rationally unwarranted, other things being equal, about feeling fear that one's prize Ming vase is about to be destroyed, but feeling neither relief when it miraculously escapes unscathed nor sadness or anger when one's fear is borne out" (Helm 2001, 68). Following Helm's approach, emotions are connected in such a way that it makes sense that in certain situations we should become sad if we had been afraid before, whereas in other situations it would make sense to be relieved instead.

Cognitivists typically assume that to account for the rational norms emotions are subject to, we need to think of emotions as being caused by, or being constituted by, complex appraisal structures. These appraisal structures also explain how emotions can be type-identified with regard to their content. Richard Lazarus's account of the individuation of core relational themes is a prime example of such an account. It will thus prove instructive to have a closer look at Lazarus's approach involving core relational themes to see how it can account for the emotions' being subject to rational norms.

Lazarus thinks of emotions as judgments that are the result of prior cognitive processing, including background beliefs about how things are, and appraisals about what it means for oneself that things are thus. The appraisals made in the emotional process determine which emotion is triggered. To describe this process, Lazarus develops an appraisal structure that ranges from general to more specific evaluations. He thereby distinguishes between primary and secondary appraisals. The *primary appraisal* is concerned with whether something of relevance to the individual has happened and includes three components:

1. *Goal relevance*: Goal relevance is a necessary condition for every emotion. Emotions only occur if we evaluate something as being relevant for us and our goals.
2. *Goal congruence/incongruence*: Emotions are either positive or negative. Positive emotions are those elicited if the situation is appraised to be congruent with our goals, and negative are those elicited if the situation is incongruent with our goals.
3. *Type of ego involvement*: We can care about our own or another's well-being in reference to background values, ego ideals, or social esteem. There are, accordingly, six types of ego involvement that can trigger different emotions: one's own well-being with regard to (a) background values, (b) ego ideals, or (c) social esteem, and the other's well-being with regard to (d) background values, (e) ego ideals, or (f) social esteem.

The *secondary appraisal*, which also involves three components, concerns the coping options of the subject in the emotional situation:

4. *Blame/credit*: An important appraisal for several emotions, but not all, concerns the ascription of blame or credit to oneself or others. In guilt we blame ourselves, when angry we usually blame others, and pride concerns credit to one's own deeds.
5. *Coping potential*: Emotions are motivating; they highlight certain options for action. Anger motivates one to react aggressively while disgust motivates one to avoid certain things or persons.
6. *Future expectations*: Further thoughts about the possible effects of an emotional reaction might influence whether a certain kind of behavior, to which one feels motivated, is enacted or inhibited.

These appraisal components are ordered, proceeding from very broad evaluations to more narrowly defined appraisals of the environment. Lazarus uses this appraisal structure to distinguish among different types of emotions and their particular contents. The appraisals start with a general distinction between emotional and nonemotional encounters, then between negative and positive emotions, and finally distinguish the emotion in question from all other emotions. With regard to the appraisal structure, Lazarus's aim is to analyze how the content of an emotion is individuated on a cognitive level. The goal-appraisal hierarchy is the tool for this analysis. For example, the situations in which anger is triggered can be described in the following schema:

1. If there is goal relevance, then any emotion is possible, including anger.
2. If there is goal-incongruence, then only negative emotions are possible, including anger.
3. If the type of ego-involvement engaged is to preserve or enhance the self- or social-esteem aspect of one's ego-identity, then the emotion possibilities include anger, anxiety, and pride.
4. If there is blame (which derives from the knowledge that someone is accountable for the harmful actions, and they could have been controlled) then anger occurs. If the blame is to another, the anger is directed externally; if to oneself, the anger is directed internally.
5. If the coping potential favors an attack as viable, then anger is facilitated.
6. If future expectancy is positive about the environmental response to attack, then anger is facilitated.

<div align="right">(Lazarus 1991, 226)</div>

In anger, blame is directed at others or ourselves, and blaming requires the background-belief that the blamed person could have acted differently. While the belief that a person acted offensively at will is necessary for anger, blame is the appraisal that grows out of it in the context of threat and frustration. Furthermore, it is a particular goal—the preservation or enhancement of ego identity—and not a general goal that must be frustrated in anger. The core relational theme of anger is thus, according to Lazarus, a "demeaning offense against me and mine." Anger reacts not to any kind of frustration of a goal but to those that carry a special significance: a slight or injury to one's identity.

To name another example, the core relational theme of sadness is irrevocable loss. If there is a loss to any type of ego involvement—esteem, moral value, ego ideal, meanings and ideas, persons and their well-being, or life goals—sadness is possible. If there is no blame then sadness is likely, but if there is coping potential (i.e., something that can be done to undo the harm or restore the loss) the emotion will not be sadness but one that involves a struggle to change a goal-incongruent condition. Only if nothing can be done, sadness is the indicated emotion, which is implied in the core relational theme for sadness, irrevocable loss.[2] Among the positive emotions,

happiness results from reasonable progress toward the realization of our goals. The type of ego involvement is less relevant to happiness than to pride. If there is ego involvement in happiness it is probably in the background sense of fulfillment, security, and well-being. In this way, Lazarus uses the tree structure of the different appraisals to define the core relational themes for the twelve emotions described in Table 1.1:

All the emotions listed concern relevant goals and are therefore triggered in situations where something of relevance with respect to our goals occurs. The emotions furthermore divide into those that occur when happenings in the environment are congruent or incongruent with our goals. The negative emotions triggered by goal incongruences are anger, fright, anxiety, guilt, shame, sadness, envy, jealousy, and disgust. While the frustration of a goal may be followed by any negative emotion, which particular emotion is elicited depends on further appraisals. If there is the possibility of a future harm, the likely response is anxiety; when goal frustration can be blamed on another person, the likely response is anger; when it is blamed on oneself, the likely response is guilt; and when there is no one to blame, sadness.

We started with the cognitivist assumption that emotions are subject to rational norms because they stand in rational relations to each other and

Table 1.1   Lazarus's Table of Core Relational Themes

Emotion	Core Relational Theme
Anger	A demeaning offense against me and mine
Anxiety	Facing an uncertain, existential threat
Fright	Facing an immediate, concrete and overwhelming physical danger
Guilt	Having transgressed a moral imperative
Shame	Having failed to live up to an ego ideal
Sadness	Having experienced an irrevocable loss
Envy	Wanting a valued object or achievement that someone else has
Jealousy	Resenting a third party for loss or threat to another's affection
Disgust	Taking in or being too close to an indigestible object or idea (metaphorically speaking)
Happiness	Making reasonable progress toward the realization of a goal
Pride	Enhancement of one's ego identity by taking credit for a valued object or achievement, either one's own or that of some group with whom we identify
Relief	A distressing goal-incongruent condition that has changed for the better or gone away
Hope	Fearing the worst but yearning for better
Love	Desiring or participating in affection, usually but not necessarily reciprocated
Compassion	Being moved by another's suffering and wanting to help

to other mental states. By spelling out the appraisal structure that triggers emotions, Lazarus gives an explanation of the rational relations in question and explains how emotion types can be individuated. The appraisal structure allows for distinguishing different emotions with regard to their content and explains the occurrence of emotional representations by defining which appraisals have to be made before a core relational theme can be represented. At the same time it also explains how emotions can be subject to rational norms. When somebody seriously claims to envy herself for her brand-new car, we are at a loss to understand her. It is a conceptual truth that we can only envy the goods that belong to others, and it can be read off the appraisal structure that the rationally appropriate emotions in this case could either be joy or pride because these are emotions that relate valued objects or achievements to the self and not to others.

At the same time, Lazarus's account strongly overintellectualizes the phenomenon. If emotions are judgments triggered by other judgments, how could we think of infants or animals as having emotions? How could we explain emotions that contradict other judgments? I will come back to these objections later. For now I will turn to a further normative dimension that emotions display.

4. SOCIAL NORMS

After having introduced semantic and rational norms, we have a good idea of why it might be inappropriate to envy somebody for the job she has (e.g., in the case where her working conditions are actually horrible) or why it makes no sense to envy oneself (because it violates the rational constraint that envy only makes sense as an other-directed emotion). But Descartes's discussion of envy suggests that envy can also be socially or morally appropriate or inappropriate. Emotions such as guilt and shame are often labeled as "moral emotions," and it has been frequently suggested that emotions in general are "about values" (see Deonna and Teroni 2012). Yet not all values are moral values and not all emotions are labeled as "moral emotions." Therefore, in trying to capture the normative dimension of emotions, one might start to wonder what kind of values emotions are about.

To further clarify whether and, if so, to what degree emotions involve a social or moral dimension, one might take a look at the objects of the different emotion types. If we agree that emotions are about core relational themes and want to find out the degree to which emotions are about values, moral values in particular, we can go through Lazarus's list of core relational themes to see what kinds of values the list involves. If we think of fear as a response to "an immediate, concrete, and overwhelming physical danger," it might make sense to think of danger as something that is of *disvalue* for the organism and therefore *should* be avoided. But this introduces a normative dimension that has nothing to do with morality but only with bodily

well-being of the subject in question that can be understood as a biological value. If somebody is afraid that a new law might be the cause of injustice, this might be a token of a moral emotion, but the emotion type fear has no moral dimension in itself. But there are other emotions for which the claim that they are about biological values or bodily well-being has no plausibility at all. Jealousy, guilt, and shame are essentially concerned with social rules and norms (jealousy according to Lazarus is about "resenting a third party for loss or threat to another's affection," while guilt is about "having transgressed a moral imperative"). This is not to say that the emotions in question have no roots in evolution or that they have not been fitness-increasing. On the contrary, the example of shame and pride as rank-hierarchy-related emotions in apes (see Chapter 2) suggests that these emotions might have developed because they happened to be functional in social groups. The point is rather that certain emotions such as guilt and pride differ from others such as fear and disgust in the sense that for the latter the social context in which they occur is a *sine qua non*. What would guilt be outside of a social situation? What could its intentional content and its function be?[3] Contrary to emotions such as fear and disgust, which can be described as being about basic needs or biological norms, these emotions have social rule violations as their objects. Another way to make this point is to say that we can imagine a species of animals without any social abilities having emotions such as fear and disgust (since these emotions do not presuppose living in a social context or being able to recognize members of one's own species and interact with them in particular ways). But it is conceptually impossible to think of animals able to be proud or ashamed that do not have any social abilities and do not live in a socially structured group, where it is possible to achieve goods or violate rules with regard to, for example, the rank hierarchy. The rank hierarchy is a social structure and the regular behavior that animals living in rank hierarchies show establishes certain norms. This structure has to be in place so that it is possible to violate norms of rank hierarchy at all. It is conceptually possible though to describe social animals as showing pride, guilt, or shame without describing moral norms to them. Contrary to social norms, moral norms are usually defined as unconditional imperatives that are obligatory and context-independent. It is hard to imagine an animal society that establishes norms that are unconditional in the sense that they couldn't change if the conditions that make the norm in question functional change. I argue in Chapter 4 that a naturalist approach to emotions can do with biological and social norms and need not include moral norms.

When we judge whether a certain case of guilt is an appropriate reaction, we can refer to semantic and rational norms by asking whether the represented scenario really contains a rule violation committed by the self. But we can also ask whether the rule violation in question was "really bad," and when we do so, we instead refer to social or moral norms. Some emotions appear to be about moral norms while others are rather concerned with

mere conventions. Embarrassment is an emotion that has a certain kind of rule violation as its object but it is one of etiquette rather than of morality. Guilt is an emotion that we typically describe as being concerned with moral rule violations such as harm-norms. Lazarus calls the core relational theme of guilt "having transgressed a moral imperative." I argue in Chapter 4 that this might be a way to capture the way we often talk about feelings of guilt but that the definition is too narrow if we want to capture cases of guilt in infants. Guilt in infants, I argue, responds to locally established rules by caregivers that often lack the general validity and the context independent obligatory character of moral rules or values. We can take it as a *prima facie* result that the values emotions are about range from self-related values of bodily well-being to conventional and social norms. I will come back to the relation between social and moral norms and values in Chapter 4.

A theory of emotion has to explain how emotions can be about all these different values and norms. A cognitivist theory is apparently in a better position to do so. If emotions are taken to be cognitive evaluations triggered by other cognitive evaluations, we just have to spell out the kinds of evaluations that each emotion involves. A noncognitivist approach is in a less comfortable position, since it is difficult to see what, other than cognitive evaluations, could detect social rules and norms. Emotions such as jealousy, guilt, and pride have often been labeled as "higher cognitive emotions" mainly for this reason (Griffiths 1997). It has been argued that being able to entertain higher cognitive emotions presupposes complex cognitive abilities, such as having an explicit concept of the self and the other and having explicit concepts of social rules and norms (Lewis 2006). The idea behind this hypothesis is that higher cognitive emotions cannot occur before an infant is able to understand that she has made a mistake—that is, violated a social norm—and that she should try to make amends for what she has done in order to repair her relationship to others. It is arguable whether the emotions in question really need to be "higher cognitive" emotions in the sense that infants need to have certain *explicit concepts* before being able to experience these emotions. But explaining how emotions can be appropriate or inappropriate with regard to social rules and norms certainly constitutes a challenge for all noncognitive theories of emotions. Explaining how emotions can be about normative properties is furthermore a challenge for a naturalist account. I argue that emotions are about relational properties that are good or bad for us with regard to biological and social norms. I also argue that while we often talk about emotions such as guilt and shame as moral emotions, a comprehensive view of the norms that these emotions refer to should use the broader notion of social norms.

I have developed the normative structure of emotions by relying on several views from the cognitivist camp that highlight that emotions have correctness conditions, that they are rationally constrained, and that they represent values that concern not only biological well-being but also social norms. To explain these features, cognitivists point out that emotions are

complex representations that stand in relation to other representations. It is the principal merit of cognitivist theories of emotion that they have highlighted the complex normative dimension of emotions. Yet a theory that captures all these features in a comprehensive appraisal theory is also the most radical cognitivist approach one can think of. Such a theory takes emotions to entail representational content, to be triggered by judgments, and to be judgments themselves. There are obvious objections that make such a cognitivist theory untenable. I discuss these objections and discuss the possibility of more moderate versions of cognitivism in the following section.

5. OBJECTIONS AGAINST COGNITIVISM

In recent debates, traditional cognitivism about emotion has become the object of heavy criticism. Several objections have been raised from different directions. The critique is usually not directed against the claim that emotions have a normative dimension with the previously characterized features. Instead, the general aim of the critique is to show that cognitivists overintellectualize the phenomenon in particular if they characterize emotions as judgments. The ability to emote demands explanations that do not treat emotions as if they were linguaform propositional attitudes that can be used in inferential reasoning. In the following section, I briefly introduce the most important objections, discuss how cognitivists can respond to those, and explain what conclusions can be drawn with regard to a possible non-cognitivist approach to emotions.

1. *Phenomenal Aspect:* Emotions apparently involve certain feelings and, according to many, feelings are an essential component of emotions that cannot be captured adequately by cognitivism. Peter Goldie (2000, 50f.) argues for this point, calling it the "Mr. Spock objection." According to Goldie, it would be possible to entertain all the beliefs and desires cognitivism describes as the constitutive parts of emotions without feeling anything—just like *Star Trek*'s Spock. Entertaining certain beliefs and desires is not sufficient for having an emotion. Furthermore, it is not satisfying to simply claim that emotions, apart from beliefs and desires, entail a specific phenomenal component as well, if all this does it merely to add the feeling as one more component without giving it any significant functional role in emotional processing. Yet many versions of cognitivism leave us precisely with this sort of picture, on which emotions involve a certain phenomenal aspect, without clarity on what role that phenomenal aspect plays or how it is related to the beliefs or judgments that constitute the emotion.[4] Therefore, the phenomenal aspect looks more like a superfluous add-on than an essential component. What is characteristic and interesting

about emotions, though, is that the evaluation of a situation and the particular feeling appear to be inseparable. Emotions can be described as pleasurable or painful evaluations, where their special character is exactly the coincidence of the two elements (Helm 2001). Classical cognitivism frequently seems to forget about the phenomenology altogether, and when it is added to the theory, it remains unclear what the role of the phenomenal aspect of emotions is.

Authors such as Nussbaum, Solomon, and Lazarus would certainly deny that feelings play no role in their theories. It is true, however, that although Lazarus occasionally highlights that phenomenology is an important part of emotions, for the core of his theory, as it was presented earlier, the inclusion of feelings makes no difference. This critique alone is no reason to refuse cognitivist approaches, though. Noncognitivists need to come up with a better theory that gives the phenomenology a functional role in the explanation of the normative structure of emotions.

2. *Passivity/Automaticity:* Emotions have been described again and again as passive states that "happen to the organism" and are not under the control of the will.[5] Griffiths further characterizes (several) emotions as affect programs that are automated, meaning that they unfold in a coordinated fashion without the need for conscious direction (Griffiths 1997, 77). These characterizations do not fit well with the claim that emotions are judgments, since judgments appear to be— at least to a much higher degree than emotions—active states that can be governed by the will and do not arise automatically or in a bottom-up fashion.

This argument, however, can easily be turned against its defender, since it is far from clear that "passive" is a status that applies to emotions but not judgments or beliefs. After all, the claim that beliefs and judgments are states that are actively brought about commits us to doxastic voluntarism, that is, to the view that one can choose one's own beliefs and judgments. But doxastic voluntarism overestimates the role of the will in belief formation. Beliefs and judgments are formed as a result of encounters with external facts and how they appear to us, as well as in coherence with the beliefs we already have. Belief and judgment formation can therefore be described as processes that unfold both automatically and in a coordinated fashion. There might be finer criteria for determining the degree to which beliefs and emotions are under the control of the individual, but to simply claim that emotions are passive and happen to us while beliefs are active, in the sense that we choose them, seems to be a rather questionable position. The difference between beliefs and judgments, on one hand, and emotions, on the other, is thereby blurred. Furthermore, cognitivists tend to hold that emotions are judgments precisely because they want to attack the traditional

view of emotions as passive feelings occurring on a cognitively inaccessible physical level: "It is true, of course, that one cannot simply choose to be angry or not to be angry, but can make himself angry or cease being angry only by performing other activities. But this is true of judgments in general: I cannot simply choose to judge a situation fortunate, awkward or dangerous" (Solomon 2003, 11).

Solomon explicitly argues that emotions are judgments, and that we are responsible for our emotions just as we are responsible for our judgments. This does not mean that we actively choose either judgments or emotions. But we can influence them via reasoning and are therefore responsible for them in a very general sense. Therefore, it might be true that emotions occur automatically, yet to turn this into a strong objection to cognitivism it is necessary to explain in more detail what kind of difference this makes compared to more complex cognitive states.

3. *Motivation:* While being triggered automatically and thus being "passive," emotions seem to directly motivate action at the same time. Somebody who is scared feels tempted to flee, somebody who is angry feels prepared for an attack, and somebody who is deeply ashamed feels a strong tendency to hide herself from others (Frijda 1986). This action tendency cannot be explained with reference to emotions being judgments or having a cognitive content alone. Instead, the action tendency inherent in emotions can best be explained with reference to a desire, to the phenomenology, or to the involved bodily arousal. All these components can contribute to the generation of an action. A judgment or belief alone cannot.

There are again different possible strategies to deal with the objection: Solomon simply takes judgments to *be* actions: "Judgments are actions. Like all actions they are aimed at changing the world . . . In other words, emotions are purposive, serve the ends of the subject, and consequently can be explained by reasons or 'in-order-to' explanations" (Solomon 2003, 12). According to this definition, emotions can be regarded as intrinsically motivating. The problem with this claim is that it relies on a completely idiosyncratic understanding of what a judgment is and therefore turns the question into a terminological matter. This is a problem with the works of Solomon, in particular, and many cognitivists, in general.[6]

Another strategy to deal with this objection is to see emotions as "belief–desire pairs." This certainly solves the problem insofar as desires are motivating states aimed at changing the world, yet it only strengthens the suspicion that cognitivism is overintellectualizing the phenomenon since an emotion is now a combination of two cognitive states instead of a single one.

4. *Cognitive Impenetrability:* A judgment such as "I'm in terrible danger!" can usually be overturned easily, were one to receive convincing

information that undermined the judgment by showing that the content of the judgment was not the case. But emotions frequently do not disappear just because one learns that one has no good reason to be scared or offended. Emotions feature cognitive impenetrability, as is nicely illustrated in a passage of David Hume's *Treatise*:

> Let us consider the case of a man, who being hung out from a high tower in a cage of iron cannot forbear trembling, when he surveys the precipice below him, tho' he knows himself to be perfectly secure from falling, by his experience of the solidity of the iron, which supports him . . . The circumstance of depth and descent strike so strongly upon him, that their influence cannot be destroy'd by the contrary circumstances of support and solidity, which ought to give him perfect security. (Hume 1739/2007, I.iii.13)

The man in the cage is undergoing an intense state of fear while at the same time having reason to believe that he is not in danger. This puts the cognitivist in the uncomfortable position of having to explain how we can entertain contradictory judgments at the same time while being fully aware that the judgments in question do contradict each other.[7] If we take judgments to be mental states that figure in inferential processes and are the subject of logical reasoning, the claim that emotions are judgments seems hard to defend.

Cognitivists have several possible ways to deal with this objection. One option is to simply claim that judgments can contradict each other, as long as we do not consciously compare them and find out that they are incompatible. Solomon, for example, argues that this would be a "judgment about a judgment" and that emotions tend to vanish, just as judgments do, in such cases (Solomon 2003, 6; see also Nussbaum 2001, 35f.). A different strategy is to argue that not all emotions have to be conscious and that unconscious states can contradict conscious states since they are the products of two distinct modules of the mind or of two different modes of meaning generation. In this vain, Lazarus argues that many emotions are products of an automatic mode of meaning generation that is unintentional, involuntary, effortless, autonomous, and outside awareness. "If we take the distinction between the two modes of meaning generation seriously, we can believe contradictory things simultaneously—for example, that flying in an airplane is safe and that it is also very dangerous" (Lazarus 1991, 159).

Both answers are problematic. The first one is problematic because emotions can occur together with a further judgment that judges them to be inappropriate, as in Hume's example of a person in a cage who is afraid of the depth although he knows that he cannot fall. Both judgments are present and conscious at the same time, but do not change each other. The person could stay in the cage for hours, maybe days, judging his emotions to be false without his judgments having any effects on his emotions. Lazarus's

argument has a much better chance to capture this phenomenon adequately. It seems much more likely that emotions and judgments can be processed separately and that therefore not all judgments can directly influence emotions. Yet if the claim is that emotions are judgments themselves and are triggered by complex inferential reasoning, how can emotions and judgments be processed separately? The claim that mental states are cognitively impenetrable can be explained through the idea of modularity of thought, yet the information processed in modules is usually thought of as being nonconceptual in format. The reason that this information can be cognitively impenetrable is precisely because this kind of information cannot be processed in the same way as judgments are processed. The objection that emotions are cognitively impenetrable, therefore, seems to be a good reason to refuse cognitivism, at least as long as cognitivism entails the claim that emotions themselves *are* judgments.

5. *Infants and Animals:* Several nonhuman animals exhibit many elements of emotional reactions that occur in a coordinated form. Nonhuman animals show behavioral patterns and bodily expressions like fleeing, attacking, avoiding, shrinking in posture, and wide opening of the eyes, which are similar to our own emotional behaviors and expressions. Neuroscientists study the emotional systems in the brains of rats and monkeys, expecting them to be homologous to ours.[8] While there may be doubt about whether nonhuman animals consciously experience emotions[9] or whether they can show cognitively more complex emotions such as guilt or envy, the claim that such animals do not show any emotions appears to be rather implausible and in contradiction to current data (I discuss some data on emotions in animals in the next chapter). As I elaborate further in the Chapter 4, young infants display emotional reactions as well. Within the first few weeks, infants show signs of pleasure when seeing their mother's face, anger when being frustrated, and react with fear to loud noises. Again, there is disagreement over whether infants are able to consciously experience emotions and over which capacities account for their ability to show apparently complex emotions such as guilt and envy later on. But the claim that infants entertain no emotions until they are able to express themselves in well-formed beliefs and judgments is rare and hard to defend. However, when we try to describe the cognitive capacities underlying the emotions of infants and animals, it certainly makes no sense to say that they entertain judgments, at least not in terms of linguaform propositional attitudes. It is also hard to see how any of the background knowledge that Lazarus, for instance, presupposes, or the kind of evaluative dimension that Nussbaum highlights, could be ascribed to infants. Therefore, it is hard to see how traditional cognitivist approaches can account for the emotional capacities of infants and animals.[10]

There are, again, different strategies to deal with this objection. While Lazarus generally speaks of "adult emotions" and remains silent about infants and animals altogether, neither Solomon nor Nussbaum seems to see any problems in simply including them in their approaches. Solomon in one of his late papers writes, "I take it as uncontroversial that animals make all sorts of judgments (e.g. whether something is worth eating, or worth chasing, or worth courting), but none of these are articulated or 'spelled out,' nor are they subject to reflection. We make nonreflective, nondeliberative, inarticulate judgments, for instance, kinesthetic judgments, all the time" (Solomon 2001/2003, 187).

That animals make "all kinds of judgments" is far from being uncontroversial, as is the definition of judgments as possibly nonconscious, nondeliberative, nonreflective, and inarticulate states. Yet in a similar vein, Nussbaum describes the observable behavior of an animal by ascribing all kinds of judgments to the animal (Nussbaum 2001, 89f.). Again, the only justification for doing so is that judgments can be understood in a very broad way, such that assent, for instance, is not a cognitive attitude but merely a way of accepting what is there that is common to all animals, infants, and adults alike. But there is—again—an obvious problem with this kind of argument: Lazarus and Nussbaum simply stretch the notion of "judgment" to defend their approaches against objections. This makes it complicated to discuss issues without getting lost in terminological confusions. In the end it seems that if we apply Nussbaum's or Solomon's notion of what a judgment is, cognitivism turns into an either trivial or unfalsifiable claim (Scarantino 2010).

6. VICES AND VIRTUES OF COGNITIVISM

Andrea Scarantino suggests distinguishing three forms of cognitivism, or three different possible claims with regard to what role cognitive elements play in emotional processing, that have frequently been confused: "constitutive cognitivism," that is, the claim that emotions are judgments; "etiological cognitivism," that is, the claim that emotions are triggered by judgments; and "representational cognitivism," that is, the claim that emotions have a representational content.

Constitutive cognitivism can be defined like this:

> *Constitutive cognitivism:* For all emotions E, E is theoretically identified with the judgment that the formal object of E is instantiated (by some material object O).

An emotion, such as fear, is identified with a judgment concerning the presence of a formal object, such as danger. Given that there is a material object, such as an upcoming exam, a possible job loss, or a killer in your house, you might judge the situation to be dangerous and thereby entertain a state of fear.

Constitutive cognitivism is not the only approach focusing on the cognitive dimension of emotions. When talking about the "appraisals" involved in emotions many psychologists do not mean to capture the emotion as a whole but only the cognitive evaluative component that causes the emotion. Scarantino calls this second variant "etiological cognitivism":

> *Etiological cognitivism:* For all emotions E, E is caused by the appraisal that the formal object of E is instantiated (by some material object O).

According to etiological cognitivism, an emotion, such as fear, is not identified with a judgment, such as "I am in danger." Instead, the hypothesis is that emotions such as fear are always triggered by some kind of appraisal concerning the danger of the situation at hand. This appraisal then triggers the other components such as the feeling or the physiological arousal involved in emotions. Lazarus is of course a prominent defender of etiological cognitivism.

The third variant of cognitivism that Scarantino characterizes can be labeled as representational cognitivism. Representational cognitivism claims that emotions have a particular kind of representational content. More precisely,

> *Representational cognitivism*: For all emotions E, E *represents* the formal object of E (as instantiated by some material object O).

Fear represents danger. Whether fear is a judgment, another kind of mental state, or triggered by a judgment or another kind of mental state doesn't matter as long as the mental state in question has the representational powers to indicate danger. The main claim of representational cognitivism is that emotions are intentional states of some sort; they have aboutness. Scarantino points out that cognitivists have frequently confused the three varieties of cognitivism he distinguishes and which he takes to be rather different in what they claim and in the kinds of arguments they need.

Since emotions have multiple components, such as the physiological, the phenomenological, and the motivational, and involve simple bodily responses and cognitively more complex reactions, constitutive cognitivism (i.e., the identification of emotions with judgments) leads to two equally undesirable consequences. First, with a conservative understanding of judgments as a linguistically complex form of conscious assent in the background, constitutive cognitivism can easily be proved wrong. Second, with a more liberal understanding of what judgments are, as in Solomon's and Nussbaum's approaches, constitutive cognitivism becomes trivially true or even unfalsifiable. Constitutive cognitivism adds very little to our understanding of the phenomenon, since it introduces a seemingly unitary construct, namely, judgment, which is then described as an umbrella term capturing evaluative, phenomenological, physiological, and behavioral components.

While the unity here is merely terminological and lacks any definition, the diversity or the multiplicity of the heterogeneous parts is equally poorly captured; constitutive cognitivism offers no description of which components contribute to which particular aspects of the phenomenon.

Etiological cognitivism is a more promising approach, since it offers the interesting insight that emotions are always caused by an appraisal. Yet etiological cognitivism has to answer obvious questions: What is meant by an appraisal, and how complex does it have to be? A first possible claim simply says that some sort of information processing must precede every emotional reaction and that even very simple emotional reactions therefore require a cognitive evaluation of some sort. The first part of the claim is uncontroversial. Neither James nor Watson, neither feeling theory nor behaviorism, would deny that even a simple fear reaction in a newborn as a response to an unexpected loud noise requires some sort of processing of the sensory input, before a bodily response can be triggered.

Yet with "cognition," "evaluation," and "appraisal," cognitivists usually mean more than that a stimulus is discriminated from other stimuli thanks to some innate adaptive mechanisms. Lazarus is very clear with regard to this question:

> Innate discriminations and the biologically fixed responses tied to them . . . should probably not be considered cognition even though they constitute an adaptive distinction between danger and no danger. . . . The discriminative meaning is, in effect, a property of the nervous system . . . and presumably requires little or no actual experience and learning.
>
> (Lazarus 1991, 128)

Lazarus goes on to distinguish between two modes of appraisal: one automatic, unreflective, and unconscious or preconscious; the other deliberate and conscious (Lazarus 1991, 128). Yet automatic and unreflective appraisals are still learned cognitive responses and not hardwired neural discriminations.

The distinction drawn by Lazarus has therefore been rejected by several authors. Emotions can be triggered on a simple reflex-like level when they occur as a response to an unexpected loud noise or a quickly approaching object. Emotions can also be triggered through facial expressions and drugs and are frequently caused through neural pathways dedicated to affective processing that, according to Lazarus's definition, should not be called "cognition" or "appraisal." Scarantino is therefore right in suggesting that the main task for etiological cognitivism is to come up with plausible categories of more and less complex appraisals:

> A viable Etiological Cognitivism, I suggest, needs to conceive of "appraisal" as lying on a continuum between primitive and sophisticated forms of information processing. Understood in these terms, Etiological Cognitivism

> contains an important insight: stimuli cause emotions not per se, but
> contingently upon how they are interpreted.
>
> (Scarantino 2010, 754)

Yet there are two problems that have to be kept in mind. First, if what
Lazarus calls "innate discriminations" are not counted as being appraisals,
many instances of emotions, such as fear in response to a loss of balance or
joy in response to the facial feedback from a smile, are excluded. Second, the
talk about "inarticulated, nonconscious, nondeliberate" cognitive processes
must be based on a theory of different levels of cognitive processing. The
way in which Lazarus and many other cognitivists refer to a "lower level of
cognition" appears to be somewhat of an unexplained explainer: cognition
on this level works just as the inferential processes do on the conscious level,
only it is not conscious, contradictions can occur there, the representations
processed there are not articulated, and so on. To simply claim that on a
lower level of cognition everything is basically the same and also completely
different from cognition on a higher level doesn't seem to explain much
about this "level" of cognition.

Representational cognitivism captures the most basic and most impor-
tant insights cognitivism brings forward: emotions have an intentional
content that has to be explained, and because of this content, emotions
cannot be reduced to mere bodily feelings or behavioral reactions. As I have
shown earlier, in cognitivist approaches the claim about the intentionality
of emotions comes together with the claim about their intensionality, and
the explanation of their intensionality usually comes together with massive
claims about the complexity of the cognitive processes involved in emotions
and frequently leads both to the claim that emotions are judgments, and to
the claim that emotions are caused by judgments.[11] A task for the following
chapters is to examine the possibility of ascribing a less demanding sort of
nonconceptual content to emotions.

However, with an account that combines all three varieties of cognitiv-
ism, Lazarus fully captures the normative assessability of emotions and
develops an impressive taxonomy of the emotions. He captures the rela-
tional meaning of each emotion type and precisely defines what kind of
background knowledge and appraisals have to be made to elicit an emotion
with such a content. At the same time, Lazarus's account is a target of all five
objections that have been raised against cognitivism. He only speaks about
"adult emotions," leaving open the question of how emotions in infants and
animals should be explained. He tries to explain cognitive impenetrability
with reference to an unconscious level of thought without explaining why
it should be possible for contradictions to occur on such a level. He admits
that emotions do have a phenomenological component but treats it as a side
effect without any function. With regard to the passivity of emotions, Laza-
rus would probably say that the judgments involved in eliciting an emo-
tion can work automatically. With regard to the motivational component,

Lazarus claims that emotions are usually combined with a desire and are therefore motivating. But this adds another complex cognitive component to the approach.

As an upshot of his multisided cognitivist approach, Lazarus compiles an impressive taxonomy that explains emotional content while both respecting the normative assessability of the emotions and also showing how the normative assessability of each emotion type can be explained with regard to the background knowledge and appraisals involved in the production of the emotion. Yet the price that Lazarus pays is a strong overintellectualization of the phenomenon. An emotion such as anger appears to be present in infants and animals and can be present in a person that is perfectly aware of the fact that there is no reason to be offended. Think of an infant reacting with anger when being captured in too tight an embrace or of a mother being terribly annoyed by the behavior of her child for hours and then "erupting" in anger while being perfectly aware that there is no reason to blame the child. These simple examples are hard to reconcile with the view presented in detail earlier that anger represents a demeaning offense against me and mine that is triggered when we judge a situation to be incongruent with our goals. Lazarus's view becomes even more implausible when we add all of the following to every instance of anger: that we have to judge that in this situation we should preserve or enhance our self- or social esteem; that we have to blame another person for how she acts against us or other people, that we have to judge that she is accountable for the harmful actions she is committing and that she could have controlled them, and, finally, that before we react with anger we check the coping potential and the future expectancies of the current situation.

Let me briefly take stock of the kind of background knowledge this definition of emotion requires that (1) one must have a concept of one's goals and (2) one must be able to check whether current situations are congruent with them or not. (3) One must have a concept of oneself and other persons and be able to understand the relations between oneself and others. This entails that (4) one must have a concept of responsibility for the actions one commits. Furthermore, the social relations in question are normative ones, so (5) one must have an understanding of social rules, norms, and values, such that one can understand in which cases social esteem needs to be preserved versus the cases in which somebody violated a rule so that blaming her is appropriate. Finally, to check coping potential and future expectancies, (6) one must be able to think about possible future situations and therefore must be capable of modal reasoning.

As far as I know, not even the toughest rationalist or nativist has ever claimed that these abilities are present and fully developed in infants before the age of three. Lazarus might argue that all these processes occur on an automated, inarticulate, and unconscious level. Yet who would seriously claim that infants and animals engage in normative or modal reasoning on an unconscious, automated, and unarticulated level? Lazarus's view has

the deeply implausible consequence that an infant that reacts angrily to an embrace that is too tight is seriously blaming her caregiver, judging that the caregiver is responsible for the harmful action she has just committed. The problems of this view cannot be solved by the claim that the complex reasoning processes that seem to involve all kinds of conceptual knowledge could be unconscious, unarticulated, and automated.

7. CONCLUSION

Cognitivists rightly point out that emotions have an intentional content with an intensional shape, are subject to semantic norms, are rationally constrained, and are sometimes concerned with social norms. In particular, the work of Lazarus shows, in an exemplary way and with these ideas in the background, that a very precise taxonomy of emotions can be developed. Yet his work also shows that a detailed taxonomy of the emotions and their cognitive contents leads to a heavy overintellectualization of the phenomenon. His account cannot offer any explanation of emotions in infants and animals and his attempt to explain cognitive impenetrability with reference to an unconscious automatic level of cognition leaves us with an unsatisfying picture of what cognition on such unconscious, automatic levels could be like.

Other cognitivist accounts, however, have stretched the notion of what a judgment is so far that the claim that emotions are judgments becomes completely trivial. Yet, as will become clear in Chapter 3, more recent and moderate approaches usually fail to account for normative assessability. The main question that remains, therefore, is how an approach that can account for the normative assessability of emotions can avoid either constitutive cognitivism, a strong version of etiological cognitivism (claiming that all emotions are caused by cognitive processing), or a strong version of representational cognitivism (claiming that emotions essentially involve a conceptual content).

For an embodied account the main task is one of replacement (Shapiro 2010): what cognitivists describe in terms of symbolic processing has to be replaced with a description of how the skillful body and the structured environment interact in an intelligent way. In the next chapter I suggest that emotions are constituted by embodied reactions, which can be type-identified with regard to their adaptive functions. Such an approach can avoid the overintellectualizing claims altogether by describing emotions as ecological responses whose function is to react to situations with particular significance for the organism. Yet it should now be clear that such an approach will have a great deal of work to do to develop a taxonomy that can be compared with Lazarus's work. Such an account has to explain what constitutes the emotions' normative assessability if it is not background knowledge and appraisal hierarchies of the sort described earlier. This is the task for the chapters ahead.

NOTES

1 This interpretation of James is in itself not uncontroversial; see, for example, Ratcliffe (2005) for an interpretation of James in the broader context of his works, which argues that James indeed sees emotions as intentional states.

2 One might wonder whether "irrevocable" isn't too strong here since the formulation excludes many cases of sadness where the loss one suffers from is not irrevocable.

3 For a defense of this claim see, for example, Parkinson, Fischer, and Manstead (2005).

4 The approaches of Solomon, Nussbaum, and Lazarus discussed earlier are all good examples of such a view.

5 Catherine Newmark (2008) convincingly shows that emotions have been characterized as passive states in the philosophical tradition ranging from Aristotle to Spinoza and Kant. See Prinz (2006b) for a current account defending that position.

6 Scarantino convincingly argues that Nussbaum and Solomon in particular stretch the notion of what a judgment is so that it suits all emotions but becomes a rather empty notion. He calls this the elastic strategy (Scarantino 2010).

7 See, for example, Tappolet (2000, 147).

8 See, for example, Panksepp (1998).

9 Damasio (1999), for example, argues that emotional processing is not always conscious in humans and never becomes conscious in animals.

10 See, for example, Tappolet (2000, 145) and De Lancey (2002, 41).

11 Lazarus explicitly claims both at the same time while Solomon claims both without ever distinguishing them.

2   Appraising Arousal
Emotions and the Body

1. INTRODUCTION

"What kind of an emotion of fear would be left," William James famously asks, "if the feelings neither of quickened heart-beats nor of shallow breathing, neither of trembling lips nor of weakened limbs, neither of goose-flesh nor of visceral stirrings, were present, it is quite impossible to think" (James 1884, 194f.). The claim that emotions are embodied could appear rather trivial at first, since it has often been remarked and is hard to deny that emotions somehow involve bodily reactions: our hearts race when we are afraid, we blush and hang our heads in shame, and we cry when we are sad. Aristotle describes the blood as boiling around the heart in anger (*De Anima* 1.1.403), and Descartes, in *The Passions of the Soul*, observes in a remarkably detailed way the different bodily impulses involved in emotion:

> In joy . . . the pulse is even and quicker than ordinary, but not so strong, nor so grave as in love, and . . . a man feels a pleasant heat, which is not only in the breast, but spreads itself over all exterior parts of the body with the blood, which is seen to flow abundantly thither. And meanwhile, he sometimes loses his appetite because his digestion is less than usual.

> (Descartes 1649/1988, § 99)

In the works of Charles Darwin we find careful descriptions of all kinds of bodily and expressive reactions involved in emotions in animals including humans. Darwin aims to show the common causes and origins of, for example, trembling in fear, hair erection and muscle tension in rage, and the common human expressive form for sadness and grieving:

> Persons suffering from excessive grief often seek relief by violent and almost frantic movements . . . but when their suffering is somewhat mitigated, yet prolonged, they no longer wish for action, but remain motionless and passive . . . The circulation becomes languid; the face pale; the muscles flaccid; the eyelids droop; the head hangs on the

contracted chest; the lips, cheeks, and lower jaw all sink downwards from their own weight.

(Darwin 1872/2009, 165)

Yet the discussion of cognitivist approaches in the last chapter showed that cognitivist authors in recent decades have frequently reduced the status of the bodily and behavioral reactions involved in emotions to by-products that neither are necessary nor play any explanatory role. With regard to the traditional views ranging from Aristotle to Darwin referred to above, it is noteworthy that all of these authors ascribe an important role to bodily processes in emotions, yet the concept of what a body is in the first place differs depending on the theoretical background of the different authors. The different theoretical frameworks generate diverse approaches to the precise explanatory role bodily reactions are supposed to fulfill. My own view has its origins in the works of Darwin and James, which is to say in early pragmatic and naturalistic reasoning about the mind. The idea that mental states do not just mirror or represent the world but are like tools that help to guide action has its origin here, as does the claim that complex organisms, with all their cognitive abilities, incrementally evolved out of simple ones. What this implies for the study of emotions is expressed very clearly in James's famous article on emotions:

> [T]he nervous system of every living being is but a bundle of predispositions to react in particular ways upon the contact of particular features of the environment. As surely as the hermit crab's abdomen presupposes the existence of empty whelk-shells somewhere to be found, so surely do the hound's olfactories imply the existence, on the one hand, of deer's and foxes' feet, and on the other, the tendency to follow up their tracks . . . our wrath at snakes and our fear of precipices, may all be described similarly, as instances of the way in which peculiarly conformed pieces of the world's furniture will fatally call forth most particular mental and bodily reactions, in advance of, and often in direct opposition to, the verdict of our deliberate reason concerning them.
>
> (James 1884, 190)

What is suggested here is a general view of emotions as constituted by homeostatic reactions concerned with the organism's well-being that evolved out of simple reflex-like bodily reactions and are adapted to certain cues in the environment, so that once we meet the features in question, in advance of deliberate reasoning, we feel motivated to react in a certain way. Emotions, according to James, are "nervous anticipations" called forth directly by the perception of certain facts and preparing the whole organism for action (James 1884, 191).

What follows places into a framework of contemporary research on embodied cognition the basic Jamesian idea, that emotions are realized by

embodied reactions of the whole nervous system, which are quicker than and independent of rational deliberation, and the Darwinian idea that emotional reactions have common origins in history. Yet, as will become clear in the discussion of empirical studies, the claim that emotions somehow involve bodily arousal is as obvious and easy to defend as it is difficult to come up with data explaining the precise role of the bodily reactions involved in emotions. Therefore, most of the questions that have been raised with regard to James's original claim remain subjects of heated debates: Do all emotions have a characteristic pattern of bodily arousal such that different types of emotions can be individuated with regard to the involved bodily reactions? Is the pattern of bodily arousal a panculturally stable feature, or does it vary among cultures or among individuals? How are the bodily reactions connected to the brain, to feelings, and to rational thinking? And, if emotions are evolutionarily developed reactions, what is their adaptive function?

It is therefore one aim of this chapter to sketch the classical Jamesian hypothesis and discuss objections from both psychology and philosophical cognitivism, as well as to discuss in which form the Jamesian claim can still be defended. For James and the discussion he initiated, the question about the form in which emotions are embodied has been closely intermingled with the question whether emotions can be type-identified with regard to the bodily reactions they involve. I refer to the claim that they can be type-identified with regard to their bodily profile as the "Jamesian hypothesis." The claim that the bodily reactions and behaviors involved in emotions have a common origin I label the "Darwinian hypothesis." Finally, I take embodied emotions to be constituted by bodily response patterns that are dynamically organized and realize an intelligent access to the world. I call the claim that such an account can be reasonably defended the "embodied cognition hypothesis."

In what follows I present the Jamesian hypothesis and objections that have been raised against it. I then argue that the Jamesian view has to be put into a broader theoretical framework, mainly to explain that emotions are multiply realizable. I argue that the Jamesian hypothesis cannot be defended without defending the Darwinian hypothesis, and *vice versa*. The bodily reactions involved in emotions cannot be seen as essential ingredients since they are multiply realizable. To type-identify emotions we therefore need functional descriptions of what role they play for the organism. On the other hand, functional descriptions cannot be constructed without considering the implementation level and the evolutionary origins of a trait. We need to know how an emotion is realized in actual living organisms. In humans as it turns out emotions are evolutionarily acquired, highly plastic and dynamically organized patterns that represent core relational themes and prepare for action. While I argue in this chapter that different emotion types are realized by different functional patterns of bodily arousal, it is my task in the following chapters to show in what way these functional mechanisms realize a certain kind of intelligent access to the world.

2. THE JAMESIAN HYPOTHESIS AND ITS PROBLEMS

Taking a closer look at current theories on emotions, it is striking that the precise explanatory role bodily reactions play in theories of emotions radically varies. On one end of the spectrum, Nussbaum (2001) argues that while emotions involve bodily sensations of many kinds, there are no constant correlations between specific bodily processes and certain kinds of feelings, thus bodily reactions appear to be a coincidental by-product of emotions and not an essential ingredient. I call this the "coincidental by-product view." Nussbaum's position represents a broad consensus among cognitivists. While cognitivists tend to argue that emotions are constituted or caused by cognitive states such as beliefs or judgments, they tend to downplay the role of the body in emotional processing.[1]

James, on the contrary, in his classic essay "What Is an Emotion?" (1884), argues that the feeling of bodily arousal is what constitutes an emotion. The commonsense view might be that emotions simply cause bodily arousal; when we are sad, we therefore start crying or when we feel embarrassed, we blush. Yet James inverts the commonsense view and argues that in most cases, we perceive something, like a dangerous predator or an offensive opponent, and this perception directly triggers bodily arousal, which is felt only afterward. An emotion is the perception of autonomic nervous system changes. James's central argument for this is of a phenomenological character: if you imagine a strong emotion and try to subtract the feeling of all involved bodily symptoms, nothing remains "but a cold and neutral mental state of intellectual perception" (James 1884, 193). I call this the "subtraction argument."

On a closer look, however, James makes two related claims. The first claim is that a mental state that does not involve the perception of bodily arousal would not be classified as an emotion at all. The second claim is that different kinds of emotions owe their special character to the different kinds of bodily arousal that constitute them. Both claims are nicely illustrated in the following quote: "What kind of an emotion of fear would be left, if the feelings neither of quickened heart-beats nor of shallow breathing, neither of trembling lips nor of weakened limbs, neither of goose-flesh nor of visceral stirrings, were present, it is impossible to think" (James 1884, 193f.). While these are some of the physiological elements involved in fear, rage, according to James, is unimaginable with "no ebullition of it in the chest, no flushing of the face, no dilatation of the nostrils, no clenching of the teeth, no impulse vigorous to action, but in their stead limp muscles, calm breathing, and a placid face" (James 1884, 194). The bodily processes involved in emotions are their essential ingredients. Subtract them away and the mental state that remains is classifiable neither as a state of fear nor as an emotion at all. I call this the "essential-ingredient view." James's essential-ingredient view has been the target of many objections, the implications of which I discuss later.

The general claim that an emotion without any bodily arousal involved is a nonentity has been denied by both a conceptual and an empirical argument. The conceptual argument, as Nussbaum articulates it, employs the possibility of bodiless creatures having emotions: "Whether we believe that bodiless substances exist or not, the reason it makes sense to imagine a bodiless substance having genuine emotions is that it makes sense to imagine that a thinking being, whether realized in matter or not, could care deeply about something in the world, and have thoughts and intentions associated with such attachments. And that's all we really require for emotion" (Nussbaum 2001, 60). The argument claims that it is conceptually possible to think of an emotion as an intense state of mind that is not realized in any kind of matter at all. Actually, James happily admits that it is no "contradiction in the nature of things, or that pure spirits are necessarily condemned to cold intellectual lives" (1884, 194). So, with regard to this particular argument there is no real disagreement between the coincidental by-product and the essential ingredient view. What James instead wants to claim is that "*for us*, emotion dissociated from all bodily feeling is inconceivable" (1884, 194, emphasis added). After all, from a naturalist point of view, the mere conceptual possibility of disembodied emotions in possible worlds where thinking substances without a body exist, can be conceded. But what is of interest is how emotions are realized in living organisms in the actual world. And here, the hypothesis at stake is that to "care deeply" about something or someone involves specific bodily processing such as the release of oxytocin.

Therefore, it is the empirical argument that really is of importance for the evaluation of the Jamesian account. Walter Cannon (1927) did research to test the Jamesian hypothesis as early as the 1920s. In his experiments, the spinal cord and vagus nerve in dogs and cats were cut to inhibit feedback from the body to the brain. Then the animals were tested with regard to their reactions to emotional stimuli like threatening shocks. Their ability to react emotionally to such stimuli appeared to be intact. According to Cannon, the animal tests prove that even if you cut the bodily feedback entirely, feelings do not seem to be reduced. Even today, Cannon's article ranks high on the list of the most quoted evidence for all variations of the coincidental by-product view.[2] It is often combined in one vein with a study by Chwalisz, Diener, and Gallagher (1988) who tested patients with spinal cord injuries. The outcome suggests that their emotional lives remain relatively stable. However, as William James already knew, cutting the vagus nerve and the spinal cord alone is not enough to prove him wrong:

> A crucial test of the truth of the hypothesis is quite as hard to obtain as its decisive refutation. A case of complete internal and external corporeal anaesthesia, without motor-alteration or alteration of intelligence except emotional apathy would afford, if not a crucial test, at least a strong presumption, in favour of the truth of the view we have set forth;

whilst the persistence of strong emotional feeling in such a case would completely overthrow our case.

(1884, 203)

More recently, Antonio Damasio (1999) carefully detailed the paths bodily feedback can take and which forms of damage could eliminate or reduce it. While it is true that James focuses mainly on feedback from the viscera, an embodied account can include more than that—in particular, musculoskeletal feedback from the body and the face is apparently quicker and contributes to interoception during emotions. That spinal cord–injury patients do not lose the ability to feel emotions can first be explained by the fact that in these patients the vagus nerve usually is still transporting information. Furthermore, only a fraction of the enactment of emotions depends on the spinal cord. A large number of processes are mediated by cranial nerves at the brain-stem level that can act on the face and on the viscera and by other brain-stem nuclei that can act directly on the parts of the brain above their own level. Also, a significant part of bodily feedback does not travel via the nerves at all, but by way of the bloodstream from where it directly enters the brainstem. Nevertheless, that the reduced amount of bodily feedback might also reduce emotional feelings is suggested by Hohmann (1966), who interviewed patients with a spinal cord lesion and found that they reported subdued experience for many emotions except sadness and found a correlation between the location of the spinal lesion and the degree to which their emotional experience was reduced. But even if the location of the lesion is high on the spine, this does not suggest that patients should entirely lack emotional feelings since the body does not end above the neck. The face, skull, oral cavity, and tongue provide a massive input into the brain that enters at brain-stem level and cannot be inhibited by any spinal cord injury. Furthermore, most emotions express themselves in changes of the facial musculature, the musculature of the throat, and in autonomic changes of the skin in the face and scalp. The representation of these changes in the brain remains available even with spinal cord injuries (Damasio 1999, 290).

Similar arguments can be brought up against Cannon's interpretation of the animal experiments. Why should Cannon have expected his prediction, that cats or dogs with severed spinal cord and vagus nerve should have a complete loss of emotional display, to be true in the first place? A severed vagus nerve and spinal cord do not impede the pathways for the facial responses that come from the brain stem and are mediated by cranial nerves, which were not compromised in Sherrington's or Cannon's experiments. It's therefore not surprising that the cats and dogs in the study showed anger expressions, even if they could not move their bodies (Damasio 1999, 291).

Furthermore, the specific Jamesian claim is that emotions are feelings of bodily arousal. This is a rather narrow way of capturing the phenomenon. What I want to capture with the term *emotion* is a complex pattern of bodily and behavioral reactions, which realize an action-oriented

representation in situations that are relevant for the organism's well-being. Feelings, just like cognitive appraisals, are important parts of these patterns but are not necessary ones since emotions can remain unconscious and can be elicited by unconditioned stimuli without involving cognitive processing of any sort. This view is one that Cannon, I assume, did not oppose, since he pointed out that peripheral reactions are of great importance in readying particular organ systems for certain kinds of action. Fear, for example, arises in situations where an animal is likely to get hurt and should therefore flee. The release of adrenaline, according to Cannon, not only prepares the muscles for flight but also enhances the coagulation of blood in case the animal is wounded (Cannon 1936). In fact, the hypothesis defended here, that emotions are embodied, is committed to the claim that bodily changes of all kinds as they appear in emotions assume an important functional role. This does not necessarily exclude the Jamesian point that bodily arousal can be felt, but the question becomes more of a sideline to the debate than its central issue.

Another anti-Jamesian argument is that bodily reactions cannot be constitutive of emotions, since the reduction of bodily feedback does not always lead to reduced emotional experience. Yet the studies cited to prove this claim are, at the very least, highly ambivalent. Reisenzein and Stephan (2014, 43) argue, for example, that studies in which facial feedback has been partly blocked by Botox injections (Davis, Sengas, Brandt, Ochsner 2010) do not support the claim that facial feedback is necessary for emotions. Yet neither James's nor my version of an embodied theory is committed to the claim that facial feedback is necessary for the occurrence of an emotion. The claim is that emotions (and thereby emotional feelings) are constituted through patterns of bodily arousal. The absence of one component of the bodily reaction should have an effect on the realization of the emotion and might affect how the emotion feels but should not necessarily prevent its occurrence. This theoretical prediction fits with the results of the studies quite well. They show that subjects who received a Botox injection showed emotional muting in response to a mildly amusing video clip, while no such response was observed in strongly negative or positive clips. A plausible interpretation of this result is that in the case of strongly emotion-eliciting clips there is no remarkable difference in experience because there are many bodily reactions going on, while in the case of mild joy the feedback that stems from a smile is a central factor of the bodily reaction so that its absence makes a difference. Similar arguments can be made for other cases of reduced bodily feedback as well.[3]

An open question, certainly, is whether all states that the folk concept of emotion is supposed to capture really include the suggested kinds of bodily reaction. Sometimes we say that we are angry about something we read in the newspaper, when what we mean is rather an intellectual stance of disapproval. It is not my aim to capture all these states in a definition that precisely mirrors the folk concept of emotions (whatever this might be). But

it certainly is my aim to develop an empirically well-supported definition giving a description that captures a great deal of our everyday understanding of what emotions are.

With regard to the second claim, that particular kinds of emotions owe their special character to the pattern of bodily arousal that constitutes them, the disagreement between the essential-ingredient and the coincidental by-product views is grounded in phenomenological and empirical arguments. While, as we have seen, James considers the reactions of the nervous system in fear and rage to be distinct, according to Nussbaum the feeling of bodily arousal involved in an emotion can vary. While "many men report experiencing anger in connection with a boiling feeling," Nussbaum states, "my own experience of anger is that it is associated with tension at the back of the neck, or a headache that appears the next day" (Nussbaum 2001, 61). This phenomenological self-observation is meant to refute the idea that every type of emotion is associated with a specific kind of bodily arousal. Whether rage is associated with boiling blood, or tensed muscles around the neck, or another bodily reaction, or no bodily reaction at all can vary from person to person and is therefore no criterion by which to individuate emotions. From merely the phenomenological point of view the disagreement between the essential-ingredient view and the coincidental by-product view therefore seems hard to decide because reports in favor of either view can be found.

Yet Nussbaum's argument appears to find support in the studies done by Stanley Schachter and Jerome Singer (1962), who injected subjects with adrenaline and exposed them to different scenarios eliciting mild anger or happiness. The result was that test subjects showed signs of anger when they were given an insulting questionnaire and signs of happiness when brought into a room where a stooge was doing silly things. The interesting point is that the reactions of the test group differed from those of the control group, who were not given an anterior adrenaline injection: members of the control group became less emotionally involved with the presented scenarios. Schachter and Singer concluded that autonomic arousal plays a role in emotion insofar as it is perceived as undifferentiated arousal and creates an evaluative need. The subject, feeling her heart beating faster and finding herself in front of an annoying questionnaire, labels her own state as anger and reacts accordingly. Thus, Schachter and Singer's experiment can be seen, as it has been for several decades, as evidence for the claim that emotions cannot be individuated according to a type of bodily arousal they alone involve.

Yet the Schachter–Singer study has been shown to be flawed in several ways. To repeat only one of the arguments brought up against it, it is highly doubtful whether the mere injection of adrenaline causes the same sort of physiological arousal in various subjects in the minutes that follow. The participants might be more sensitive to the happiness- or anger-inducing scenarios because of the adrenaline but because of the perception of the anger-eliciting task together with the adrenaline in the blood might cause

further bodily changes that are anger-specific, while the adrenaline combined with the funny situation causes bodily arousal that is specific to joy (see Prinz 2004, 70f.).[4] So the frequent claim that empirical studies prove that emotion types cannot be distinguished by the bodily reactions they involve is not as well supported by evidence as it might seem at first.[5] This will become clear in the discussion of recent evidence in the next section. A closer view at the results of recent research render it highly likely that emotions involve something like core patterns of bodily reactions that are evolutionary acquired and can be found cross-culturally.

3. EMPIRICAL EVIDENCE FOR BODILY REACTIONS IN EMOTIONS

The possible sources of evidence of the bodily reactions involved in emotions are numerous and far from having been tested and measured in detail. Therefore, even detailed overviews of recent studies are far from issuing clear and stable results. In the following I give a brief and highly selective overview, which aims to show that there is clear evidence against the coincidental by-product view, good evidence for evolutionarily acquired distinct patterns of bodily reactions in many emotions and that it is an open question how closely emotion types and discrete physiological responses are related. As has already been mentioned, the place to look for bodily reactions is not only the inner organs and the visceral feedback stemming from them, but also the endocrine system and the somatic nervous system that innervates skeletal muscles, including those of the face.

Indeed, research focused on facial expressions (and more recently on body posture) has been most successful. Therefore, I start there and selectively extend the discussion into other fields, without aiming to cover the range of existing data completely. Paul Ekman has investigated the facial expressions accompanying emotions. By carrying out research on facial expressions for more than thirty years he has established a theory of basic emotions, which are expressed in ways that can be recognized panculturally. Ekman and his colleagues developed the facial action coding system, which codes observable facial muscle-movements, to test whether their configurations could be identified with the expression of specific emotions across cultures (Ekman and Rosenberg 1997). The research on reidentifiable facial expressions in many cultures, including some that barely had any contact with the Western world, resulted in the hypothesis that several facial expressions occur and can be recognized universally. The emotions for which Ekman found stable recurring patterns are fear, anger, joy, sadness, disgust, and surprise (Ekman 1971; see also Matsumoto et al. [2010] for a current synopsis).

Apart from the visual observation of facial expressions, which Ekman and colleagues performed, nowadays there is a growing amount of research using facial electromyography. Facial electromyography (EMG) directly

measures muscle activity by detecting and amplifying the electrical impulses generated by muscle fibers when they contract. The lowering and contraction of the brows, for example, has been observed as part of the expressive reaction to unpleasant stimuli. It is involved in producing frowns. Therefore, the firing of motor units in this muscle region can be expected if a stimulus is judged to be unpleasant. But facial EMG studies have also found that activity of the corrugator muscle, which lowers and contracts the eyebrow, varies inversely with the valence of the observed stimuli and the reported emotional state, which means that the muscle, in neutral situations, is under slight tension and relaxes in the presence of pleasant stimuli (Schwartz et al. 1979).[6] This supports the conclusion that the corrugator muscle serves as a permanent bodily valence marker by sending proprioceptive feedback to the brain, which indicates the presence of pleasant, unpleasant, and neutral stimuli.

While the research on facial expressions alone only led to the discovery of the six aforementioned basic emotions (or, rather, seven, since Ekman later added contempt to the list), recent studies have extended the view to combinations of facial expression and body posture and found evidence for cross-culturally stable patterns of pride and shame as well (Tracy and Robbins 2007a, 2007b and Clark 2010; see the following discussion). It is likely that further research will reveal evidence for even more emotion types. Facial expressions do not occur in isolation but have been shown to be systematically associated with physiological responses, such as skin conductance, cardiovascular activation, and somatic activity (Mauss et al. 2005), and with neuroendocrine activity (Lerner et al. 2005). Lerner et al. were able to show that fear expressions are associated with elevated heartbeat and cortisol levels, while anger and disgust show reduced levels.

With regard to the reactions of the autonomic nervous system, the evidence is still ambiguous. Yet hardly anybody would claim today that there is no covariance between reactions of the nervous system and particular features of emotions at all. Even Lisa Feldman-Barrett (2006), who assumes that there are no emotion kinds with a unique and invariant autonomic signature, suggests that patterns of physiological responses can be divided into more general dimensions of threat and challenge, and positive and negative valence (see also Taylor 1991, Cacioppo et al. 2000). If we look at emotion *patterns* rather than *individual variables* of autonomic reactions, *no* current position would deny differences among coarse-grained emotional dimensions (Mauss and Robinson 2009, Harrison et al. 2013).

Recent overviews even suggest that emotion types can be much better individuated via patterns of autonomic nervous system (ANS) arousal than was previously thought. A main reason for this is its greater sensitivity to fine-grained differences and to the various sources of those differences. While most studies are, to a large degree, based on measures of cardiovascular, respiratory, and skin conductance responses, when it comes to individuation a lot depends on measuring more and more correct factors to

individuate a certain emotion type. Sylvia Kreibig (2010) points out that, in the case of disgust, measurements of the increased heart rate and skin conductance tend to resemble other negative emotions like fear and anger. Surprisingly, there are few tests that look at gastrointestinal responses that might correspond to disgust, although lay association could draw such a connection, and the few studies that tested for stomach electrical activity could actually find reactions.

Fear is among those emotions for which there are stable results for a unique ANS pattern already. Fear studies have used a number of different fear-induction paradigms and identified a broad pattern of sympathetic activation that includes increased heart rate, narrowing of the blood vessels and electrodermal activity, accompanied by an increase in respiratory activities associated with a decrease in blood carbon dioxide levels (Kreibig 2010, Harrison et al. 2013). Anger (elicited through harassment or personalized recall) is associated with a general increase in sympathetic activity as well; it also includes an increase in heart rate and in systolic and diastolic blood pressure, but, unlike fear, anger is associated with total peripheral resistance (Harrison et al. 2013). Some studies also found that in anger there is an increase in forehead temperature, probably due to a general increase in facial circulation (Stemmler et al. 2001). Analogously, an increase in facial blood flow in anger was found in rhesus monkeys (Nakayama et al. 2005). Happiness, like many negative emotions, is associated with cardiac activation, increased blood pressure, and increased respiratory activity. However, in contrast to many negative emotions, it is associated with peripheral vasodilation (Harrison et al. 2013). When comparing different types of emotions, there are similarities such that most positive and negative emotions involve some kind of arousal, which tends to be stronger in negative emotions such as fear and anger. However, taking a closer look at fear and anger in particular, it turns out that anger appears to be more strongly associated with vascular activity than fear, but less strongly associated with cardiac activity. These results render it highly likely that emotions involve something like core patterns of bodily reactions that are evolutionary acquired and can be found cross-culturally (see also Colombetti 2014).

It must be noted, however, that there is a huge amount of data lacking in this field and a considerable amount of confusion and seemingly contradictory results brought forth in recent decades. For example, while Cacioppo et al. (2000) found fear to be associated with a greater heart rate increase than anger, Labouvie-Vief et al. (2003) found that imagery-induced fear and anger elicited comparable increases in heart rate. A plausible explanation for such differences is that many studies neglect the importance of context-dependent aspects of emotional reactions. Different studies compared, for example, the physiology of fear in situations as different as hearing a loud noise, looking at a picture of an amputated leg, giving a public

speech, or imagining an intruder in one's house (Bradley 2007). Cardiac reactions also tend to differ when subjects process affectively similar events in different types of tasks. The results of the study by Labouvie-Vief et al., for example, might be explained by the fact that heart rate tends to decelerate during unpleasant picture viewing, while it accelerates when unpleasant mental imagery is evoked (Lang et al. 1990). The context of perception or imagination has a strong influence on the heart rate. A more obvious case is that the psychophysiology of emotion in action can differ from vigilant anticipation. Even without further evidence, one can make good guess about the difference in heart rate measured when somebody meets a snake while she is jogging through the woods compared to somebody lying on the couch and seeing a snake suddenly appearing on TV. Therefore, the parameters of each task, which have implications for the motivational action, have to be carefully distinguished before comparing studies. While one could try to come up with bodily arousal patterns relative to the situations in which they occur, finding stable patterns is further complicated by the degree to which individual differences influence patterns of physiological reactions, including interpersonal differences in gender, age, temperament, and intellectual capacity, such as imagination and memory, just to name a few.

A further complication stems from the fact that the presence of specific emotions is not only inferred on the basis of physiological reactions but also on the basis of the behavioral actions they cause, and these behaviors tend to vary depending on contextual support. In response to an electric shock, even rats show a variety of different behaviors depending on their actual environment. If there is an exit, they escape; if there is another rat, they attack; and if there is neither, they freeze (Lang et al. 1990). This tactical nature of emotional behavior, which is already present in animals far less intelligent than humans, constrains the efforts of psychophysiologists to make general inferences regarding the physiology of specific emotional states since freeze, fight, and flight reactions pose different metabolic demands and therefore influence the emotional arousal evoked by emotional processes (Harrison et al. 2013).

The preceding evidence makes the Jamesian hypothesis, that a specific emotion has a specific physiological pattern, look too general and difficult to test since the physiological pattern of an emotion depends on the individual, on its concrete circumstances, and on the kind of task in which the emotion-eliciting cue appears. This not only calls for clean parameters to be defined in studies on the psychophysiology of emotions. It also raises the question of whether it makes sense to think of the bodily arousal involved in emotions in terms of an essential ingredient, that is, in terms of a pattern with stable features that can be reidentified in the various occurrences of an emotion-type in a subject over time as well as among different subjects.

Although there is considerable evidence for recurring patterns of bodily reactions in emotions, the essential ingredient view has no answer to a

simple functionalist argument such as the one Nussbaum has constructed to support her view:

> All human experiences are embodied, and thus realized in some kind of material process. In that sense, human emotions are all bodily processes. But the question is, are there any bodily states or processes that are constantly correlated with our experiences of emotion, in such a way that we will want to put that particular bodily state into the definition of a given emotion-type? And here we run up against an issue well-known to biological researchers: the plasticity of the human organism, or, in other words, the multiple realizability of mental states.
>
> (2001, 58f.)

Nussbaum here refers to the functionalist paradigm in philosophy of mind that was elaborated by Hilary Putnam (1960) and others in the context of computationalist theories of mind. At the heart of functionalism lies the idea that mental states like ours might be realized in several ways and should not be type-identified with reference to their particular neural underpinnings but, instead, with reference to their functional role.

Multiple realizability is a central concept in this context. It can be understood in two ways: first, with regard to an individual and, second, with regard to different kinds of living organisms. With regard to the individual, the findings on the plasticity of the brain suggest that the brain is a functional organ where, for example, after a lesion the functions realized by the damaged areas can come to be realized by other areas of the brain. Therefore, it does not make sense to claim that a particular arrangement of neurons in a particular individual realizes one particular type of mental state. Across species there are obviously functionally identical mental states, like pain or hunger, in most mammals, while at the same time we would not expect the neural machinery realizing a pain state in a rat to be precisely the same as in humans. Even with regard to individuals from the same culture Nussbaum argues, "if we said that grief is always of necessity accompanied by the firing of so-and-so many neurons of such and such type, we would be likely to find hundreds of cases for which this just isn't quite right" (2001, 59). Nussbaum here takes the central nervous system to be part of the bodily realization of an emotion, which in a broader sense is, of course, correct. But the Jamesian hypothesis in particular takes emotions to be perceptions of arousal in the peripheral nervous system. So the question is whether we find distinct patterns of bodily arousal there that can be perceived by certain parts of the brain (James thought these to be subcortical sensorimotor areas) and not whether the neurons in the brain form distinct patterns that can be identified.[7]

But Nussbaum's argument can be turned into an argument against James: it is not only the brain that is plastic but also the whole nervous system. Therefore, one can have doubts about the possibilities of finding exactly the

same reactions, for example, in the endocrine system, in the same person over several years. And one can have doubts about whether we find the same reactions in the endocrine system when observing fear in a female baboon compared with an old man. Put this way, it becomes easy to see what position Nussbaum is denying: she does not want to commit herself to the claim that there is a single pattern of neuronal activity or bodily arousal that is activated whenever somebody experiences a certain type of emotion. Instead, the functionalist idea Nussbaum refers to claims that a fear state can be realized in various ways and that emotions should therefore be type-identified with regard to the mental evaluations they involve.

Yet what remains unclear is what explanatory role is left for the body in such a functionalist account. On one hand, Nussbaum claims, "all human experiences are embodied, and thus realized in some kind of material process. In that sense, human emotions are all bodily processes" (2001, 58). So Nussbaum seems to commit herself to some version of materialism, as most functionalists do. But on the other hand, Nussbaum takes it to be a virtue of her definition that it includes bodiless beings such as gods and angels: "Whether we believe that bodiless substances exist or not, the reason it makes sense to imagine a bodiless substance having genuine emotions is that it makes sense to imagine that a thinking being, whether realized in matter or not, could care deeply about something in the world, and have thoughts and intentions associated with such attachments. And that's all we really require for emotion" (2001, 60). We have already seen that not even James would deny the mere conceptual possibility of such bodiless emotional creatures. Yet for Nussbaum, the reference to them is meant to support her view that emotions should be type-identified with regard to the evaluation that emotions involve.

One may find it confusing, though, that Nussbaum refers to the functionalist argument of multiple realizability, but at the same time claims that thinking beings do not need to be realized in matter at all. Most functionalist theories that argue for multiple realizability do not deny the important causal role of matter in the making of mental states. They simply claim that several different material arrangements can fulfill the same causal roles. Multiple realizability is not at all equivalent to the complete arbitrariness of the realizing arrangement of matter. Yet what Nussbaum seems to have in mind is a position that claims that emotions are mental evaluations that can be realized in all kinds of ways. Coincidentally, in all animals, including humans, emotions are realized in their brains and bodies. But the ways in which they are realized vary enormously so that it does not make sense to refer to them when looking for features to classify emotions.

This appears to be a radical version of what Lawrence Shapiro (2004) calls the "separability thesis." According to the separability thesis, it is perfectly possible for a humanlike mind to exist in a nonhumanlike body. Shapiro rejects this view, arguing that, for example, visual perception in humans depends fundamentally on the way we move our heads to gain information

about the distance of objects. Shapiro makes similar points with regard to the other sensory modalities and concludes that "a description of various perceptual capacities cannot maintain body-neutrality and . . . an organism with a nonhuman body will have non-visual and auditory psychologies" (2004, 190). What Shapiro means by body-neutrality is precisely what Nussbaum is arguing for: the view that the mind is a program that can be realized in several kinds of bodies and can be characterized in abstraction from the body. Like Shapiro, I think that the ways in which the body contributes to our special ways of perceiving the world have to be taken into account. There are a variety of ways in which this idea can be explained, by giving more or less weight to the precise role of the autonomic nervous system, several brain structures, and cognitive processes. But for any kind of naturalist theory, functionalist or not, Nussbaum's coincidental by-product view is a slightly too arbitrary description of the role of bodily reactions in emotions.

Furthermore, the evidence discussed so far clearly shows that the bodily reactions are far from arbitrary. Consider Nussbaum's example of rage again. Nussbaum reports that while many men say that they associate rage with a boiling feeling, she associates rage with a tense feeling around the neck and sometimes a headache occurring the next day. With regard to the studies discussed earlier, there might be several reasons for these differences: gender reasons might be important, but there is also evidence that people have different abilities when it comes to sensing their own heart rate (Barrett et al. 2004), and the same might prove true for sensing one's own body temperature. Yet tensed muscles, increased heart rate, and elevated body temperature all fit into the prototypical pattern of bodily arousal that is generally associated with rage. What would be disturbing is if Nussbaum reported feeling her heartbeat slow down whenever she is angry, or if she associated it with a relaxation of her muscles from the corrugator all the way down to the belly and the toes. My guess is that in such a case we would doubt whether the feeling that Nussbaum describes could be labeled as a case of anger at all. This intuition, together with the evidence discussed above, should suffice to disqualify an approach that denies any significant functional role of the bodily changes involved in emotions altogether.

Still, Nussbaum is making an important point: if James's claim, that there are specific patterns of bodily arousal in humans that can be used to type-identify emotions, is taken literally, it seems hard to reconcile with the functionalist idea of multiple realizability. Yet the idea that all fear states in all mammals under all circumstances should involve exactly the same pattern of arousal sounds unlikely. We need a better definition of what these patterns are and what it means to reidentify them in various organisms in various situations. The Jamesian question needs to be situated in a broader theoretical framework in order to be answered. In the following, I examine approaches that take emotions to have a common origin and, therefore, an underlying causal mechanism or function that allow us to type-identify them. The Jamesian question is thereby merged with the Darwinian question.

4. COMMON ORIGINS OF AROUSAL PATTERNS

Thus far, we have seen that there is solid evidence for recurrent patterns of bodily arousal in different emotion types. Yet the main reason that the evidence is still somewhat ambivalent is that bodily arousal tends to vary strongly among persons and situations. The theoretical task arising from this is to account for both the fact that bodily reactions in emotions are not random, or mere side effects, and the fact that emotions are multiply realizable. James's essential-ingredient view is too narrow to account for multiple realizability. What I want to argue for in the following is that the Jamesian question cannot be answered outside of a Darwinian framework, which refers to the common origin of the bodily reactions that belong to a particular emotion type.

A current theory that makes a version of the Darwinian claim is the "affect program theory," an approach that gives a precise definition of how emotions can be type-identified with regard to an evolved underlying causal mechanism that triggers the involved bodily and behavioral reactions. The affect program theory has been most notably defended by Paul Griffiths (1997). Griffiths's main source of evidence is Ekman's work on facial expressions, introduced earlier. The broader theoretical framework Ekman developed suggests that the facial expressions he discovered are components of basic emotions. Basic emotions are evolutionarily acquired reactions caused by automatic neural appraisal systems, which then trigger a complex pattern of reactions including facial expressions; musculoskeletal responses, such as flinching; vocal changes, such as a tremulous voice; endocrine system changes and consequent changes in the level of hormones, such as the release of adrenaline or cortisol; autonomic nervous system reactions such as heart rate or body temperature changes; and feelings (Ekman 2003).

Like James's essential-ingredient view, the affect program theory depends on empirical evidence supporting the claim that basic emotions involve specific patterns of bodily arousal. But Griffiths and Ekman abandon the focus on visceral reactions and explicitly list a broad set of bodily and behavioral reactions that are involved in an emotion. More important, Ekman and Griffiths do not claim that emotions *are* the feeling of bodily changes; instead, they argue that the feeling caused by the ongoing bodily arousal is only one aspect of the emotional process and not even a necessary one, since emotions can remain unconscious as well. As I mentioned before, this is a claim that is widely agreed on in the empirical sciences.

Finally, the affect program theory is not committed to James's somewhat vague claim that emotional arousal is directly elicited by perceptions. Instead, the affect program theory distinguishes between the input and the output side of an emotion. While the crude neural evaluation of the stimuli constitutes the input-side of an emotion, the listed pattern of reactions forms their output side. With regard to the input side, Ekman and Griffiths both claim that unconditioned stimuli such as loud noises or loss of balance can trigger emotions as can an infinite range of complex cognitive considerations.

Griffiths argues that only the six basic emotions that Ekman describes can be explained in terms of the affect program theory. These emotions are modular[8] mechanisms that are evolutionarily old and can be activated by unconditioned stimuli without complex cognitive evaluations. Fear reactions in newborns, for example, can be triggered by a loud noise or the loss of balance, while prolonged restraint triggers rage and gentle skin stimulation triggers pleasure (Watson 1930/1970). Although basic emotions can be triggered by an infinite variety of stimuli later in life, and are therefore not domain-specific in the strict sense, the fact that emotional learning mechanisms are not general-purpose mechanisms either might suffice for the modularity thesis. Affect programs also meet the innateness criterion and seem to be realized in dedicated neural circuitry: as far as we know they are subserved by neural circuits in the limbic system and the brain stem which form the phylogenetically old portion of the cortex, the amygdala, and the hypothalamus (Damasio 1994, 1999, 2010). Furthermore, affect programs are triggered and unfolded automatically; that is, people often get disgusted or afraid whether they want to or not. Emotional processing is largely inaccessible to higher cognitive processes. People are aware of the outputs, that is, the emotional responses themselves, but not of the processes producing them. Most important, emotions show a high degree of cognitive impenetrability; that is, they are informationally encapsulated: fear of flying does not vanish when you remind yourself that you wholeheartedly believe flying not to be dangerous. Many responses to early-learned stimuli like phobic reactions to spiders or dogs seem to be irreversible irrespective of whether we think of these reactions as warranted or not. Informational encapsulation is generally seen as the most important criterion for a modular system, since it captures the idea of a system that is separated from general reasoning processes.

Those emotional reactions that can be classified as modular affect programs, according to Griffiths, form a natural kind. They are evolutionarily old mechanisms, hardwired in all humans in a similar way. But Griffiths does not conclude from this that there is a set of necessary or essential elements, like a set of bodily and behavioral reactions, which, when present at once, constitute an emotion. Rather than sharing a certain amount of intrinsic properties, what all fear states have in common is that they are triggered and organized by an affect program that has its origin in evolution.

To make this point, Griffiths endorses Boyd's (1989, 1991) definition of natural kinds. Boyd has developed a more liberal definition of natural kinds that suits biological categories and not just chemical ones. For a chemical kind, it might be a satisfying definition to say that each member of the kind must have some underlying microstructural property that explains all its other projectable properties. The kindhood of biological taxa is very different, however, since biological taxa are historical entities. They are groups of organisms held together by their common history rather than any intrinsic resemblance. Boyd's suggestion, therefore, is that natural kinds are

homeostatic property clusters. A category brings together a set of objects with correlated properties. The category has causal homeostasis if this set of correlations has some underlying explanation that makes it projectable. Instead of simply noting the existence of a cluster of properties in a given object, the order in which these properties occur is explained by an underlying causal mechanism. Homeostatic property clusters need not have a common essence. It is not necessary for all the members of a species to share the same genetic code and neither is it necessary for all instantiations of a certain emotion to include a certain set of bodily and behavioral reactions. They only need to share similar properties bound by a causal mechanism with a common origin. If a state of fear does not include the usual acceleration of the heart rate, because the subject in question is running or angry and therefore the heart is already beating relatively fast, this does not call into question the fact that she is undergoing a state of fear, since the same affect program was triggered, although one of its features could not be realized in the normal way. Therefore, the notion of causal homeostasis offers a good explanation for the patterns of bodily arousal in emotion: they do not have to be exactly the same each time they occur, yet they are caused and organized by the same causal mechanism that has its origin in evolution.[9]

This description differs from the Jamesian essential-ingredient view and can answer Nussbaum's concern that we should not expect emotions to be realized in exactly the same way in us and all other mammals. It identifies affect programs with the causal mechanisms underlying emotional reactions in primates and, more broadly, in other vertebrates. The claim, however, is not that our ancestors all have the same neural circuits and nervous systems as we do, but rather that our nervous system shares a common origin with theirs. The objects the theory describes, like the affect program for fear, are products of a particular sequence of evolutionary events. The affect program theory assumes that human emotions will resemble those of animals roughly in proportion to the degree of relationship between humans and other species.

Different instantiations of affect programs, therefore, are homologies of one another in the Darwinian sense. Homologies are traits that organisms share due to their common descent from organisms with similar traits; analogies are traits that share a common function. Homologies can occur across different species and functional categories as long as there is a continuous line of decent. Analogies are the result of parallel evolution and can occur across different lines of decent. While the wings of birds and bats share the same broad function—they allow for flying—they do not share the same descent. They are therefore analogues of one another but not homologues. Bat wings and human hands on the contrary do not share the same function, but they have evolved from a common descent and are therefore homologues of one another.

There are two approaches to classify psychological traits in biology that focus on two different levels of explanation: the *cladistic* approach, which

focuses on the level of material realization or computation of a trait, and the *ecological* approach, which focuses on the functional level and on the way in which the organism interacts with its environment. The cladistic approach, according to Griffiths, elicits more projectable categories. The ecological approach remains at the task-description level since the causal homeostatic mechanism of each ecological category is a set of adaptive forces, not a set of bodily reactions and behaviors. The idea is that disgust, an ecological category, can be used for explanation and induction because natural selection has imposed the same requirements on all responses that were designed to prevent the ingestion of noxious things. But computational and neurological knowledge about disgust in rats cannot be extrapolated to Martians or octopuses, in fact not even to birds (Griffiths 1997, 234).

Griffiths argues that research on the actual realization structure of a biological trait, which includes the study of its decent, that is, its homologies in other species, is more telling than the search for functional categories across species. Furthermore, well-formed functional categories depend on information from the cladistic level:

> The functional-ecological categories found in progressive research programs in biology are typically historically constrained, and sound functional explanations are typically also historical ones. This suggests . . . that any features of emotion which *are* best explained in terms of adaptive convergence should be studied as an extension of the natural historical approach which treats emotions as evolutionary homologies.
>
> (Griffiths 1997, 236)

This leaves Griffiths with the result that several emotions are affect programs and can be understood as a natural kind, in the sense that they share underlying causal mechanisms. However, according to Griffiths, there are other emotions, namely, higher cognitive emotions, for which current research does not yet have a method for coming up with projectable categories, since they cannot be classified as affect programs and ecological psychology is not able to produce projectable categories. This account can be challenged for several reasons. I first consider objections to the neat distinction between basic and higher cognitive emotions, then discuss Griffiths's critique of the notion of function and ecological categories, and finally come up with a suggestion of how to capture the embodied dimension of emotions and classify them in functionalist terms.

5. AGAINST BASIC EMOTIONS

The neat definition of affect programs as a natural kind appears plausible and well founded on empirical data at first sight, yet it does not fit well as a model for all emotional phenomena. Affect programs are brief, highly

stereotyped reactions; they are not involved in higher cognitive processes and their behavioral output unfolds in a largely autonomic way. However, romantic love might last for years and is not a quick and automatic reaction; pride about an achieved goal might occur in the absence of a particular pattern of bodily arousal; jealousy or shame seem to presuppose higher cognitive abilities, like an explicit understanding of social relations, rules, and norms. It is less clear whether these emotions are encapsulated (i.e., separated from general reasoning processes), and it is likely that they can be type-identified only with reference to the social context in which they occur. Not only do certain types of emotions like jealousy, guilt, or pride not fit with the affect program theory, there are also instances of emotions such as fear or disgust that do not involve affect programs, like constant stress before an exam, which is triggered by internal reasoning and lasts for weeks, or moral disgust in response to an article on racist violence, which involves complex cognitive processing and is not necessarily accompanied by autonomic reactions or facial expressions.

Griffiths concludes that while basic emotions form a natural kind that was already present in our ancestors, higher cognitive emotions should be seen as a distinct class of mental states that rely on certain cognitive abilities unique to humans and are underlain by a different causal mechanism. The distinction between higher cognitive and basic emotions crosses folk categories insofar as not all things that we label as fear are identical to an affect program. For example, a calm state of disapproval might be labeled as anger according to the folk concept but lack the pattern of arousal that would allow us to classify it as an instantiation of an affect program.

This so-called disunity thesis has been criticized by many authors (e.g., Prinz 2004; Zinck and Newen 2008; Colombetti 2014). While most of these authors criticize Griffiths on the basis of their own theories and develop other ways to classify emotions, recent studies suggest that there might be basic forms or homologies for far more emotions than Griffiths considered. Jason Clark (2010) cogently argues that shame has a basic form and that its higher cognitive form has deep commonalities with its basic form. Shame appears to have emerged from a rank-related emotion in nonhuman mammals, in which its primary function was to signal subordinance to a dominant in order to appease her. This emotion can be elicited simply by being in the presence of a more highly ranked individual without presupposing a theory of mind (Fessler 2004). Although shame does not seem to have a pancultural facial expression, it does have characteristic expressive and behavioral features, which include shrinking in posture, gaze aversion, flight, bent-knee gait, avoiding social contact, and hiding (Fessler 2004). Shame responses also involve patterned physiological activity in the hypothalamic–pituitary–adrenal axis (HPA axis), known, in part, to regulate stress and proinflammatory immune system responses serving basic physiological protective and reparative functions (Dickerson and Kemeny 2004). Activation of the HPA axis associated with submission behavior has been observed in a wide range

of animals, among them baboons, rats, fish, and humans. Subordinate animals tend to have higher levels of basal corticosteroids than their dominant peers, as well as enhanced proinflammatory immune activity. With these data in mind, there is little reason to think of shame as a human-specific higher cognitive reaction without any roots in the evolutionary history of other species. Shame in human adults might include further cognitive abilities and might have changed or extended its function, but there is no reason for the radical distinction drawn by Griffiths. This evidence also supports the claim that all emotions, not simply basic emotions, possess stereotypical physiological, behavioral, and expressive characteristics: "The view that Higher Cognitive Emotions do not have any distinct neural, physiological, expressive, or behavioral characteristics is widely held, but poorly supported" (Clark 2010, 81). Not only is the disunity thesis poorly supported, as I argue in the following; a strong case can also be made for the unity thesis that emotions do belong to the same natural kind.

Further support for Clark's thesis comes from recent psychological studies on pride: Jessica Tracy and Richard Robbins (2004, 2007a, 2007b; Tracy et al. 2013) argue that pride is cross-culturally associated with a distinct, recognizable nonverbal expression. The most prototypical pride expression includes an expanded body posture, a slightly backward-tilted head, arms akimbo with hands on the hips, and a small smile on the face. Not all components are necessary, but at least one prototypical element other than the small smile must be present so that pride can be successfully distinguished from other pleasant emotions such as happiness. It is likely that evidence for other stereotypical expressions of so-called higher cognitive emotions will be found once research is extended from its narrow focus on facial muscle movements to position of the head and body posture. While pride seems to include an expanded posture, the shame expression seems to include a downward head tilt and a shrunken posture. Fessler (2004) even views shame as the antithesis of the pride display: the core components of these two displays (direction of gaze, erection of posture, and gait) are components of displays employed by nonhuman primates (and other mammals) during the negotiation or affirmation of relative rank.

The neat distinction between basic emotions, which are evolutionarily acquired affect programs, and higher cognitive emotions, which do not have a unique pattern of bodily reactions but presuppose higher cognitive abilities, is, to say the least, much too strict. Shame is a prime example of a higher cognitive emotion. Most people think of it as a uniquely human emotion that develops relatively late since it presupposes complex cognitive abilities, such as a theory of mind. Yet if shame has a cross-culturally occurring expression, is associated with certain reactions of the endocrine-system that already occur in many animals, and so on, it seems reasonable to think of shame as having an evolutionary basis that only develops further and changes or broadens its function in human life. The same is probably true for pride and, since research in this field is still very young, we can expect similar evidence for many, if not all, other so-called higher cognitive emotions.

This calls for a reconsideration of the concept of "basic emotions" that Ekman and Griffiths defend, since what remains to distinguish emotions such as fear and disgust from emotions such as pride and shame are only very vague and gradual features. This becomes clear when reexamining Griffiths's points about the modularity of affect programs, on one hand, and the higher cognitive capacities involved in emotions such as shame and pride, on the other hand. Emotions such as fear and disgust are supposed to be domain-specific, in the modest sense that they have some unconditioned stimuli but can be triggered by all kinds of complex cognitive thoughts later on. Yet if we think of the most basic forms of shame in animals, it seems that they can be triggered by the perception of the animals' status in relation to others and that this perception does not require anything like a theory of mind and is instead entirely hardwired. That shame can be triggered by all kinds of reasoning about social rules and norms in human adults is analogous to the fact that fear in human adults can be triggered by cognitively complex processes such as an imagined job loss or a glance at one's bank account. It is also considerably too early to conclude that emotions such as shame and pride are not realized by dedicated neural circuitry—recent studies have produced some evidence suggesting that shame and other so-called self-conscious emotions show typical patterns in neuroimaging studies (Beer 2007).[10] It is certainly true for emotions such as shame and pride, too, that they usually unfold automatically, that is, that we often feel proud or guilty whether we want to or not. Typical instances of higher cognitive emotions are also very often cognitively impenetrable: guilt does not always vanish just because we tell ourselves that we have not done anything wrong and we often feel proud when somebody compliments us, even if we think that something like "You look pretty good today!" is actually nothing to be particularly proud of.

One remaining distinction that Griffiths mentions is the claim that higher cognitive emotions can last for weeks, months, and years, while affect programs are brief, automatic, and coordinated reactions. Yet, as Clark has shown, this distinction is certainly not a strict dichotomy. If we think of the physiological arousal involved in emotions only in terms of spikes in skin conductance or increased heartbeat, then it certainly makes sense to see them as brief reactions, since these are not the kinds of things that last for weeks. But alterations in the thresholds or cyclical patterns of such transient patterns can last for longer periods. We can be "on the verge of tears" for hours, maybe days, with only short bursts of actual crying; we can be permanently stressed, nervous, and "tense" in the weeks before an exam. There are physiological elements that can persist over long periods, such as alterations in hormonal or neurotransmitter levels (Clark 2010, 82). Again, what follows from these arguments is that it does not make much sense to draw a categorical distinction between basic emotions and higher cognitive emotions because of their differences in duration.

To sum up, the evidence for a categorical distinction between basic and higher cognitive emotions is pretty thin.[11] A methodological problem with

the postulation of basic emotions could be that there is a tendency in the research literature to postulate a fixed set of basic emotions extrapolated from one's own research in a single field. This is exactly what Ekman does in postulating that there are six affect programs because this fits with his finding that there are six cross-culturally occurring facial expressions. Yet, as we have seen, the inclusion of even just body posture allows us to specify a stable and cross-culturally occurring expression for pride that can be distinguished from similar emotions such as joy. In a similar fashion, Jaak Panksepp (1998) suggests that there are four basic emotions (fear, play, care, anger) constituted by four different emotional systems in the hypothalamus. However, while it might be true that these are the main systems in the hypothalamus, this does not mean that the structure of the hypothalamus is the only basis for distinguishing emotion types. Different emotional elements could, on the contrary, trigger these four systems during the unfolding of an emotional episode, yet they could be differentiated along other neural and bodily levels and in the way they interact with the environment (McNaughton 1989, 19f.; Colombetti 2014).

The presence of these discipline-specific blind spots shows that research on basic emotions certainly must be a cross-disciplinary affair in which no discipline claims to have privileged access to the phenomenon. But as soon as we compare the different claims that scientists from different fields have made, the talk of some emotion types being basic while others are derived or higher cognitive becomes problematic. There seem to be evolutionarily acquired elements and learned components in each emotion. The challenging task for a differentiated theory of emotions would be to untangle emotions in order to classify them along a gradual spectrum, where some might be evolutionarily older, or more hardwired, or less sensitive to social context than others.

Yet in spite of the evidence for the similar basic components in fear and shame, one might still claim that the concept of an affect program implies a certain kind of causal mechanism, which is a certain way in which the components that realize an emotion are connected. Griffiths and Ekman both describe affect programs as automatically unfolding coordinated programs. Ekman further claims that affect programs are literally inscribed into our neural circuits, are triggered by a stimulus, and account for the fact that emotions are modular, quick, and automatic responses that are complex and coordinated at the same time (Ekman 2003). If this was the case for some evolutionarily hardwired emotions but not for others, which were instead more plastic and sensitive to social influences during development and more flexible in their unfolding, there would be good reason to place these two kinds of emotions in different categories and label one as more basic than the other. However, there is no evidence for this claim—either for those emotions that Ekman considers to be basic or for any others. The idea of a neural script is hypothetical, more an educated guess as to what the mechanism behind the unfolding of an emotional episode might be. But

there is room for other educated guesses here. An entirely different option is to think of emotional episodes as dynamically organized patterns that are realized by the whole organism. This is the view that I want to develop and argue for in the next section.

6. EMOTIONS AS DYNAMICAL PATTERNS

Dynamic systems theory is well established in physics. Dynamical approaches to cognition go back at least to the early cybernetics literature of the 1940s (see Clark [2014] for an overview). The broadest definition of a dynamic system is a system that changes over time. In the cognitive science domain, dynamic approaches are especially attractive for understanding complex systems composed of many continuously interacting parts. They can also capture the aspects of adaptive behavior that depend on complex, circular causal exchanges. A famous example is the explanation of how we learn to walk (Thelen and Smith 1994). Traditional theories assume a genetically specified central set of instructions that is supposed to be neuronally realized and comes complete with timing. Thelen and Smith (1994) argue that there is no such complete and prespecified neural control system. Learning to walk, they argue, involves a complex set of interactions among neural states, the springlike properties of the leg muscles, and the local environment. The ability to walk is the result of a complex and balanced interplay of multiple factors spanning the brain, the body, and the world.

In a similar vein I suggest to abandon the idea that emotions are centrally governed by one program that is triggered by one stimulus. Instead, there might be a number of neural, bodily, and environmental elements organized in parallel to each other that respond to various features of the stimulus situation. These different elements could synchronize to one standard form of an emotional episode through a process of entrainment. In fear, one trigger could stimulate the autonomic system while another could independently stimulate the musculoskeletal system, and the two reactions could get "used to each other," such that the presence of one of them easily triggers the other. What allows us to see the separate effector systems as a unified response of the organism is an ecological analysis that ascribes a function to the episode as a whole.

However, there could still be an inner connection between emotional components. Ekman et al. (1983) found that facial expressions can produce autonomic changes. There is a similar link between body posture and autonomic arousal and a direct link between autonomic arousal and experience that also has feedback effects on cognitive attention (McNaughton 1989). Giovanna Colombetti (2014, 53ff.) argues that there is good evidence for such dynamical self-organization in emotions at least in (1) facial expressions, where we find coordinated muscle-structures that mutually constrain and trigger each other; (2) neural systems underpinning the

microorganization of emotional episodes; and (3) interpersonal patterns of behavior that get established in the early infant–environment interaction.

In addition to the evidence supporting the dynamical organization of these emotional elements, there might be feedback loops where facial expression modifies autonomic reactions, autonomic reactions modify cognitive assessment, and cognitive assessment modifies facial expression. In such feedback loops, the unfolding of emotional episodes can be explained as an effect of the emotions being adapted to the environment and of the components being adapted to each other such that they can stimulate each other without there being a fixed sequence or a governing central state. Instead, emotional processing at each step of the interaction can produce different results of emotional endurance and intensity and can involve more or fewer components. The unfolding of the emotional sequence does not depend on central processing, nor does an affect program govern it; rather, it is the result of a well-adapted dialectical interaction of the various inner components with each other and with the environment.

The usual co-occurrence of the components alone is no justification for treating them as linked. Furthermore, evidence for some correlations between some components of emotions does not imply that there are correlations between all components of all emotions. A central state triggering an internal emotional program is a possible but not a necessary explanation for the unfolding of emotional episodes. On a pragmatic level, where we ask which theoretical assumptions should guide future research, Neil McNaughton suggests that until centralization is demonstrated experimentally, we should simply assume that "the interaction between different components of emotion . . . could provide emotional phenomena with sufficient coherence and integration that even without an obvious central control state we would want to see individual emotions as separate functional entities" (McNaughton 1989, 141).

There is, however, an argument in favor of the dynamic approach: it is commonplace that evolution does not result in well-designed, optimal engineering solutions but is more adequately described as a permanent process of tinkering. Evolution is conservative insofar as completely new features that result from mutation are very rare. The more complex a feature, the more unlikely it is that it resulted from one single mutation. Evolution works by incremental changes that offer small increases of fitness to the organism. The hormone oxytocin, for example, changed or extended its functions several times, as it predates mammals. In terrestrial animals it is involved in the regulation of water and minerals; in female mammals it is upregulated during pregnancy and triggers the letdown reflex later on in lactating mothers, but it is also involved in the motivation of maternal caring behavior (Porges and Carter 2011). Emotions are incrementally developed assemblies. They are complex dynamic patterns that include various elements, such as expressions, autonomic reactions, hormonal changes, and cognitive appraisals, some of which are evolutionarily older than others and most of which had previously independent functions. The evolution of

"neural scripts" of the sort that Ekman seems to have in mind—scripts that centrally govern the unfolding of an emotional process—are not excluded by such considerations, but they appear less likely. The dynamical approach fits into the broad framework of evolutionary development better because it describes the development of affect programs in a way that relies on the slow and permanent adaptation of an organism to its environment. An animal's autonomic nervous system responses and its facial expressions in a dangerous situation might have developed for different reasons and at different times. The development of a neural program like the one Ekman describes is certainly possible. But the evolution of parallel systems that are well coordinated when they occur in a structured environment is much more likely. Such parallel organization can also better explain why bodily arousal shows not only certain consistencies but also so much variation within a single organism and among various organisms. Adaptations to a certain niche as well as the current situation can differ, as can the various components that make up the normal affect program.

Although aware of its advantages, Griffiths rejects the dynamical approach. According to him, it is less convincing for highly flexible organisms in complex environments such as humans. Humans permanently cope with novel stimulus situations. Therefore, "the only thing which all eliciting situations for fear have in common is the extremely abstract property that, in the light of the organism's past learning history, they can be evaluated as dangerous" (Griffiths 1997, 86). Griffiths assumes that emotions in humans require the cognitive evaluation of the stimulus. Because the stimuli triggering the same reaction are different, the evaluation classifies stimuli as "dangerous" or "offensive" in the first place. I argue against this assumption in Chapters 4 and 5 and thereby defend the dynamical view. The dynamical view is often used not only to discard a neural program that governs a response but also as a replacement for cognitive explanations that assume complex inner representations. As I argue against Colombetti in Chapter 3 it is not (yet) possible to give a convincing account of emotions as dynamical patterns that makes the assumption that emotions are representations that stand in relation to other representations superfluous.

Griffiths has another obvious reason for disagreeing with the dynamic approach that he does not make explicit: what allows for a type identification of emotions in the absence of a neural program or cognitive appraisal is the presupposition of a function that the emotion fulfills for the organism. As I have shown, Griffiths thinks that affect programs should be identified by research on the level of computation and implementation: only research on homological traits on these levels could produce projectable categories. Ecological analysis, which operates on the task-description level and suggests adaptive functions for different emotion types, is not able to produce comparably stable results, or so Griffiths believes. In the following section I argue that homological categories cannot be established without a notion of function that helps to define them, in opposition to Griffiths's position.

7. HOMOLOGIES AND FUNCTIONS

Ecological categories are more closely related to our vernacular usage of emotion terms than the ones brought forward by the affect program theory. Griffiths, metaphysical revisionist that he is, does not see this as an advantage but instead comes up with three reasons why ecological categories cannot be projectable:

1. They do not have the capacity to explain facts about emotions at the computational or implementation level. They can explain facts only at the level of task description.
2. There are no ecological categories that successfully categorize emotions with regard to their adaptive functions and there are reasons to believe that they will never exist.
3. Current work in this field proceeds by taking folk concepts and assuming that they somehow correspond to categories in cognitive ecology. There are no grounds for reasoning of this sort.

Griffiths certainly makes an important point here: reasoning about the adaptive functions of certain traits should be based on the actual evidence for the historical development of these traits. Classical functionalist approaches have the tendency to think about cognitive abilities, perceptions, and feelings as if there were a realm of functional categories and various instantiations of these categories realized by all kinds of material arrangements. Also, I would agree with Griffiths if his critique were mainly directed against the tendency in evolutionary psychology to come up with hypotheses about the adaptivity of certain cognitive mechanisms that are extremely difficult to test empirically.

However, I think that Griffiths is mistaken in his view that homological categories can be constructed without first considering the task description-level and reasoning about adaptive forces and resulting functions of the items that a category aims to capture. There are several arguments for this claim, the first of which is a methodological one. Millikan claims that we need a notion of function to classify various activities as kinds of behaviors. She argues that without the notion of function, an animal's activities can be described in a potentially indefinite number of ways:

> Should we explain why Amos's eyes blinked just before a piece of dust struck his closed eyelids, when the clock said 2:37:08, just as Amos's whiskers twitched, or just as the end of Amos's tail passed the fifth blue square of the kitchen linoleum? Indeed, did Amos blink or was it just that his upper eyelashes removed themselves, in an arc, away from his eyebrows or moved to point at his navel or his nose or his toes?
>
> (Millikan 1993, 142)

The eye-blink reflex is properly described as a blink, or closing of the eyes, for only as such does it have a biological function. The notion of function is necessary to distinguish behaviors from all other dispositions that animals might have. While a chameleon has a disposition to turn brown when placed in a brown box, a mouse has the disposition to brown nicely when placed in an oven at 350 degrees Fahrenheit (Millikan 1993, 146). There are many lawlike dispositions, like cockroaches becoming torpid when the temperature drops too low or a blow below the kneecap causing a kick. However, most of these dispositions should not be categorized as behaviors that serve a certain purpose for the animal. They might be "spandrels," that is, results of the system's architecture that are accidental relative to its functional design. It is the ethologist's job to come up with categories of behavior, and there is no possible way of describing types of behavior or picking them out of the world without applying a notion of function. Millikan's argument not only applies to behavior but to all kinds of traits and to complex patterns of bodily and behavioral reactions that are involved in emotions as well. If we did not apply the notion of function to the features and reactions we are investigating, we would not know how to describe them or which things to include in the description. While this argument appears to imply that functional reasoning has its place merely in preempirical considerations on how to capture one's subject and plays no role in the actual formulation of projectable categories, that is, on the ontological level, there is a further argument closely connected to the first one, concerning the notion of normality that clearly has ontological implications.

Millikan (1993) and Neander (2002) argue that we need the notion of function in biological categorization in order to come up with normative categories that explain what an item was selected for, that is, what it is supposed to do under normal conditions. Such a notion of normal conditions enables us to distinguish between well- and malformed members of a category: a heart that does not pump blood still belongs to the biological category of hearts, since it shares a common decent with other items that have been selected for the same purpose. Yet it is a malformed member of the category since it cannot fulfill the function it was selected for. With regard to behavior, Millikan argues that

> [t]he very subject of behavioral study, the intact animal, is defined by reference to proper or normal function. Behavioral dispositions are dispositions not just of any old chunk of warm matter but of a chunk having a normal constitution, where this is defined relative to its (historically defined) proper functions . . . Most of the dispositions of Amos (as a chunk of matter, R.H.) are chemical and physical, not psychological. To find the psychological ones, we must make a necessary reference to the functions of Amos's . . . dispositions.
>
> (Millikan 1993, 146)

This is not merely a methodological or epistemological claim about the usefulness of the concept of function in structuring a theory; it is a claim about what organisms are, how they develop, and how they are structured. For Millikan, it is clearly an ontological claim that organisms can be in a normal condition and that organs, mechanisms, and behaviors form functional kinds with well-functioning and malfunctioning members. Applying this kind of reasoning to the affect program approach, one might wonder how we can identify the causal homeostatic mechanism underlying each affect program and how we determine which components belong to it without involving any kind of functional reasoning. One might also wonder how to distinguish between the normal functioning of that mechanism and abnormalities (e.g., where an organism is undergoing a state of fear but cannot be adequately prepared for flight because autonomic arousal is inhibited for some reason). If this argument is correct, there is no need to prove, contra Griffiths, that ecological approaches can come up with *additional* plausible categories, since functional reasoning is *already* part and parcel of research accounts that focus mainly on the implementation level.

Yet Griffiths argues against Neander and Millikan that functional kinds only derive their ability to type-identify well-functioning and damaged or diseased items from their logical relations to cladistic kinds:

> Biological functions attach to items in virtue of their evolutionary history. So in order to ascribe a function to an item, it is necessary to establish that it is a member of lineage with a unique evolutionary history, or, in other words, a clade. Functional kinds are therefore either coextensive with cladistic kinds or with disjunctions of cladistic kinds.
>
> (Griffiths 1997, 216)

I think that Griffiths's argument is flawed. What is right is that an item that developed through a mutation has no biological function in the first place. Only if they are useful for the organism, and therefore are reproduced, can items of following generations be said to have a function since functions are established by means of selection. Therefore, a first instantiation of a feature with certain causal powers must exist before following generations of the same item can be said to perform a function when exhibiting the same causal powers. So if there is any such thing as a "logical priority," it concerns an item with certain causal powers being there before it can be said to have a function. A *first member* of a clade must exist before there can be a first member of a functional category. Yet it is nonsense to say that *a clade* exists from the moment the first item exits, without any selectively established function. A cladistic category can only be established with reference to a history of reproduction, and this historical process necessarily involves adaptive and selective processes in which functions and normal conditions for the members of the category become established as well. What Griffiths implies is that there are cladistic categories whose members, on the implementation

level, show certain causal powers due to their common descent. Then, on the task description level, one can guess about the functions of the traits or behaviors in question and, if one is right, discover a cladistic category or a disjunction of several cladistic categories. But Griffiths bases this argument on the assumption of a "logical priority" of homological categories, which, as far as I can see, implies that historical selective processes establish categories with members possessing certain causal powers, without already being subject to selective processes that determine reproductive fitness and establish functions.

Looking back to what I said about different levels of explanation in the introduction, one could say that an explanation of affect programs as incrementally evolved mechanisms can hardly do without a notion of function that allows a description of the intact animal on the task description level. The task-description level is an explanatory tool that could, in principle, be reduced to the implementation level. Instead, we need the notion of function in our theory because it accounts for how the categories we are looking for were established in a selective process. The research on homologies alone is just as much a hunt for a chimera as the search for pure functional categories, which is restricted to the task-description level alone. The idea that we can search for psychological categories without observing on the task-description level what certain mechanisms do and what functions these activities might have, appears to be just as blind as pure functional categories are empty.

8. ASSEMBLING THE PIECES

The starting point for this chapter was the Jamesian hypothesis that emotions—contrary to cognitivist claims—can be type-identified with regard to the bodily reactions they involve. Psychological research after James has shifted the focus from the mental feeling of visceral arousal to the whole psychophysiological reaction, asking whether autonomous arousal and hormonal changes, together with facial and bodily expressions, form a complex pattern that can be reidentified across different species, individuals, and situations. I have argued that while there is clear evidence for prototypical features of emotions—like a facial expression, a certain kind of autonomic arousal, and tensed muscles in anger, or a body posture and typical hormonal changes in shame—these cannot be described as a set of necessary features or essential ingredients, since they tend to vary over situations and between different organisms. There is also no evidence for a common neural program that triggers the automatic unfolding of an anger reaction. Since emotions are complex reaction patterns, it seems rather likely that they evolved incrementally and are not governed by one central state but rather are a set of dynamically organized components that are adapted to each other and trigger and deactivate each other during interaction with the

environment. Such a description of emotional reactions unfolding in feed-back loops can hardly do without considering the task-description level and the questions of how the organism is adapted to its environment and what the environment affords it. The normal, or functional, behavior that can be observed on the task-description level is not something that we can simply invent for explanatory reasons. We need an account of functions to figure out biological and psychological categories because these categories have been established through selection processes.

Given all this, a positive account keeps James's idea that emotions are "nervous anticipations" by claiming that the bodily reactions involved in emotions constitute action-oriented representations with the function to prepare the organism for action. This combines the insights from the Jamesian and the Darwinian perspectives. Furthermore, the claim that the patterned bodily reactions involved in emotions serve a function, namely, to prepare the organism for a certain kind of action, is a first hint of how to answer the question from embodied cognition. As I argued in the introduction, interesting cases of embodied cognition are those where bodily reactions or sensorimotor skills enable the organism to interact in an intelligent way with the environment. With regard to perception, O'Regan and Noë (2001) develop a "sensorimotor enactive view." O'Regan and Noë claim that perception fundamentally involves sensorimotor knowledge of the effects of movement on sensory stimulation. In the following chapters, I argue that the bodily reactions involved in emotions do indeed work in a similar manner. While there is much to say about how to make such an account work, there is one objection that immediately springs to mind. Perception, according to O'Regan and Noë, is partly constituted by the exercise of certain bodily skills and therefore depends "on the possession of the sorts of bodies that can encompass those skills, for only a creature with such a body could have those skills. To perceive like us, it follows, you must have a body like us" (Noë 2004, 25). As this concerns the intentional object *and* the phenomenological aspect of perception, Noë and O'Regan are committed to an essential-ingredient view with regard to perception: perceptual experience depends on the kinds of skillful bodies we have.

Yet the view that it is our bodies, in their very particular uniqueness, that enable us to experience perceptions or emotions the way we do appears to be a kind of sensorimotor chauvinism (Clark 2008b). The body, or sensory apparatus, certainly makes a special kind of functional contribution to the constitution of emotions; differences in implementation are likely to cause real differences in experience, too. Thus, at this point in the research, functionalism and full sensitivity to the details of embodiment cannot completely be reconciled. Sensorimotor chauvinists make a mistake in moving from the premise, "bodily structures and worldly interventions can be active and crucial participants in cognitive processes," to the conclusion, "bodily structures and worldly interventions must in all cases play a special role

such that sameness of mental state requires sameness of bodily structure" (Clark 2008b, 18).

Embodied and embedded approaches can be seen as extending rather than undermining functionalism. Such an extended functionalist approach analyzes cognitive processes as sequences consisting of less intelligent subprocesses. These subprocesses can be described using recognizable computational concepts, but applying those concepts not only to neural hardware but also to various parts of a larger organizational whole that involves the body and the environment. The profound contributions that the body and the environment make to cognitive processes can thereby be recognized, and the abstract roles of bodily and neural operations in real time can become the object of research. "Bodily actions are thus part of the means by which certain computational and representational operations are implemented. But what makes the cognitive process the one that it is, is simply its functional profile" (Clark 2008b, 11).

Adapting Clark's view to a theory of emotions means abandoning the original Jamesian hypothesis that emotions can be type-identified with reference to their bodily profile *alone*. It also means abandoning the affect program view as Griffiths defends it, with its strict distinction between basic and higher cognitive emotions and its critique of functional categories. While I agree with Griffiths that a functionalist approach needs to integrate the historical dimension and pay attention to the actual forms of realization, the extended functionalist approach can accommodate the role of the concrete forms of realization and can easily be extended to capture the historical dimensions of a trait as well. Selectively established functions, after all, are nothing but a subclass of causal role functions. While ascribing a causal role function to a system remains neutral with regard to adaptive values and simply names the effect a mechanism is expected to produce in a given system, to ascribe a selectively established function to an item means identifying the effect of the item on the organism as well, except that it is assumed that the item has gained the causal role function as a result of a reproductive history.

The present approach combines insights from biosemantics with an embodied functionalism. The biosemantic claim is that functions are historically developed and objectively existing features of psychological (and other) categories. The claim from embodied functionalism is that, to understand actual cognizers, that is, living organisms, one must pay attention to the implementation level, which can include the body and the environment of an organism. To individuate a cognitive trait one must look at the trait's functional profile. The two approaches complement each other nicely since they highlight different aspects of the same story: Clark would never deny that most causal role functions that are active in human psychology have been selectively established. On the contrary, the extended functionalist approach understands emotions as incrementally evolved reaction patterns, whose components are closely connected but not centrally governed;

emotions cannot be type-identified with regard to an inner appraisal, a neural program, or a causal mechanism. Emotions should be individuated as whole-organism responses with reference to their functional profile: fear is realized by a complex psychophysiological mechanism that prepares the organism to deal with a dangerous situation. Its function is to help the organism avoid danger. Such a functional state can be realized in several ways, and we should expect it to be realized differently in rats, humans, and robots. Yet in characterizing the functional profile of an emotion as it is realized in humans, the precise role of the bodily arousal involved is a central research issue. Characterizing the functional role of fear, for example, includes finding out how the autonomic nervous system, the musculoskeletal system, facial and vocal expressions, and so on realize fear reactions *and* what other means of implementation would grant the same kind of reaction. Bodily realization matters insofar as it is only certain kinds of mechanisms that can come to fulfill certain kinds of functions.

NOTES

1   In this spirit, Robert Solomon, with regard to his early works from the 1970s, remarks, "In my original theory, it was by no means clear, that the body had any essential role in emotions" (Solomon 2001/2003, 189).`

2   See, for example, Solomon (2001/2003, 26).

3   See Laird and Lacasse (2014) for further arguments.

4   For further arguments against the Schachter–Singer experiment, see Griffiths (1997) and McNaughton (1989).

5   Walter Cannon (1927) also argues that visceral feedback does not suffice to distinguish various kinds of emotions. However, this objection can be met with the same kind of argument that I have already mentioned, that bodily reactions involve much more than just visceral feedback from the spine and the vagus nerve.

6   See also Laird (2007) and Matsumoto et al. (2010) for general overviews of the feedback generated by facial expressions.

7   I have already mentioned that the view I want to defend here departs from the Jamesian hypothesis. But with respect to this particular point, the embodied-emotions hypothesis I am discussing here, just like the Jamesian hypothesis, does not depend on the question of whether there is a particular brain state that realizes an emotion but rather on whether there is a particular set of bodily reactions that does so. However, Nussbaum denies both claims.

8   According to Jerry Fodor's (1983) highly influential definition, several parts of the mind can be seen as modular insofar as they (1) are localized in a fixed neural architecture; (2) show specific breakdown patterns; (3) are mandatory, that is, cannot be controlled voluntarily; (4) operate quickly; (5) produce shallow outputs, that is, nonconceptual representations or simple concepts that do not suppose much background knowledge; (6) are inaccessible for huge parts of higher cognitive processing; (7) are informationally encapsulated or cognitively impenetrable in the sense that the information in question cannot be directly influenced by further beliefs; (8) are highly ontogenetically determined; (9) and domain-specific; that is, they only process a restricted kind of input.

9   It is worth noting that Griffiths switches from grave metaphysical talk about what and how things really are to a rather pragmatic scientific perspective in which natural kinds are no longer conceived of as subjects of the fundamental laws of nature. Instead, they are simply nonarbitrary ways of grouping natural phenomena. These kinds are nonarbitrary (or natural) because they have some degree of projectability (Griffiths 1997, 213).

10   See also Colombetti (2014), who argues that recent studies do not justify a strict distinction between basic and higher cognitive emotions.

11   Clark himself does not suggest dropping the basic–higher cognitive distinction entirely but merely reserving it mainly for differences within emotion types. There is a basic form of shame that is present in other animals and a higher cognitive form with a slightly different function that is unique to humans and which presupposes certain cognitive abilities. I consider this to be mainly a conceptual difference between Clark's approach and mine.

3   Embodiment and the Intentionality of Emotions

1. INTRODUCTION

The result of the last two chapters is that while an embodied functionalist approach can shed light on the central role of bodily reactions in emotion, it confronts the same general objection as the feeling theory developed by James. Since an embodied functionalist approach explains emotions as primarily realized by bodily mechanisms, it is not *prima facie* clear how emotions could be intentional, let alone how they could account for the normative assessability of emotions. This raises the question of how such an approach can include any kind of "aboutness" of emotions such that the bodily reactions involved in emotions play a fundamental role in the constitution of the access to the world that emotions establish.

In this chapter I discuss perceptual approaches to emotions and argue that these accounts do well in pointing out that emotions have a nonconceptual content. Yet the analogy with perception hides some obvious differences between perception and emotion, and I argue that perceptualist approaches do not offer a convincing explanation of the role that bodily arousal, motivation, and valence play in emotions. I then further pursue the idea that emotions have a nonconceptual content and argue from a phenomenological perspective that the emotions' intentionality and the way they feel are best explained by seeing them as constituted by bodily arousal. This gives some *prima facie* evidence for embodied and enactive accounts that I discuss in detail. I argue that all current embodied and enactive approaches fail to account for the normative structure of emotions. This is mainly because they do not make the ontological commitments needed to hold *diachronic* and *synchronic environmental externalism* (DEE and SEE, respectively). In conclusion, I argue that emotions are representations, although an embodied and action-oriented type of representation.

2. PERCEPTUAL THEORIES

Cognitivists claim that emotions have intentional objects and are subject to semantic norms. They also highlight that emotions are about things that

matter or are of value to us. The frequent conclusion is that emotions are cognitive evaluations. Yet, of course, there are many alternative explanations for the emotions' aboutness, and there is no consensus about how to account for the evaluative character of emotions. The objections against cognitivism discussed in the last chapter point out that theories of emotion are in need of alternative explanations of the emotions' aboutness and evaluative character.

Many contemporary authors suggest seeing emotions as perceptions or perception-like states. This is not surprising given that the main objections raised against cognitivism (see Chapter 1) have been brought forward in a very similar way to defend the position that perceptions do not have conceptual content, but rather nonconceptual content. In the literature defending the nonconceptual content of perception we find (1) the argument that animals and infants have perceptions but not concepts (Peacocke 2001), (2) the argument that propositional knowledge and perceptions can contradict each other and that perceptions show the feature of cognitive impenetrability or belief independence (Evans 1982, 123; Crane 1988), (3) the argument that perceptions are passive states of mind that we cannot avoid happening to us—that is, we cannot avoid perceiving red when seeing a ripe tomato (Prinz 2004, 236), and (4) the argument that perceptions have a certain phenomenological character that cannot be captured by a simple appeal to a judgment (Evans 1982).

Perceptual theories all aim to meet the objections that classical cognitivism raised with regard to feeling theories of emotions, but at the same time they share the impression that cognitivism has been overintellectualizing emotions by claiming that they have conceptual content. Perceptualists tend to focus on the particular phenomenal character of emotions, arguing that emotions represent concerns, not in the form of a judgment, but more in the form of a Wittgensteinian perceiving-as: in pride we perceive ourselves *as* being worthy of praise; in anger we perceive the other person *as* being offensive (Roberts 2003). This form of affective perception could also be described as *Gestalt*-like, arguing that emotions, like perceptions, present complex situations to us in which certain salient patterns are immediately perceived as such. Just as in perception we perceive a vertical line with a dot above it immediately as the letter *i*, so in emotions we immediately perceive what is of value for us (Döring 2007). Emotions are perceptions with a normative content and a motivating component; they are direct perceptions of values or ways of seeing what to do. For Sabine Döring, this is the reason that emotions, when endorsed by the subject, can account for motivational internalism. They are inputs to content-involving practical reasoning, without the need for any inference in the occurrence of the emotion. The content of an emotion is the noninferential justification of a moral judgment's content. Döring, like most perceptualists, thereby clearly embraces representational cognitivism while denying constitutive and etiological cognitivism. Emotions are not constituted by judgments and can be directly triggered by perceiving the world; they need not be preceded by judgments.

Perceptual theories capture the phenomenology of emotions far better than cognitivist approaches do. Trembling with fear is a way of seeing a situation as dangerous, and swelling with pride is a way of seeing oneself as praiseworthy. From a first-person point of view it does not seem as if we are drawing appraisals along a personal goal hierarchy when feeling scared or proud. The claim that emotions are ways of seeing-as that highlight salient features of a situation comes much closer to an adequate description of experience.

An obvious problem with the perceptualist account, however, is that most of its defenders, from De Sousa (1987) to Döring, use the notion of perception merely as an analogy to highlight the noninferential character of emotional content. The authors neither precisely define what perception is nor do they point out which aspects perceptions and emotions share and which they do not. Emotions are described as noninferential ways of interpreting (normatively) relevant scenarios or patterns of salience. The main point of the perceptualist account seems to be the claim that emotions constitute direct access to the world through which properties of normative relevance can be directly perceived.

Perceptualist accounts claim to successfully bypass all the objections raised against cognitivism: perceptions are present in infants and animals, they show the feature of cognitive impenetrability, they have a specific phenomenological aspect, and they occur automatically. Yet at the same time, perceptualists, just like cognitivists, hold that emotions are ways of interpreting the world or of focusing on what matters for us or even of evaluating situations that are of normative relevance. With the rather vague notion of what it means to be perception-like, the picture of how these perception-like states can fulfill such complex tasks remains blurry. Lazarus might ask, "What could it mean that emotional perceptions are not in need of any inference in their occurrence?" and Solomon might ask, "How can emotions be normatively assessable if their content is nonconceptual?" The claim that emotions are a form of seeing-as, in which seemingly complex content is represented in a nonconceptual format, without any previous inferences, is rather *ad hoc* so long as the definition of being perception-like remains unclear. Yet perceptualists are in need of a definition, since there are obvious differences between perception and emotion:

1. Emotions are valent; that is, they are positive or negative in the way that all affective phenomena are. Perceptions are only valent when they are colored by affective reactions.
2. Emotions are motivating. Fear strongly motivates to flee, anger to attack. Perceptions do not seem to motivate in the same way.
3. Emotions include patterns of bodily reaction such as increased heartbeat, release of adrenaline, widened eyes, and so on in fear. Comparable forms of inner arousal or expressive behavior do not accompany perceptions.
4. Perceptions are direct responses to external stimuli that reach the organism through a sense organ. Emotions do not have specific sense organs.

5. Emotions are follow-up mental states. They need to be preceded by another mental state; for example, we have to perceive a loud noise before getting scared or think about a forgotten birthday before feeling guilty.[1] This is not the case with perceptions.

None of the defenders of perceptual accounts denies these differences between perceptions and emotions. However, perceptual accounts leave the question of how to account for these differences entirely open. Furthermore, for naturalist accounts there is one main problem with perceptual theories. Perceptual theories assume that emotions are direct perceptions of values. Yet emotions do not appear to have any specific sense organs through which the organism could extract any kind of information. Therefore, in claiming that emotional content is given to us neither through causal processing nor through cognitive inferences, perceptualists thereby postulate a rather opaque emotional faculty without giving us any idea about either a cognitive or a physiological route through which the relevant information can be grasped. This is certainly problematic for any theory but definitely a "no-go" for a naturalist account.

The account developed by Prinz (2004) is a naturalist account that identifies emotions with perceptions and also offers a definition what perceptions are. Prinz defends a neo-Jamesian approach, according to which we perceive things of value through the bodily arousal involved in emotions. The bodily reactions do the work that the sense organs do in sense perception: they transfer information. That emotions have external intentional objects that can be tracked by inner bodily arousal is explained through a teleosemantic framework. I discuss this framework and its role for embodied accounts in the fourth section. To give bodily reactions a constitutive role in emotion is particularly interesting because such a theory could potentially account for the motivating and valent qualities of emotions. Emotions, according to such an approach, can be seen as motivating or action-oriented, because the bodily arousal involved in emotions prepares the organism for action. Emotions can be seen as valent in the way that bodily sensations are valent. Finally, an embodied account can also account for the particular phenomenology of emotions, their drive, and their intensity. In emotional experience we see something as dangerous *and* feel tempted to flee the situation at the same time, since the pattern of bodily arousal prepares the body for action. The representation and feeling resulting from the pattern of bodily feedback are evaluative and directive in themselves.

3. NONCONCEPTUAL CONTENT AND EMOTIONAL EXPERIENCE

While the comparison of emotions and perceptions leaves many questions unanswered, it certainly points in the right direction by saying that emotions can be intentional without having conceptual content. The debate over the

nonconceptual content of perception highlights the tidy connection between perceptions being about something and feeling a certain way. From phenomenological descriptions we gain a good understanding of the particular way in which perceptions present their objects to us. This is true for emotions as well. A description of how emotions feel offers a good *prima facie* understanding of the representational format that emotions have and the way that bodily reactions, aboutness, motivation, and valence are connected in emotions.

The character of perceptual representation has often been described as different in kind from conceptual representation. The kind of information we receive through perception simply cannot be captured in words. We cannot explain what a pineapple tastes like, or how a toothache feels, or what the sea looks like to somebody who has never had these impressions herself, yet a pineapple has a special taste, and once familiar with it, we can recognize it from among hundreds of other tasty things without difficulty (Hume 1739/2007, I.i.1). The same is true of emotions. Even apparently similar emotions, such as embarrassment and shame, clearly differ in terms of how they feel, yet the difference can hardly be explained to somebody who has never felt the emotions herself.

Another claim, about visual perception in particular, is that it is more fine-grained than conceptual content. We can capture a part of what we see at a certain moment in a phrase like "there is a cup on the table," but such an utterance will always reduce what we have seen to a very particular aspect. We have also seen the shape and color of the cup and the table, the room in the background, the light conditions, and so on. And even if we were to describe such a visual experience in detail, it could never be detailed enough to communicate to somebody else what we see (Dretske 1981).

One might think, however, that it is misleading to say that the special phenomenal character of emotions has to do with their being somehow more "fine-grained" than judgments. Emotions are intense states that grab our attention, much like pain or hunger. It is often said that emotions "color" everything we perceive: when we are sad, we perceive the world as a gray fountain of despair; happiness, on the other hand, produces the proverbial view of the world through "rose-colored glasses" (such metaphors also make clear that emotions are indirect states of mind that highlight what is perceived through a prior state). But one might say that it is not a rich variety of shades of gray that we experience in sadness. Rather, it is the same dull depressing feeling again and again that constantly colors our view of the world. So, what does it mean to say that we represent emotional content in a fine-grained format? And, if bodily arousal is supposed to be constitutive of emotions, what role does it play in their being fine-grained? After all, the entire cognitivist tradition argues that Jamesian feeling theory is wrong because the bodily arousal involved in emotions is rather vague and cannot even account for the felt distinction between positive and negative emotions.[2]

With regard to the first point, it is misleading to think of emotional content as being presented with no fine-grained structure at all. The fine-grained structure of emotional feelings differs from visual perception insofar as it is not constituted by the processing of the rich amount of data entering the retina, but mainly by the rich amount of information stemming from the various sources of bodily feedback involved in emotions. This makes intensity a central factor. Sadness can come in degrees; it can come with desperate sobbing, silent tears, or just a typical posture of the body and an expression of the face. Yet sobbing and silent tears differ not only in the degree of intensity but also strongly differ with regard to the bodily feedback they provide.[3] Fear can feel more or less intense as well. And, again, "intensity" is an abstract notion for particular bodily reactions that do not all have to be present and can vary in degree. Sometimes we feel our muscles tremble and our heartbeat increase; sometimes we feel only a prickling sensation in the stomach when we are afraid. Fear can vary in duration, depending on the intensity of the bodily reaction constituting it. Together with the heartbeat and muscle tension, breathing rhythm is another factor that contributes to the bodily orchestra constituting the feeling of being afraid. Breathing is, furthermore, a phenomenon that nicely illustrates the direct influence between bodily factors and feelings even in self-observation: fear can decrease when we breathe slowly and deeply, while quick and shallow breathing can produce a panic attack (Laird 2007).

A standard objection to such an account is that bodily arousal is too vague or indifferent to account for the huge variety of emotional feelings. As has been argued in detail in the previous chapter, the view that the bodily reactions involved in emotions are just a random kind of arousal is incorrect with regard to the underlying concept of what bodily arousal is. Bodily arousal is frequently reduced to visceral feedback by cognitivists, but an embodied account of emotions can include fine-grained changes in facial expression, bodily posture, endocrine level, respiration patterns, and so on. Emotions obviously include a large number of bodily experiences such as trembling, sweating, feeling choked, being on the verge of tears, having a lump in one's throat, blushing, having butterflies in one's stomach, having one's legs turn to jelly, and so on.

The large number of bodily reactions involved in emotions is what explains the fine-grainedness of emotional feelings (at least to a large degree). Fleshing out a theory of emotional experience means describing the calmness and relaxation associated with sadness as an effect of decreased circulation (Kreibig 2010), of the head hanging down on the contracted chest, and of the lips, cheeks, and jaw all sinking downward to make the whole body feel heavy and immobile rather than just calm and relaxed. A typical case of a fine-grained difference between two emotions is the difference between embarrassment and shame. It is again a difference that can be described as a difference in intensity. One can blush in a more or less intense way and the stress reactions guided by the HPA axis might also be more or less intense.

But there might also be differences in behavior: embarrassment is typically accompanied by a coy smile, while shame is accompanied by a shrinking posture. The difference in intensity, plus the differences in behavior, can account for a difference in how embarrassment and shame present their objects and in how they feel.

The description of emotions as representing things as fine-grained because of the rich amount of bodily feedback from various sources might still sound incomplete. What is striking about emotional representations is not their fine-grainedness but their urgency. In this respect emotions resemble sensations such as pain or hunger that are intense, attention grabbing, and directly motivating for action. Picking up on the intense motivating component of emotional feelings, the psychologist Nico Frijda describes emotions as reactions to *urgent* situations. "Urgency is the irreflexive counterpart of felt emotional intensity. Difficulty and urgency are the situational meaning components corresponding to emotional upset, to emotion in the excited sense of the word" (Frijda 1986, 206).[4] Sensations such as pain and hunger are feelings that inform us about the current state of the body and strongly motivate us to react to them. For example, the pain caused by touching something hot forces us to better protect our hands, and the unpleasant feeling of starvation forces us to go and find something to eat immediately. Emotions work in a very similar way: they can be understood as embodied representations to situations that are urgent or matter to us. Emotional feelings owe their intense phenomenal character to the perturbations in visceral organs and adjustments of skeletal muscles constituting them. The heart beating faster, adrenaline rising, and muscles tensing constitute the feeling of fear and prepare the organism for action.

There is an obvious objection to the claim that emotions are motivating because they involve bodily arousal. The claim that emotions are directly motivating might seem plausible for extreme cases of fear and anger, but in the cases of sadness, pride, or joy, it is far from obvious that they have any motivating force at all. Usually, the conclusion drawn from this observation is that motivation is not an intrinsic component of emotions (Roberts 2003, Deonna and Teroni 2012). I disagree.

To address this concern, the notion of motivation itself needs to be clarified. The most crucial point, I think, is that emotions do not always motivate *intentional* actions, such as beating somebody who has offended you. Emotions mostly motivate simple reactions and behaviors such as staying in a pleasant situation, avoiding others' company, or smiling at everybody around you. This type of motivation is not underlain by strong bodily reactions, as in the case of urgent fear where the racing of the heart can literally be felt.[5] When a person feels proud because she has just been complimented, the pleasant state she is in might be accompanied by a slight increase in the skin conductance level and heart rate, producing a very weak sense of excitement, and by a slightly changed bodily posture generating proprioceptive feedback.[6] These two things together may suffice to trigger certain

follow-up behaviors, such as telling more jokes or simply being chattier than usual, having more wine than was previously planned, and so on. Emotions can feel more or less intense because the bodily arousals are more or less intense and the object of the emotion is thereby represented as a more or less urgent action affordance.

Another objection is that the phenomenology of emotions can include the perception of bodily changes but is not restricted to it. Rather, certain cognitive activities, such as the sense of altered or focused attention, can accompany emotions and contribute to complex human emotional feelings as well.[7] I agree, insofar as it is not my aim to argue that emotional feelings are constituted by the detection of bodily arousal *alone*. Yet I think that bodily reactions do constitute the *core constituent* of emotional feelings. Other elements, such as increased attention, might contribute largely to the fine-grained character of adult emotions. Much has been written about the degree to which language and background beliefs contribute to our feelings (see, e.g., Goldie 2000). The present approach does not deny these kinds of influences. What I am arguing for is that all emotions fundamentally involve bodily arousal, which constitutes their intentionality and how they feel. All emotions are intrinsically motivating and their motivating power is, to a large degree, constituted by the bodily reactions they involve. The case of a person who thinks of herself as laudable because of her impressive performance without feeling the slightest change in bodily posture, hormonal process, or inner arousal is not a case of pride or any kind of affective state.

To claim that bodily arousal somehow constitutes embodied action-oriented representations and thereby emotional experience gives us a good *prima facie* guess of how bodily reactions, intentionality, motivation, and valence are related in emotions. Bodily responses are organized, evolved patterns whose function is to prepare the organism for action. These responses constitute emotional aboutness and feeling. Therefore, emotions do not represent bare facts but urgent concerns or affordances for action. And, therefore, emotions feel as intense as the urge is and are pleasant or unpleasant in a way that is similar to bodily sensations. Yet it is important to note that for an emotion to be constituted by bodily reactions does not mean that it has the body as its object. Similarly, for a feeling to be a bodily feeling does not mean that it has the body as its object. The body can also enter awareness as that *through which* something else is experienced (Maiese 2011, Colombetti 2014, 113). The remaining question is therefore what kind of theoretical framework can best make sense of the body being "in the fringe of" emotional representation or the medium of emotional representation. This involves the question of what kind of approach can combine this *prima facie* construct with a theoretical framework that can make sense of the normative structure of emotions. I discuss several embodied accounts with regard to that question in the section that follows.

4. EMBODIED ACCOUNTS

Many current theories take emotions to be noncognitive and fundamentally embodied (e.g., Prinz 2004, Maiese 2011, Hutto 2012, Colombetti 2014). These approaches share the opinion that emotions are about organism–environment relationships or that things are of value for the organism's well-being. They also agree that bodily reactions play a constitutive role for emotions and that these bodily reactions, due to an adaptive history, constitute the emotions' intentionality. The general assumptions are that these theories agree upon the following:

1. Bodily reactions play a constitutive role for emotions.
2. Emotions are about organism–environment relationships that concern the organism's well-being.
3. The emotion's aboutness should be understood as being nonconceptual.
4. The emotion's aboutness should be understood in terms of an adaptive history in which bodily reactions gained the function to constitute the emotion's aboutness.

All embodied approaches rely on some kind of naturalist approach, though, as I show later, these are rather different versions of naturalism. Embodied approaches differ in radicalness. While Prinz argues that only basic emotions are constituted by distinctive patterns of bodily reactions, Colombetti holds that "all emotions come in complex organismic patterns" (2014, 72). Furthermore, a major difference between the various approaches is that some take emotions to be representational, relying on a teleo- or bioemantic framework, while others take emotions to be intentional but not essentially representational, relying on an enactivist framework. The different directions embodied approaches have taken with regard to these questions are analyzed through a discussion of Prinz's teleosemantic approach, Dan Hutto's radical enactivist approach and Colombetti's autopoietic enactivist approach. In particular, I test these theories on the question of whether they are able to account for the normative assessability of emotions.

a. The Teleosemantic Approach

Prinz calls emotions embodied "radar detectors alerting us to concerns," relying on Fred Dretske's (1986) account and his famous comparisons of representational systems with technical devices. According to Dretske, all mental states are representations and all representations carry information and can misrepresent that information. Information occurs everywhere in the world in the form of stable covariance or lawful connections. For example, since fire causes smoke, smoke indicates fire; since the chemical constitution of stars causes them to shine, the stars' shining can indicate their chemical constitution.

Representational systems have the function of detecting certain types of natural information. Representations can malfunction, and it is therefore on the level of representations (not on the level of natural information) that misrepresentation can occur. For example, the function of a fuel gauge is to inform us about the amount of petrol in the tank. If the fuel gauge is working properly and is correctly connected, the needle indicates whether the tank is full or empty; it is designed to do so. Analogously, we can say that a mental representation of fire has the function of representing fire, and not dry ice, and that fear has the function of alerting us to dangerous situations. While technical devices owe their function to their designer, biological representation systems are created by evolution and/or are calibrated by learning. The latter have biological functions that allow us to distinguish between representations and misrepresentations. If, for example, a frog is snapping for bypassing flies and bullets alike, the best explanation can be given by looking at the environmental conditions under which frogs have developed their perceptual systems: the former is an appropriate reaction, while the latter is the result of a misrepresentation.

Following Dretske, Prinz defines emotions as embodied appraisals "set up to be set off" by urgent concerns that our ancestors repeatedly met in their environments until a pattern of bodily reactions evolved that prepared them to deal with the situation in question. These "urgent concerns" or "core relational themes" are the contents that emotions have the function of representing. Misrepresentations occur, for example, when an organism reacts with fear to a situation in which there is no actual danger present. To account for the different contents of each emotion type, Prinz simply adopts Lazarus's list of core relational themes, according to which fear represents a situation as being dangerous, anger represents a "demeaning offense against me and mine," and so on (Lazarus 1991).

This leads to an obvious problem, however, because core relational themes in Lazarus's sense are represented by inner judgments triggered by a complex appraisal process. According to Prinz, however, we can say that Lazarus was right about the *content* of emotions, while his theory is misleading with regard to their *form*. Emotions represent core relational themes, such as losses and dangers, but they do not represent them in the form of a judgment or any other conceptual state. When we fear something, our racing heart indicates that we are in danger and we represent danger by registering that our heart races. A specific emotion type that has a particular content—that is, that represents its core relational theme—thus depends on the existence of specific patterns of bodily arousal that differ from one another and can therefore be perceived as different types. These patterns do not have to be identical; they only have to have a common origin and thus a common function. It is the function of fear, for example, to detect dangerous situations and this function is performed by perceiving patterns of bodily arousal that themselves are typically triggered by the presence of a core relational theme.

Core relational themes exist as the class of features to which an emotion has the function to respond. Accordingly, Prinz holds that core relational themes are relational properties. Fear represents danger and danger is a relational property that exists only *between* the organism and the world. A snake has the property of being dangerous only in relation to some organisms, just as a father has the property of being a father only in relation to his children. Still, contrary to response-dependent properties, relational properties qualify as objective facts. That the snake is dangerous to certain organisms does not depend on subjective judgments or any representational capacity. Response-dependent properties are properties that owe their identity to the way we represent them. They are properties that cannot be individuated without taking our representational capacities into account. Many people take, for example, color or value properties to be response-dependent: My cell phone has the property of being valued by me only in relation to me valuing it. The rose has the property of being red only in relation to me seeing its color.

For externalist and especially teleosemantic approaches to representation response-dependent properties are not good candidates to explain representational content. It is a central claim of externalist approaches that the extension (and not the intension) plays a defining role with regard to a representation's content. Yet response-dependent properties are intensionally defined properties. Furthermore, teleosemantic approaches assume not only that content is externally defined, but also that representations have the function of representing a certain content, because it was useful somewhere in the history of our ancestors to be able to represent this content. What DEE claims is that we acquire representational abilities in response to things that were *already* nutritious, poisonous, or dangerous to us before we were able to represent them as such. That the snake has the relational property of being dangerous to the organism explains how a fear reaction understood as a danger-detecting system evolved: our ancestors repeatedly faced dangerous situations, until they developed a detecting system to avoid them. This account rejects the claim that emotions are constituted by mental evaluations and replaces it with the claim that, in response to certain adaptive pressures in the environment, we developed bodily reactions to detect dangerous situations and adequately respond to them.

By embracing relational properties Prinz can account for DEE, that is, the claim that the environment takes on an active structuring role in the evolution of emotions. Teleosemantic approaches are committed to the claim that organisms develop representational powers through direct interaction with the world and in response to certain adaptive pressures exerted by their environments. Various organisms have developed eyes and the ability to use visual information because the ability to use visual data is of immense value for guiding the organism's behavior. This view depends on the claim that there are things in the external world that our ancestors faced repeatedly, thus making it beneficial for these organisms to develop mechanisms in response.

Prinz satisfies these requirements by claiming that "emotions are reliably caused by both bodily changes and core relational themes, but they seem to have the function of detecting only the latter" (2004, 67). Prinz assumes a chain of natural information that runs from a particular object in the world with the relational property in question, through the bodily arousal, to the perception of the bodily arousal that has the relational property as its distal content.

SEE is the claim that the world is its own best model and that since it is there and in permanent causal contact with us, we need not assume complex cognitive representations to be in place. Although Prinz does not belong to the camp of antirepresentationalists most famous for defending SEE, he does defend a Dretskian version of the claim. Prinz adopts Lazarus's list of core relational themes, claiming that Lazarus was right about the content of emotions, but not about their form. What Dretske knows, and cognitivists don't, is that complex states of affairs can be represented by something quite simple. As Prinz puts it, "[a]ppraisal theorists often mistake the complex property represented by emotions for the inner representations that constitute or precede our emotions" (2004, 65). By taking emotions to be simple representations about complex states of affairs, Prinz subscribes to the claim that the presence of the world itself can account for a great deal of complexity that we tend to ascribe to inner representations.

However, it should be noted that Prinz's account of DEE and SEE depends on the ontological commitment that core relational themes are relational properties found in the organism's environment. We are surrounded by situations that contain actual dangers and actual offenses to us. This is what a noncognitivist and externalist embodied account must include to satisfy DEE and SEE. Relational properties, according to Prinz, reliably cause emotions, and emotions were set up because it was useful for the organism to be able to represent relational properties. Such causal powers cannot be ascribed to response-dependent properties, that is, properties that owe their identity to the way we represent them. If we say that "being dangerous" is a property that an object only has because we represent it as such when being emotional, then it remains opaque how this property can reliably cause emotions and how it made it useful to have emotions before there were any.

Once the claim is made that properties such as "being dangerous" and "having been unfaithful" objectively exist in relation to the organism, it is only a small (and perhaps inevitable) step to say that these relational properties are normative properties. This appropriately captures the nature of the relational properties in question and offers an explanation of the valence of emotions: emotions are valent, that is, positive or negative, because they respond to properties that are good or bad for us.

It might therefore come as a surprise that valence has a completely different place in Prinz's theory. According to Prinz, emotions are perceptions of bodily arousal that represent a core relational theme. Furthermore, every emotion involves a separate "valence marker" that evaluates the emotion

as good or bad and then triggers further reactions (see Prinz 2004, 160ff., 2010). Valence is the driving force of emotions, motivating seeking and avoidance behavior. However, instead of calling valence an intrinsic aspect of emotional content, Prinz distinguishes the perceptual content of emotions from valence as a nonperceptual add-on. According to his account, when we are scared we represent our state of being in danger by introspecting our bodily changes, but the fact that danger is something we *should* avoid is not part of the perceptual representation itself nor is the motivation to avoid it. Valence is an inner reinforcer (Prinz 2004) or a reward/punishment marker (Prinz 2010), an unconscious mechanism that influences behavior. However, this marker is not an evaluation of the event that causes the emotion; it is an evaluation of the emotion itself. The value registered is directed at emotions rather than at the world. When fear is experienced, the core relational theme "danger" is represented through the bodily changes taking place. However, an unconscious marker that says "less of this emotion" constitutes the motivational force of the emotion. The marker evaluates the ongoing bodily changes as being bad.

A neural valence marker is certainly a possible solution for the question of valence but, particularly for an account taking emotions to be embodied appraisals, it is—to say the least—not the most elegant solution. Prinz (in a very Dretske-like way) neatly distinguishes between perception, evaluation, and the regulation of behavior and action. We perceive danger through bodily reactions, the valence marker evaluates danger as something bad, and response behaviors are triggered. This separation adopts what Susan Hurley (1998) calls the *sandwich model* of the mind on a microlevel because it separates perception, evaluation, and action. Emotions, according to Prinz, are perceptions that receive and store input from the body. However, it does not make much sense to call such perceptions embodied *appraisals*, since they merely indicate core relational themes without involving any evaluation. Fear indicates danger without simultaneously evaluating danger as something bad or motivating the organism to avoid anything of the sort. Such an account loses most of the explanatory advantages a well-understood embodied account has. The advantage of an embodied account is that it can make sense of the close connection of bodily reactions, intentionality, motivation and valence in emotions by saying that the bodily reactions constitute an intrinsically action-oriented representation that feels good or bad because it motivates one to approach or avoid something.

The reason for Prinz to separate perception from evaluation and action lies in his aim of avoiding the commitment to normative realism. Prinz claims that emotions just represent objects and their properties, which is not normative. The representations are then internally evaluated as being either positive or negative. But it remains an open question whether the objects of emotions can be understood as value-free properties in a reasonable way at all. How can we represent the relational property of "being dangerous" but insist that further evaluation is required to describe something dangerous

for the organism as bad for that organism? Core relational themes are always described in normative vocabulary, and as far as I can see, there are no fitting redescriptions of them in nonnormative terms, which I take to be a strong indicator that accounts that deny the normative dimension of emotional content are inadequately reductionist.

Prinz distinguishes between basic emotions such as fear, anger, and disgust that have been set up by evolution, and higher cognitive emotions such as jealousy, shame, and guilt that are the result of a learning history. When a basic emotion (e.g., fear) occurs several times with a certain type of judgment (e.g., "My lover is unfaithful to me"), we develop a mental file (a "calibration file") that can result in a frequent triggering of an emotion such as jealousy. Once such a file is installed in our long-term memory, that emotion can be triggered without being preceded by a judgment. For instance, the smell of an unfamiliar perfume on a lover's clothes can be an immediate trigger of jealousy. Still, the essential elements of higher cognitive emotions remain bodily arousal and the core relational theme to which it is directed; these two features also suffice to individuate the emotion. Yet we need to be able to entertain judgments, such as "I have been betrayed by my loved one," before jealousy as an emotion type can be set up.

Based on the evidence presented thus far, we are in a good position to question the ability of the approach to account for the normative assessability of emotions. With regard to semantic norms, Prinz has a favorable starting point as he merely needs to refer to the teleosemantic framework and the function this framework ascribes to emotions. The evolutionarily acquired function of representing something explains the normative character of representation. We can say that a representation *fails* to represent what it is supposed to; for example, that the frog is misrepresenting the bullet passing by to be a fly because fly representations have the evolved function of representing flies and not bullets, even if they are caused by bullets.

The notion of a function can make sense of the occurrence of misrepresentations and, at the same time, provide a criterion to distinguish between representation and misrepresentation. This is also true of emotions given that they have adequacy conditions. Emotions need to be described as states that have the function of alerting us to concerns in certain situations; if the concern is really present and urgent, then the emotion is adequate.

It is hard to tell what Prinz would say about the emotions' being subject to rational norms, since he doesn't address the issue himself. One option is to refer to the emotions' being in close connection to other mental representations, including judgments, that can frequently trigger or accompany emotions. Emotions themselves are embodied appraisals. That they appear to be subject to rational norms standing in certain logical relations to us, the world, and other emotions can be explained by the narrative structures they are embedded in. Emotions occur in social contexts and are governed by certain conventions that make it appear to be reasonable that we feel relieved when the object of our fear is removed or jealous when a loved one

turns to another person. But this would shift a lot of the explanatory burden from teleosemantics and the content of the emotions themselves to other cognitive structures that closely interact with emotions.

Also, Prinz can account for certain features that Lazarus aims to capture with the appraisal structure. Prinz claims that emotions are individuated with regard to their core relational themes and the pattern of bodily arousal set up to be set off by the core relational theme. Kenny says that norms of appropriateness, not causal laws, restrict emotions' intentional objects. It is because of such norms, and not causal laws, that we cannot envy ourselves. In teleosemantic terms, the object is restricted by biological regularities that concern the function of the representation. It is appropriate to envy others because this is the biological function the emotion was set up for. It does not have the function to envy oneself. Yet this is inadequately reductionist when applied to an emotion such as envy. Furthermore, it does not capture that it is a conceptual truth that it is always inappropriate to envy oneself. Therefore, it seems that Prinz cannot account for the emotions' being subject to rational norms. I come back to the question of how a naturalist theory of representation can account for emotions' being subject to rational norms in Chapter 5.

I argued earlier that some emotions constitutively respond to a certain social rule or norm. For embodied accounts, it is the social emotions that are most difficult to explain. Traditionally, the group of social emotions has been categorized as higher cognitive emotions because these emotions seem to involve complex cognitive evaluations of the self or others in relation to certain social rules, which makes it hard to see how the emotions in question could be noncognitive embodied responses.

Prinz holds the view that we only acquire these emotions after we have acquired concepts and are able to judge a situation—for example, by thinking, "My lover is unfaithful to me." Higher cognitive emotions share the embodied reactions of other emotions. Guilt, for example, entails the bodily reactions of sadness. At a later stage, however, higher cognitive emotions are "recalibrated" through judgments to represent a new core relational theme. The idea is that the perception of the bodily reactions that represent loss in sadness come to represent the violation of a social rule when triggered through a certain mental file that was installed through learning.

What is confusing about this account is that it introduces judgments only to capture the complexity of the appraisals that seem to trigger these emotions. Prinz takes it that the relational properties in question individuate the emotion type. Whether something is a case of guilt or sadness can only be determined by looking at the relational property that the emotion is set up to represent. If Prinz were to abandon this claim it would make him a demi-cognitivist; he would have to claim that judgments concerning social rules are essential parts of higher cognitive emotions, not merely one possible trigger among others. An embodied account here is in need of a certain form of normative realism. It would be an option to claim that

properties such as "having transgressed a social norm" are relational properties. I argue for this option in the next chapter. Yet Prinz in his later works explicitly asserts that normativity is entirely subjective. Morality, according to Prinz (2007), is constituted by emotional reactions and not by the things emotions respond to. But this claim is not consistent with a noncognitive embodied theory of emotions. It raises the question of how to account for DEE and SEE. Furthermore, it also questions the main point of his sentimentalism. How can emotions constitute morality if their content is constituted by judgments about the transgressions of moral norms in the first place? According to such an account, moral judgments arise prior to emotions and not the other way around.[8]

In conclusion, the teleosemantic approach, as it is defended by Prinz, cannot account for emotions being subject to rational norms and for emotions being about social norms without fundamentally changing its ontological commitments.

b. The Radical Enactive Approach

Daniel Hutto has developed an enactive account of basic states of the mind as intentional but without content (Hutto 2005, Hutto and Myin 2013). He applies this account specifically to emotions (Hutto 2012). This application relies heavily on Prinz's claims that emotions are bodily feelings and, at the same time, intentional and that emotions are about core relational themes. However, Hutto departs from Prinz by defending the view that basic mental states such as sensations, perceptions, and emotions are intentional yet not representational, since these states do not aim at truth.

Hutto therefore suggests a tweak of Prinz's embodied account that requires "giving up the idea, for example, that in the right conditions my anger *represents truly* that there has been a demeaning offense against me and mine" (Hutto 2012, 178). Emotions do not represent situations in a truth-evaluable way; when we are emotional, we respond to core relational themes in distinctive ways that implicate experiences of bodily feelings. Hutto argues that even accounts that assume nonconceptual content or speak about the organism "enacting meaning" (as in Colombetti 2014) commit themselves to a certain type of truth or adequacy conditions and thereby to norms that make it possible to differentiate true from false, or adequate from inadequate, emotions. Hutto assumes that the commitment to such norms is something hardly reconcilable with naturalism and for that reason denies their relevance on the level of basic cognition altogether (Hutto and Myin 2013).

Hutto not only pleads for parsimony on the ontological level but also further argues that the ascription of content on the level of basic cognition is not needed for an adequate description of the phenomenon. According to Hutto, basic cognition, including emotion, should be understood as (1) reliably caused by (or nomically dependent upon) the occurrence of certain

external features, (2) disposed to produce certain effects (under specific con-
ditions), and (3) selected because of its propensities for (1) and (2). This
requires a separate explanation of why a state should qualify as represen-
tational and not simply as possessing the properties (1) through (3). Only a
state that has the proper function of "saying or indicating that things stand
thus and so, and being consumed by other systems because of what it says
or indicates" (Hutto 2012, 177) should be taken to be a representation.
A representation according to Hutto is a mental state that has the function
to indicate states of affairs in a way that can play an informative role in
a larger cognitive system. Perceptions, sensations, and emotions appear to
have the proper function of guiding a system's responses to specific kinds of
worldly offerings.

Contrary to Prinz, the approach presented by Hutto denies that emotions
are subject to semantic norms. Hutto takes emotions to be intentional but
not *intensional*: they do not have content in the sense that they represent
something *as* something. Accordingly, they do not have adequacy condi-
tions or the function of representing something. Such a characterization of
emotions leads to an unacceptably reductionist result. If we try to charac-
terize fear as a reaction that is reliably caused by the presence of certain
dangerous stimuli that under normal conditions cause typical bodily reac-
tions (such as the heart beating faster, adrenaline being released, and so on)
and finally think of the whole process as a mechanism created by evolu-
tion, we end up with a crude and simplistic behaviorist model of emotions.
Such a model takes emotions to be nothing other than evolutionarily created
input–output patterns.

If the aforementioned three criteria suffice for characterization, then
emotions are just as intelligent as simple reflex-like reactions. Consider the
blinking reflex: whenever something quickly approaches the eye, it causes
the eye to blink. Such a reaction is (1) reliably caused by external features,
(2) disposed to produce a certain response, and (3) selected to do just this.
By paralleling emotions with reflex-like reactions, Hutto is subject to the
most basic objection that cognitivism has brought forward, that a charac-
terization of emotions as mere bodily reactions and behaviors completely
misses the normative dimension of emotions. Fear reactions are not just
automatic responses to a given stimulus. They are reactions that can be
adequate or inadequate depending on whether the situation in question is
really dangerous or not.

Hutto might respond that taking emotions to have adequacy conditions
is an overintellectualization that is a result of our way of talking about emo-
tions in everyday language and that it suffices to talk about emotions as
being intentional but not intensional. To give a clearer idea of what this is
supposed to mean we should have a look at particular emotions. The usual
way to describe emotional content is to say that in fear we present the dog
in front of us *as being dangerous*. This has led cognitivists to claim that
emotions are not simply "about something": in fear we perceive something

to have a certain quality—namely, that it is dangerous. Prinz captures this approach by saying that emotions represent core relational themes and that these themes are relational properties. The dog in front of us has the property of being dangerous for us. Fear is a representation that was set up to detect this particular property of the dog.[9] For Hutto, the trouble begins with the question of what emotions are specifically about. When we are afraid of a dog, are we afraid of the dog's response-dependent property of being dangerous to us, or is it the property of having huge teeth? Or is it the fact that the dog barks loudly? Or is the property somehow supervening on the whole situation?

Intentionality alone, as Fodor (1990) points out, does not give us any precise distal content. It seems that Hutto is willing to "bite the bullet," but the resulting account, in its reductionist aim, is hard to distinguish from classical behaviorist approaches. As long as the intentional object cannot be specified at all, it does not make a huge difference whether we say that an organism shows an automatic response to a certain kind of external stimulus or whether we speak of a reaction that entails intentionality.

The matter becomes even more confusing when we take into account the fact that Hutto wants emotions to be "reliably caused by external features," yet at the same time he thinks that it is not necessary to talk about core relational themes as objectively existing properties, suggesting that they can also be understood as response-dependent properties.[10] Following such an approach, there is nothing that the things we happen to be afraid of have in common except that they all have the property of appearing dangerous to us. However, this means that there is no class of external features that causes fear reactions. "Being dangerous" is a property that owes its identity to the fact that we happen to represent things as dangerous. At the very least, this is how the term "response-dependent property" is traditionally defined. Therefore, response-dependent properties must be defined intensionally rather than extensionally; they owe their identity to the way we represent things. Things are funny, elegant, or dangerous because we see them that way. In any case, how can we say that "scary things" are defined intensionally if we claim, on the other hand, that fear reactions do not have intensionality?

Here we see again how embodied approaches rely on DEE and SEE. If the claim really is that emotions are intentional without being intensional, then I take it that the class of stimuli that an emotion is prone to respond to *must* be defined extensionally because I do not see how response-dependent properties could do that kind of job (since they rely by definition on responses by the organism). Thus, Hutto's approach is not only inadequately reductionist (an objection that might be acceptable) but could also become contradictory if Hutto really wants to claim that emotions represent response-dependent properties while at the same time holding that emotions must be causally triggered by external stimuli. The claim that emotions represent response-dependent properties means that there is no relevant class of

external stimuli that could count as the group of typical intentional objects of emotions. Therefore, Hutto must either revise his claim that emotions are reliably caused by external features or accept the ontological claim that the objects of emotions are relational properties, such as being dangerous or offensive.

Hutto's account is also not particularly well equipped to account for the emotions' being subject to rational norms. If emotions are not even representations involving adequacy conditions, it is hard to see how they could stand in rational relations to us, the world, and other mental states. For Hutto, the criterion for calling something a representation is that it has the function of indicating things as being thus and so in a larger cognitive architecture. This definition leaves room for interpretation: when precisely is it the case that a representation has such a function and how large does this cognitive architecture have to be? Emotions are often preceded and accompanied by all kinds of mental states; they can be directly influenced by them, and directly influence them. What matters most is that emotions appear to stand in relations to other emotions and other mental states. For example, if I am afraid that somebody might delete a document that I only saved on a public computer, I will be relieved if it is still there the next day and sad or angry (depending on further circumstances) if it is not. Does that not count as a sign that emotions, apart from their main function to prepare for action, can also function to indicate to other mental states that things are thus and so? Finally, since, according to my argument, Hutto cannot account for emotions having an intentional object, I take for granted that he cannot account for emotions having social rules and norms as their intentional objects. Radical enactivism, if understood "Hutto style," is not able to account for a single feature of the normative structure of emotions.

c. The Autopoietic Enactivist Approach

Colombetti (2014) develops a comprehensive approach that grounds the enactive model in the affective sciences and convincingly argues that an enactive approach to emotions also comes with a solid empirical foundation from various sources. Colombetti relies on the framework of autopoietic enactivism (Di Paolo 2005, Thompson 2007, Di Paolo and Thompson 2014), which broadens the notion of cognition to include all kinds of coordinated interactions that organisms perform in their environment. Such structured interactions are called *sense-making*, which, according to enactivist approaches, is the mark of the mental and an activity that all living organisms exhibit as autonomous and adaptive systems.

Autonomous systems are inherently purposeful in the sense that they generate ends or purposes within themselves in order to maintain themselves. A system that is autonomous and adapted to its particular environment brings forth a structured environment. Such an environment entails features

of different valence for the maintenance of the organism. It monitors itself regarding its maintenance conditions within its environment and thereby develops a graded scale of values that structure the environment: some substances are good for nutrition, some are even better, and some are not so good or are to be avoided entirely (Di Paolo 2005, Colombetti 2010, Di Paolo and Thompson 2014).

The very idea of sense-making is that organisms, even simple cells, transform the world around them into a meaningful place simply by interacting with that environment in order to regulate their self-maintenance. The ascription of meaning and value on the level of simple cells rests on the claim that all living organisms are purposeful systems owing to the sheer fact of their struggle for survival. As such, all creatures need to care about the world around them and actively draw distinctions:

> Immanent purposefulness and care are required to discriminate what matters to the organism; at the same time, the world takes on significance and value precisely in relation to what the organism is concerned about and striving for—there is no meaningful environment for the indifferent, nonmotivated being.
>
> (Colombetti 2014, 19)

Colombetti labels this most basic level of "caring" as *primordial affectivity*. This basic capacity grounds all further and more complex forms of sense-making, including emotions. In this framework emotions are understood as responses of the whole organism that enact meaning in the structured interaction with their surroundings.

According to the enactivist understanding of cognitive processes all emotions are cognitive. It can therefore be misleading to label Colombetti's account as a "noncognitivist" approach. Yet I introduced cognitivism as an approach that takes emotions to essentially involve judgments, be essentially caused by judgments, or essentially entail representations. All three claims are rejected by Colombetti, and she argues in detail against even current modest versions of cognitivist appraisal theories like the one defended by Klaus Scherer (see Colombetti 2014, Chapter 4).

Colombetti agrees with Prinz and Hutto that emotions are embodied appraisals and that the bodily reactions involved in an emotion constitute a feeling and a simple form of intentionality. Colombetti and Hutto both criticize Prinz's account for assuming that the vehicles of emotional content are wholly embrained. This is a view enforced by traditional brain-centered theories of representation, which take neural states to be locatable vehicles that carry a certain kind of content just as letters in a book do. Enactivists of all camps agree that emotional episodes are responses by the whole organism. The main problem for traditional representationalism is that it simply assumes that intentionality is a relation between a vehicle in the brain and content in the world.

Relying on a broad range of empirical evidence, Colombetti suggests that there are coordinated patterns, which include neural activity, bodily reactions, and expressive behavior that are involved in emotions. She points out that there is no *prima facie* reason to think that these reactions must be perceived in the brain first and that only the perception of them can be a meaningful representation of the urgent situation. Instead, Colombetti argues that in emotions the nonneural body similarly forms part of the vehicle, so that the whole organism comes to fulfill a meaning-generating and consuming role. The background assumption is that of a living organism that can generate emotional episodes without the pervasive guidance of appraisals. Colombetti instead focuses on the organism as a highly complex and interconnected system in which changes in one part typically modify the rest of the system. Emotions, according to this view, are dynamic patterns without a fixed sequence of input, cognition, and output.

Emotional expressions, for example, should not be seen as the output of internal pregiven instructions but as inputs that trigger an emotional reaction or the outcome of synergistic processes of mutual constraints among muscles. The standard picture of someone getting scared when encountering a barking dog (input) and then showing certain bodily responses including a facial expression (output) is one-sided. Whether somebody is prone to becoming scared in a certain situation might depend on more than the cognitive evaluation of the triggering stimulus. Bodily reactions can also come to be causes themselves or might influence the overall evaluation of a situation. For example, someone in a tensed bodily posture and with a short and shallow breathing cycle might be scared more easily than she would have been under different circumstances. Colombetti departs from Prinz here, since she does not assume any kind of "calibration file" but, rather, claims that emotions are entrained patterns that can be triggered by each element they involve (see Chapter 5). Colombetti, furthermore, does not claim that the bodily reactions that constitute emotions have to be "perceived" by a further inner mental state. According to Prinz, the perception of inner arousal could be labeled as the essence of an emotion. According to Colombetti, there is no such thing as an inner essence of an emotion. An emotion is a dynamical pattern that unfolds between an organism and the structured environment and is concerned with the organism's well-being.

A third point on which Colombetti departs from Prinz is the distinction between so-called higher-cognitive and basic emotions. She cites various empirical sources that argue against a strict distinction between the two, such as neuroscientific studies that suggest that the parts of the brain involved in higher cognitive emotions are not distinct and separate from those involved in basic emotions. Colombetti suggests completely abandoning that distinction and instead regarding all emotions as complex dynamic patterns of brain and bodily events.

Like Hutto, Colombetti is rather skeptical about Prinz's commitment to relational properties and, thereby, to objective realism. Both Hutto and

Colombetti tend to think that it is not necessary to commit oneself to relational properties that are wholly objective and mind-independent. As Colombetti puts it, "[f]rom an enactive perspective . . . danger, loss, and other core relational themes . . . are not represented or detected but enacted, namely, brought forth in the world-organism encounter—in other words, they are response-dependent" (Colombetti 2014, 110–111). Both Hutto and Colombetti note their preference for response-dependent properties as a minor point in passing. Yet, as I have already noted with regard to Hutto, the ontological commitment to relational properties makes a central theoretical difference and leads to very different possibilities when it comes to explaining the normative structure of emotions. I discuss the particular difficulties that Colombetti has in accounting for DEE and SEE in the following.

For autopoietic enactivism, valence is a central notion. As pointed out earlier, Colombetti sees valence as a feature of the environment that is created by the interaction of an organism with its environment. Due to the sense-making activities of an organism, even the organism's simplest responses to its environment are valent. Organisms differentiate between things in the world that should be approached because they serve the purpose of self-maintenance and things that should be avoided. According to this theory, emotions are just another kind of response to valent features, and the properties they respond to are themselves response-dependent. Their existence not only depends on the existence of the organism but is also brought about by its activities.

It is at that point that enactivist approaches depart from the traditional definition of response-dependent properties. Response-dependent properties are usually seen as owing their identity to the way the organism *represents* the world. Funny things have nothing in common except for the fact that we represent them as funny. Since enactivists do not take emotions to be representations at all, they ought to describe response-dependent properties as properties brought into existence by the interaction of an organism with the world rather than by its way of representing the world. With regard to the value properties to which emotions respond, the latter account could be described as a kind of bioconstructivism in that it claims that biological values owe their identity entirely to us and to the way in which we draw distinctions in our interaction with the world for the purpose of self-maintenance.

Of course, one could object that enactivism need not claim that biological values are entirely made up. Certainly the way we enact values has to respond to actual features in the world that make it adaptive to represent the features in question as valuable. Yet such an approach would need account for DEE. DEE claims that emotional representations are responses to adaptive pressures. I fail to see how a response that by being enacted creates a value could be the result of an adaptive pressure, where the situation that creates the adaptive pressure could be described adequately in value-free terms. The very notion of an adaptive-pressure seems to imply that the situation in question is somehow bad for the organism and that therefore it

would be good if the organism would develop a certain disposition to deal with the situation in a way that is fitness-increasing. Therefore, enactivism needs to introduce relational properties to account for DEE in the way I develop in Chapter 4.

Bioconstructivism has further problems. Cognitivists have suggested that different types of emotions can be identified in terms of their contents. Anger represents an offense; fear, a danger; and sadness, a loss. Even Prinz, who claims that different emotion types are underlain by different patterns of bodily arousal, considers the different intentional objects to play a role in the individuation of emotion types as well. Emotions that appear to involve similar, if not identical, bodily reactions, such as embarrassment and shame, can be differentiated with regard to the properties they represent. Yet, if the properties to which emotions respond come to exist through the organisms' different ways of responding to the environment, then it is hard to see what the structuring role of the environment is in the individuation of emotions.

Colombetti does not take emotions to be nonrepresentational in the same straightforward sense as Hutto. Following Thompson (2007), she rejects the straightforward distinction between a symbol-producing inner world and an external world that can come to be adequately depicted by the inner symbols. Yet autopoietic enactivism does not deny that the nervous system does exhibit patterns of activity that reliably recur in particular contexts and that these embody meaning for the organism. But such patterns should rather be treated as endogenously created responses triggered by sensory perturbation than as representations. Colombetti does not want to exclude talk about representations in principle and would allow the notion if it referred to the enactive patterns just described (see Colombetti 2014, 214).

Does her approach, therefore, face the same objections with regard to semantic norms and intentionality? The decisive difference between Colombetti and Hutto is that Colombetti takes emotions to be processes of sense-making that enact meaning, while Hutto rejects such a definition. The question of whether such a process of sense-making gives a satisfying explanation of the way distal intentional objects, such as danger, is one that at best can be left open but certainly cannot be answered with respect to current research in dynamical systems theory. Chemero (2009) convincingly argues that there are dynamic explanations that make representational accounts superfluous. He uses the example of artificially evolved robots that are able to seek and reach a target. Sussex roboticists explain these robots and their environment as coupled dynamical systems. To explain their operation they give mathematical descriptions of the structure of each separately, and then, based on those, they give a unified account of the coupled system. Once we have the full dynamical story, we can predict the behavior of the robot in its environment completely. We can then still give an explanation of the robots behavior by using a representationalist framework, but saying that the robot moved in a certain direction because it was representing a certain object as being there and wanting to get close to it, does not really add

any interesting information to the story already given by the mathematical explanation of the robots visual field and its moving abilities. Dynamic approaches on emotions such as Colombetti's, however, are far from issuing any clear predictive results about what emotions we would have in certain situations. As long as there are no such results (and given the complexity of organisms like us, one can doubt whether we will ever have them) the representationalist framework still does a lot of explanatory work that cannot be replaced by a dynamicist approach.

What Colombetti would therefore need to clarify is how our ordinary talk of emotions being adequate or inadequate can be explained in her framework. If emotions are not representations, they do not have adequacy conditions in a semantic sense. But what does it mean, then, to criticize someone for excessive anger or groundless jealousy? If emotions are conceptualized as reactions that enact an environment by reacting to certain events and not to others, then the adequacy of a reaction can be traced back only to the reacting organism itself. It is the organism that creates the environment as a place that entails meaning and values in the first place, so the criteria for assessing the adequacy of a reaction can only be found within it. The tension between this claim and our ordinary way of talking about emotions reveals a systematic problem with this approach. The adequacy of emotions is not a phenomenon that can be traced back to the organism's perspective on the world alone. The norms according to which fear or jealousy can be adequate or inadequate are biological and social norms that the organism encounters in the environment. I argue for this position in detail in the next chapter.

Furthermore, it is hard to account for the emotions' being subject to rational norms, if they are not representations in the first place. How are we supposed to understand the role of emotions in a larger cognitive architecture if not in terms of their being representational? The arguments against Colombetti's approach are here very similar to those against Hutto's (see the earlier discussion). How can we understand that it is not rational to envy oneself when describing emotions as dynamical patterns? How can we understand that it is rational for me to be afraid that my wallet was stolen, insofar as this puts me into danger, that it is rational to turn angry, when I come to believe that it was a friend of mine who stole the wallet and guilty when I then find out that I lost the wallet myself and blamed my friend for no reason? There is something fundamental about the way emotions relate to each other and to other mental states that behaviorist or dynamicist approaches haven't been able to capture so far and that might only be capturable in terms of representations that can stand in rational relations.

The autopoietic approach has an explanatory problem with regard to social norms as well. Colombetti claims that the distinction between basic emotions and higher cognitive emotions should be abandoned since there are no significant differences between the two (Colombetti 2014, 26ff.). Yet even if one agrees entirely with her evaluation of the empirical data (which

I do), what remains to be explained is the distinction between emotions that respond to biological values and those that respond to social rules and norms. As I point out in Chapter 1, many argue that those emotions that respond to social rules and norms presuppose complex cognitive abilities such as explicit conceptual knowledge (e.g., Lewis 2014). If emotions are about response-dependent properties, then how the responses in question could be about social rules and norms without being cognitive responses needs to be explained. I take it that an embodied account can make sense of the idea that social emotions are nonconceptual embodied action-oriented representations. But this requires taking synchronic and diachronic environmental externalism seriously and developing the claim that social rules and norms are indeed given as relational properties with which organisms learn to interact from very early on. This is what I do in the next chapter.

Many current approaches, including autopoietic enactivism, point out that emotions are situated in a particular environment and argue that the emotions' unfolding in a particular situation as well as their development is, to a large degree, scaffolded by the biological and social environment (Krueger 2014, Slaby 2014, Krueger and Colombetti forthcoming). However, I think that these approaches lack some important ontological ingredients to count as diachronic and synchronic environmental externalists. If the intentional objects of emotions are assumed to be properties that highlight a situation as biologically or socially relevant, then diachronic environmental externalists need to assume that such properties exist and that organisms make contact with them again and again until they develop an adequate response. Synchronic environmental externalists need to assume that the relevant properties exist in the world, so that in the usual case we do not need complex inner representations of them to be able to respond to them. To say that "the world is its own best model" is only an option if the relevant parts of the world are part of the ontology in question. Response-dependent properties in both cases do not do the job. They cannot be what caused our ancestors to develop representational or meaningful responses, since their existence depends on the prior existence of these representational or meaningful responses. Furthermore, they cannot be part of an antirepresentationalist account that claims that "the world is its own best model," since their existence depends on there being certain representational abilities. Response-dependent properties by definition cannot explain what is called "cognitive offloading" in the debate on embedded cognition. That we can replace an inner cognitive structure with the interaction between a skillful body and the environment to which the organism is adapted, presupposes that we describe the environment in an ontologically adequate way; only what stands in a certain kind of causal or informational relation to the organism can be part of the world as its own best model. Everything else necessarily assumes an inner model of the feature in question.

5. WHY REPRESENTATIONS AT ALL?

The present account claims that emotions are representations, cognitively simple ones that are realized through patterns of bodily arousal, are highly dependent on the context in which they occur, directly motivate for action, and have a nonconceptual content. But still, they are representations, and as such, they are subject to adequacy conditions. In responding to enactivist accounts, I discussed some reasons for assuming that emotions are representations. However, since the debate on representations is hotly contested (see, e.g., Clark 1997, Gallagher 2008b, Hutto and Myin 2013, De Bruin and Kästner 2012, Ciaunica 2014), I conclude this chapter with a detailed discussion about the kind of representations I take emotions to be.

This is particularly important because I argue in Chapter 5 that emotions are about affordances. Gibson (1979) coined the notion "affordance" to capture the idea that organisms do not perceive simple objects or states of affairs but, rather, an environment that affords all kinds of actions. Gibson, and many Gibsonians generally, not only claim that direct perception can guide action without any inferential processes between but usually also deny that there are *any* representations or even inner information processing involved in affordance perception.[11] Therefore, I clarify in what follows what I mean in saying that emotions represent affordances.

The radical Gibsonian claim is that certain variants and invariants in the environment impinging on an active organism can carry information about the relations of significant distal affairs and guide motions to make use of these affairs. This implies a radical externalization of information: not only can a tree be directly perceived as a climbing affordance, but the perception also plays a causal role in allowing successful climbing to occur. The squirrel not only perceives that the tree affords climbing, but also perceives at the same time how to do so; it can perceive what specific dynamic sequence of his own bodily movements could be successful. This information is not stored in the form of complex representations in the squirrel's memory; instead, the tree strikes the squirrel as a climbing affordance with specific features that can be used for ascent. Yet the affordance "striking" the squirrel does not include any form of inner information processing or representation—the Gibsonian is instead committed to the claim that natural information occurs in the world.

The general antirepresentationalist view behind this picture assumes that representations are internal images or signs that "stand between" the organism and the environment. Representations re-present the preexisting world in an internal image that is a discrete identifiable entity, bears an internal content, requires interpretation, is passive, is detachable from its context, and can be used in inferences. While this strong notion of internal representation certainly does not fit with any kind of embodied view, I disagree with Gibson and several current proponents of enactivism and radical embodied

accounts, which argue that the concepts of internal information processing and representation should be abandoned entirely. Information, picked up by the organism reacting to it, involves some systematic alterations of inner states and such alterations occur in a patterned form that developed or evolved in the history of interactions between the organism, or its ancestors, and the environment. This patterned reaction of the organism can be seen as the vehicle of the content. Representations are not realized in the head alone and they need not be seen as anything like concrete entities in the brain. Yet it is misguided to deny that there is information processing going on in the brain and that such processing is involved in eliciting an emotion, in its becoming conscious, and in further processes, such as increased attention and active memory in fear.

What would also be misguided, however, would be to simply identify such information processing with an emotion. Emotions are constituted by reactions of the whole organism that make sense in certain scenarios as they prepare for interaction with them. Emotions cannot be individuated appropriately without reference to what they represent; their being normatively assessable can only be captured on the task-description level, where the whole organism interacts with an environment in ways that are appropriate or inappropriate. Neither a taxonomy of emotions nor their normative assessability can be explained on a mere neuroscientific level of description. Yet to make this claim, it is not necessary to deny that the brain processes information. It does so all the time. Therefore, I do not follow Gibson's suggestion that the concept of inner information processing should be denied, in emotions or in general. Furthermore, it is not reasonable to abandon the concept of representation altogether. Emotions are representations with a distal content and what they represent can be appropriate or inappropriate. They are subject to semantic norms. Yet emotional representations certainly do not meet the characteristics of representations mentioned above.

Emotional representations have an external content, namely, core relational themes, that is, certain relational properties that afford something to the organism. These properties are represented in a nonconceptual format. The vehicles carrying emotional content are not single states in the brain; they are patterned reactions of the whole organism that evolved incrementally and need not be centrally governed or elicited in the brain. These patterns evolved because they serve certain functions for the organism, namely, detecting external core relational themes and directly reacting to them. These whole-body reactions to concrete external situations fulfill none of the criteria mentioned above; they are not discrete identifiable entities, their content is not internal, they do not require interpretation, they directly motivate for action, they only occur in certain contexts, and they are not formatted to be processed in inferences.

So why, then, should such patterned alterations of inner states be called representations? A first answer lies in the anticipatory moment involved in affordance-perception: the squirrel perceiving the tree as a climbing

affordance is anticipating a future action. Such anticipation, one could argue, cannot be explained with respect to the external properties of the tree alone. That the squirrel can represent the tree as an affordance depends not only on the present features of the tree and on the present features of the squirrel but also on the history that the squirrel or its ancestors had with similar trees. It is this history that established a set of coordinated neural and bodily reactions that trigger certain behavioral patterns as a response to certain stimuli. A concept of representation that makes sense of such anticipations need not include the claim that representations are inner identifiable items that are passive and detachable from their current context. The squirrel is not able to think about the tree being climbable while sitting at home and wondering about tomorrow's meal. Therefore, affordance perceptions are best described as action-oriented, egocentric representations that are relative to the agent, context-dependent, and realized through vehicles that are spread over the brain and the body.[12] The squirrel's response to the tree is shaped by past experiences, which are responsible for the anticipatory moment in the squirrel's perception.

I develop the view that emotions represent affordances such as a-danger-to-be-avoided or as a restriction-to-be-fought in Chapter 5. What should be clear from the discussion of enactivist approaches above is that there are further reasons to assume that emotions are representations. Emotions have many functions in a larger cognitive architecture. They can be triggered not only by the direct perception of a situation but also by all other kinds of mental states. Emotions seem to hold a kind of in-between position between perceptions that cannot be detached from the situations in which they occur at all and abstract thoughts that can occur independently of the present situations. Emotions are nondetachable in infants and animals, and they are not under our control in the way that we are able (at least sometimes) to produce thoughts while sitting in the proverbial armchair. Yet they can occur while remembering or imagining something, which suggests that they can refer to a core relational theme in its absence. Antirepresentationalists cannot account for this feature, although it certainly is a feature that is essential for emotions.

6. CONCLUSION

I have argued that the analogy between emotions and perceptions is not helpful when it comes to explaining emotions' being motivating and valent, and the role that bodily reactions play in emotions. I have suggested, however, that we follow perceptualists who claim that emotions do have a nonconceptual content. A description of this content from a phenomenological perspective sheds light on the role that bodily reactions play in the constitution of emotions. I have furthermore discussed current embodied and enactivist accounts that share the claim that emotions' intentionality is

constituted by bodily arousal. The result is that none of these approaches can account for the normative structure of emotions.

The first step in developing a theory that meets the normative challenge is to find an adequate conception of emotions as representational. I have suggested that emotions are embodied action-oriented representations, or representations that are constituted by bodily arousal and represent affordances. In the following chapter, I develop the ontology that allows us to account for DEE and SEE by introducing an appropriate notion of natural information and of relational properties. This sets the stage for the theory of emotional affordances that I develop in Chapter 5 and enables me to account for the normative structure of emotions without assuming that emotions entail conceptual content.

NOTES

1 For extensive discussions of the objections against perceptual accounts, see Deonna and Teroni (2012, Chapter 6) and Salmela (2011).
2 This claim can be found, for example, in Solomon (2003) and Nussbaum (2001). See also Chapter 2.
3 See Kreibig (2010, 405) for an overview of studies on the patterned reactions of the nervous system in sadness.
4 Apart from Frijda, so-called transactional or situated approaches to emotions have recently highlighted the way in which emotional content is action-driven as well. See, for example, Griffiths and Scarantino (2009) and Parkinson (1996).
5 According to recent studies, strongly motivating emotions are probably those that involve the strong arousal of the autonomous nervous system (Laird and Lacasse 2014).
6 Kreibig (2010) reports such slight changes in the autonomic nervous system, and studies on typical pride-related postures can be found in Tracy and Robbins (2004).
7 Such a claim is backed up by neuroscientific studies that show that during reactions of fear the amygdala can alter attention and awareness by modulating the hippocampal memory system (Phelps 2004).
8 I develop this objection in Hufendiek (2012).
9 When it comes to the precise content of an emotional representation, Prinz refers to Millikan, rather than to Dretske, since Millikan defended the claim that even the simplest representations have some distal content since they are intensional in a certain sense that is defined by the consumer function (see Millikan 1993).
10 To be precise, Hutto mentions in passing that it is not necessary to follow Prinz with regard to the commitment to relational properties and implies that he could develop an account without taking core relational themes to objectively exist (Hutto 2012, 178).
11 See, for example, Chemero (2009) for a current Gibsonian approach that takes perception to be direct in the sense that it (1) doesn't involve any inferences, (2) doesn't involve any representations, (3) picks up natural information, and (4) directly guides behavior.
12 These criteria are inspired by the concept of action-oriented representations, as it is defended by Michael Wheeler (2005) and Andy Clark (1997). For critical discussion see Gallagher (2008b).

4 Embedded Emotions and the Ontology of Core Relational Themes

1. INTRODUCTION

An embedded theory of emotions needs to take the structuring role of the environment into account. An emotion is embedded wherever the environment plays a constitutive role in emotional processing such that it is the interaction with the structured environment, rather than internal symbol processing, that explains the intelligence of an emotional reaction. The discussion of embedded, situated, or scaffolded emotions has been particularly heated in recent years and has raised questions about the typical settings in which emotions occur, the effects of the environment on the development of emotions, and the typical triggers and structuring features that occur in the environment. Griffiths and Scarantino (2009) point out that most emotions occur in a socially and culturally structured environment. Our social relations toward others and the cultural artifacts we are surrounded by play a diachronic role in the development of emotions and a synchronic role in the unfolding of emotional sequences. Furthermore, Sterelny (2010) points out that scaffolding of cognition in general has three dimensions, and Colombetti and Krueger (forthcoming) adapt these to emotions.

While these accounts are helpful to clarify the particular role of certain environmental features in emotional processing, most of the discussed examples concern how the environment takes part in triggering, realizing, or regulating an emotional episode. The current debate thereby tends to lose sight of the external world as a store of information that enables intelligent interaction with the world without the need for complex internal representations. I argued in the last chapter that this causes trouble in current embodied accounts, since to explain the normative assessability of emotions we need to assume either a complex inner evaluative apparatus or a complex interplay between skillful body and structured environment, where the environment serves as its own best model. To make sense of this idea we need to say something about the skillful abilities of the body, but we also need to make certain ontological claims with regard to the organism's environment. I labeled these claims diachronic environmental externalism (DEE), that is, the claim that if representations do have a function, then there must be something in the world that was there before the representation evolved

to make it useful to have such a representation in the first place, and synchronic environmental externalism (SEE), that is, the claim that the world is out there, structured and present, so we can rely on the world itself in cognitive processes instead of making up complex internal models. I also suggested that embodied accounts of emotions need to commit themselves to a kind of normative realism to account for DEE and SEE, yet no current approach is willing to make such a commitment.

In this chapter I develop a kind of normative realism that is able to account for DEE and SEE. I distinguish normative realism from moral realism and explain why I am defending the former but not the latter with regard to emotions. As argued in Chapter 1, emotions are about things that are of value for us in a biological and/or social sense, but I suggest that emotions are not particularly set up to respond to moral values. An embedded account of emotions needs a story about the structure of the environment in which the values that emotions are about are situated, and in which we encounter them again and again. The goal is to develop an account of natural information, relational properties, and affordances not only as a meaningful structure of the environment but also as a structure that is of value for the organism situated into it.

After having distinguished normative and moral realism, I develop an account of natural information that allows us to see an organism's environment as a space that transports all kinds of emotion-relevant information through all sensory channels. Since the distal contents of emotions are relational properties, in a second step, I translate some core relational themes into relational properties that occur in the biological environment. Yet this does not give us a comprehensive explanation of the emotions' intentional objects, since only some emotions are primarily concerned with biological values while others appear to be about social norms. Before explaining, in a final step, how to understand core relational themes as relational properties that occur in the social world, I discuss the traditional understanding of the emotions in question. Emotions with a social object, such as guilt and shame, are commonly understood as emotions that can only be entertained by beings with an explicit understanding of the self and social rules and norms (Lewis 2014). I offer an alternative understanding of the emotions in question and reformulate their core relational themes. Once an appropriate understanding of what has to be explained is established, I develop an account of core relational themes as properties that steadily occur in our social environment.

2. WHY MOORE'S QUESTION IS STILL OPEN

The normative realism I am going to defend in the following is not to be confused with what is usually understood as "*moral* realism", that is, the position that moral judgments are about moral facts and that those

judgments that get the facts right are true.[1] In what follows, I am neither concerned with moral judgments nor with moral facts. I am concerned with the biological and social environment of organisms that is stuffed with things that are of value to them with regard to biological and social norms. The classical worry raised by Moore (1903) against moral realism is that it is not compatible with naturalism. The idea is that being a naturalist about morality implies thinking that moral terms can be defined correctly using terms that refer to natural properties. *Good* might be defined as "pleasant," and pleasure understood to be a natural property. Moore's objection against such approaches is that they leave the question of whether something is really good or bad open, since a mere description of natural facts can never determine whether something is good or bad. I develop an ontological approach that captures the normative dimension of emotions in a naturalist framework by claiming that things can be good and bad in relation to us with regard to biological and social (but not moral) norms. As stated in Chapter 1, I take social norms to be conventional norms that are not context-independent, unconditional, and obligatory, which is how moral norms are usually defined.

According to my view naturalist approaches are not restricted to mere descriptions of natural facts. Introducing biological and social norms to explain what emotions are about is compatible with naturalism. Still, such an approach does not intend to answer Moore's question in any satisfying sense, since the question of whether the values emotions refer to are good or bad in any moral, unconditional, intrinsic, noninstrumental or obligatory sense will be left open.

My claim, therefore, is not that emotions refer to moral values and my aim is not to defend the view that emotions play a certain epistemic role in grasping moral values. But one might wonder nevertheless, whether it would be an option at all, to establish a metaethical approach on the grounds of the present theory. Couldn't one argue in a sentimentalist manner that our ability to respond to moral values or that our understanding of the notions of good and bad are constituted by our emotions that refer to historically established biological and social values? The result could be a sentimentalist approach similar to those of Hume and Prinz with regard to the psychological mechanisms that constitute moral judgments but different from Hume and Prinz, where it comes to the ontological status of moral values. Where Hume and Prinz remain subjectivists about moral values, a metaethical approach that takes emotions to be about affordances in the sense I develop in the next chapter could ground a naturalist realist position. Only it's not that easy.

The present theory certainly offers a good starting point to argue that emotions (for contingent reasons) play a constitutive role in our development of social abilities on a phylogenetic and on an ontogenetic level. Emotions most likely played a central role in our becoming beings that are able to care about others and approve and disapprove of things in social contexts. But

these social abilities aren't moral abilities insofar as they respond to con-ditioned biological and social norms that contingently developed because they happen to serve a certain function. A sentimentalist view building on my view of emotions—if it would not want to reduce moral values to the biological and social values emotions are about, would have to do a lot of work to explain how moral values are related to the biological and social values emotions are about.

According to my view emotions are set up to respond in a quick and cognitively restrained way to situations that are of relevance with regard to bodily well-being and with regard to social norms. The difference between social and moral norms might be blurry in our everyday talk about e.g. who should and shouldn't feel guilty. Yet whenever we look at social scenarios where guilt reactions are established the different, for example, guilt reac-tions often respond to the violation of moral and not just social norms, I don't think that this is what emotions are set up for and also not what emotions are particularly good at, at least if moral norms are understood to be obligatory, context-independent and nonconventional.

Therefore, whatever a metaethical approach grounded on the present theory of emotions would look like, it would be very different from those metaethic approaches claiming that emotions constitute a moral faculty through which we acquire moral knowledge (e.g., Roeser 2010). More than anything, I defend a very different theory of what emotions are and what they are supposed to do than these authors. I take emotions to be the result of an evolutionary and social process of tinkering. They are clumsy skills that orient us in a biological and social environment not an intuitive fac-ulty that enables us to grasp abstract moral properties. I develop a realist position with regard to biological and social but not moral norms in the following.

3. NATURAL INFORMATION

Core relational themes such as "being dangerous" or "being indigestible" might be relational properties, yet this alone doesn't explain how we can grasp them via nonconceptual embodied action-oriented representations. Core relational themes do not strike us "directly" via one input channel, as the textures of surfaces strike us directly via touch or loud noises strike us directly via hearing. We perceive many things that do not strike us directly in this sense, such as velocity and distances. As many would argue, we do not need to apply conceptual reasoning to sensory data to understand that something is too far away to grasp or that something is approaching quickly and we should step aside (Evans 1982, Cussins 2003). The claim that such abstract features of the environment can be perceived "directly," that is, without any kind of inferential reasoning taking place, has been defended in the context of ecological psychology.[2] Teleosemantic approaches have

furthermore developed the hypothesis that perception can be seen as an evolutionarily established ability that involves representations but not conceptual reasoning. Both accounts rely on the claim that "in the beginning there was information" (Dretske 1981, vii), which is partly what DEE is about. Part of the assumption required to explain how organisms came to acquire simple representational systems is that there is information in the world that representational systems can come to pick up. I discuss in the following what an appropriate notion of natural information is with regard to emotions.

I introduced the notion of natural information in the context of teleosemantic approaches to emotions (Chapter 3). Yet the theory of natural information that Dretske and Prinz embrace has been criticized and further specified in the context of Millikan's (2004) biosemantics. I take it that an embedded account of emotions relies on Millikan's specified notion of natural information. Dretske's original definition relies on the Shannon–Weaver theory and suggests that a signal carries information about its source only if nomic relations determine the relationship between the signal's occurrence and its source. "The transmission of information requires, not simply a set of de facto correlations, but a network of nomic dependencies between the condition at the source and the properties of the signal" (Dretske 1981, 77).

One might wonder what precisely Dretske has in mind when talking about "nomic correlations." One way to understand nomic correlations is in terms of natural laws. But if the connections between signal and source are established by natural laws, then one might wonder whether emotions could be described as detectors that pick up natural information at all. As I argued in the introduction, biological processes and the regularities that structure an ecological niche cannot reasonably be understood as universal laws that hold between properties, independent of the time and place of occurrence. We need to think of a biological niche as a place that is filled with natural information transported through light, sound, vibrations, and so on, where most of the information recurs for a reason in this special niche, but only a minor part of this information can be described as being established through natural laws that hold between signal and source. There are, for example, no natural laws about individuals, because natural laws are supposed to be general. Individuals are of course subject to natural laws. The fact that bodies on earth have weight is explained with reference to the law of gravity. Yet the fact that a caregiver has a certain weight and will fall, when being pushed down a cliff doesn't offer any interesting information about the caregiver as an individual but only about the caregiver as a body. Therefore, if natural information is only carried by nomic relations, no infant could ever recognize its caregiver by exploiting natural information[3] but only as a body or a member of the human species. Alternatively it could be argued that the recognition of the caregiver as an individual does not depend on the direct perception of natural information alone but on further cognitive processing (Millikan 2004, Chemero 2009). Both consequences

are rather unsatisfying, since infants are able to distinguish their caregivers from other people using information that is carried by the caregiver's voice or smell from early on.

A second problem is that if natural information is restricted to information established through natural laws, no information can be carried by conventions. Conventions, by definition, hold because of public agreement, not because of natural laws. Our ecological niche is fundamentally social and, therefore, full of information that relies on conventions and social practices that recur reliably enough to successfully guide behavior (Chemero 2009). A song can be a locally recurrent natural sign for an infant that it is time to sleep; getting dressed up warmly might be a sign for the infant that it will be taken for a walk. Again, saying that there is no natural information transmitted here, or that the meaning of the situation can only be known by inferential reasoning, is rather unsatisfying.

It is an issue of debate what precisely Dretske means by "nomic relations" and how strict the probability laws he assumes have to be.[4] In the context of ecological approaches and Millikan's biosemantics, natural information is explicitly understood in a broad sense and thereby avoids the problems mentioned above. Millikan agrees with Dretske that while representations or intentional signs can be false or inadequate natural information cannot. False representations simply do not detect or carry natural information. But, according to Millikan, nomological relations are neither the only nor the most typical sources that carry natural information. Natural information can be carried by a nonaccidental connection between, for example, trails in the wood and the animal that left them or between red spots on the face and the virus that caused them. We don't need to refer to natural laws or nomic necessity as the only possible source for natural information and thereby for true representation.[5] Certain kinds of natural information steadily recur in our environments in such a way that we can grasp them and make sense of them or be guided in our behavior by them (Millikan 2004). A sign can carry "local information" whenever it correlates with some environmental feature *for a single reason*. Nomological correlations can be among such correlations but need not be and are certainly not the kinds of correlations that evolution typically builds on. The correlation in question need not necessarily hold everywhere in the same way but must hold at least in the local domain where the animal makes use of the sign. "For a system to acquire the capacity to use a natural sign either by means of natural selection or by means of learning, the sign would need to be one that recurs, and recurs with the same natural signification or meaning" (Millikan 2004, 37). With this account, Millikan aims to define natural information as something that must be appropriate for the way in which actual organisms make use of it. It must be of a form that can be detected by organisms, and it must be about the concrete and relevant circumstances of the organism and not about abstract possibilities: "The kind of knowledge that earthly creatures have is knowledge applicable in the domains they inhabit, not knowledge

for arbitrary nomically possible worlds, nor for other domains, regions or eras within the actual world" (Millikan 2004, 44).

One may wonder, then, whether there are any criteria to distinguish correlations that are sufficiently "strong" to transfer information from correlations that are too random or too vague to be of any use. Yet the question should always be whether the correlations to be found in an animal's environment are strong enough to have an actual influence on sign use either through selection or through learning. Animals use many signs that are ambivalent in their domain: a certain smell might indicate various other animals including predators but also harmless creatures. Animals might use this kind of smell as a sign for predators although they are in principle not able to distinguish predators from other animals by this sign. As Millikan argues, our abilities to represent things could not depend on natural information specified in the way Dretske suggests:

> If the capacity of an organism to represent something mentally were to depend on its ability to discriminate that thing from all others in accordance merely with natural law and logical necessity, it is clear that no organism could possibly represent anything distal.
>
> (Millikan 2004, 34)

Millikan takes natural information to be a fundamentally epistemic notion; to explain it properly it must be considered what it has to be like for the animal to be able to grasp and exploit it. No biological representation system, however, has the ability to discriminate something from all other things with regard to both natural laws and logical necessity.

Millikan's definition of natural information explicitly sees the local environments of animals as rich sources of natural information. Individuals can be recognized directly by voice, sight, or smell, and conventions can transfer natural information as well. Also, animals can receive natural information and successfully use it to guide behavior, although the information might be ambiguous because the animal could be principally unable to distinguish, for example, between two species as a possible source of information.

What role does natural information play in grasping core relational themes? Take fear, for example. In our local environments dangerous situations recur for certain reasons, as does plenty of natural information that has the dangerous situation as its source. From birth on we are disposed to react with fear when hearing an unexpected loud noise or losing our balance, for example. Since hearing loud noises and losing one's balance appear to be unconditioned stimuli for fear, the respective auditory, visual, and proprioceptive stimuli carry natural information that we are sensitive to from birth. This shows that fear can obviously be elicited through information-carrying stimuli from different perceptual channels: auditory and proprioceptive stimuli can both cause fear reactions. While there appear to be some hardwired stimuli, it must be highlighted that there are very

few of them. Fear is a highly plastic reaction that can come to be triggered by all kinds of stimuli, once we learn that something in a given situation is associated with danger. Animals that once met a predator often react with fear when one of the predator features occurs again; they can remember the predator's smell or the rhythm of the vibration of its feet on the ground or the sound of it flying through the air, if they once learned via association that this kind of natural information is a sign of a predator. In a similar fashion, infants can quickly learn what recurring signs for danger in their environment are. Realizing that the caregiver is not present is a natural sign of a dangerous situation, as is the facial expression that indicates that the caregiver is about to lose her temper.

These examples also highlight something about the character of the natural information relevant for emotions. Although all kinds of stimuli can come to be triggers for fear or anger, there are certain kinds of natural information that are far more typical than others and play a structuring role in how infants learn emotional behavior. Most of the information that recurs steadily in the infant's environment and is associated with emotions is of a social nature: all forms in which we express our attitudes, feelings, and opinions about others to them constitute a kind of social information that frequently triggers emotions.

These ways of expressing our attitudes to others can be divided into more basic and more complex stimuli, and the basic level is the one that is important for infants that associate more and more kinds of natural information with emotional reactions. On a basic level, facial expressions can be understood early, and the way the caregivers interact emotionally with their infant has a strong impact on the emotions that the infant shows. All kinds of bodily relations that the caregivers have toward the infant can be seen as typical triggers of emotions as well. Embracing an infant too tightly can evoke anger, touching her gently can produce joy, and leaving her alone can produce fear. Infants can pick up this kind of socially significant information early on, since infants come into the world well equipped for the direct understanding of others' intentions and feelings (see the following discussion). On another level, more complex stimuli come to be similarly important. For example, people react with pride when praised in public, feel guilty when seeing that they have missed someone's birthday, or feel angry when they have been insulted by someone. A perceived insult or a note about one's birthday can be seen as a natural sign, too, but it requires extensive background knowledge to be able to grasp it (Millikan 2004). Therefore, I focus on the kind of natural information that is available to the infant on a nonconceptual level.

Fear and disgust are probably the most prominent examples for evolutionarily "old" emotions with basic homeostatic functions that regulate the organism's well-being. The standard description of fear situations is therefore that of meeting a predator in the woods (e.g., Cosmides and Tooby 2010). While I don't doubt that fear and disgust are evolutionarily set up

emotions with the function of representing dangerous and indigestible features, it is worth highlighting that these emotions also have a social dimension, and that these emotions can be socially shaped on a nonconceptual level by social impact perceived through locally recurrent natural information. From an ontogenetic (i.e., developmental) point of view, early fear reactions certainly do not occur in the presence of predators and do not prepare the infant to run away. The most primitive fear is probably a fear of abandonment, lacking guidance and comfort in an uncertain environment. Reacting fearfully in early infancy, if anything, has the function of communicating being afraid in order to solicit a certain kind of help (Parkinson, Fischer, and Manstead 2005, 205). From a phylogenetic (i.e., evolutionary) perspective, being prepared to flee a predator is usually described as the primary function of fear, while communicating signs of fear might be a secondary function. But things are the other way around from an ontogenetic viewpoint: first, we learn to communicate that we are afraid, and then we learn to react with flight behavior ourselves. This is important insofar as we can expect early fear conditioning, as well as the development of fear regulation and output patterns, to be largely shaped by social interaction.

Infants are able to understand facial expressions and bodily gestures of other people from early on (see the following discussion), and these are among the most important natural signs enabling the infant to grasp situations that are of emotional relevance. Beginning at nine months, infants start to check their caregiver's faces for emotional expressions to double-check how to understand a situation (Campos and Stenberg 1981). If the caregiver shows a disgusted face with regard to a certain toy, the infant is more likely avoid playing with it. Disgusted faces can thereby be seen as natural signs for indigestible and poisonous things, and they constitute a first channel through which the set of things to which an infant tends to react with disgust is broadened. Yet the occurrence of facial disgust expressions on certain occasions is also a first filter through which personal and cultural differences can be shaped.

Anger, on the other hand, seems to be a basic emotion set up through evolution, which appears to have a fundamental social character at the same time. Among the earliest triggers of anger are bodily restrictions such as embraces that are too tight. Anger seems to be set up to defend one's place in the social arena against literal and metaphorical restrictions, just as fear and disgust are, on a basic level, concerned with stimuli that literally affect the organism's well-being. The struggle against a too-tight embrace is also a literal struggle for the space one needs to be able to breathe, whereas more complex triggers of anger later on might react to situations that concern one's social status, or the well-being of closely related persons, or even one's car.

While I have given a definition of natural information that allows all the stimuli mentioned to carry natural information, there remains an obvious question. A characteristic of natural information is that it can be picked

up directly by perception, and thus it occurs in the world in forms that are modality-specific. We can pick up certain information by vision or by touch. Once we combine tactile and visual information into one representation of an object, people hesitate to call this a case of direct perception. This is even more pronounced if the tactile and visual perception do not happen at the same time and do not concern the concrete object but rather two different objects that are recognized as belonging to the same kind. Emotions represent core relational themes, and obviously core relational themes can be recognized through different modalities. For example, we can be scared by loud noises, or by a loss of balance, or by seeing an object approaching us quickly.

Is it still direct perception even though the information that travels through different channels is recognized as being of the same kind? Or is that conceptual knowledge? This need not be concluded. One reason to differentiate between emotions and perceptions is that emotions are not direct in the way that perceptions are; they have to be triggered by a perception or by another type of mental state (see Chapter 3). Yet the fact that emotions are not direct does not mean that they must be conceptual. For a Dretskean this might be hard to swallow, since for Dretske, any kind of cognitive processing and conceptualization go together. Prinz solves this problem by arguing that emotions involve a "mental calibration file." In this file some stimuli might be stored from birth, others might be learned early on via association, and still others might be added later in life through all kinds of experiences and conceptual reasoning. Calibration files are necessary parts of emotions insofar as *every* emotion is triggered via the calibration file. Prinz takes the bodily responses, and the perception they trigger, to be the essential part of an emotion, which is what allows him to call his account an embodied approach; the perception of the bodily arousal is what counts as the appraisal of the external situation. Yet one could adopt Prinz's account, change just the detail of calling the calibration file the essential part of the emotion that elicits bodily reactions and a perception of them, and have an account not very different from a classical appraisal theory, where bodily reactions are a by-product of a neural appraisal system with no significant explanatory role.[6] Dretske, if he were to consider Prinz's theory of emotions, would probably take it to be a nice story about emotions and their *conceptual* content.

A more radical embodied view, according to which the bodily reactions involved in emotions function as appraisals themselves and thereby perform part of the intelligent reaction, needs to be tied to an embedded and action-oriented view. What Prinz introduces as the calibration files is a central intracranial component that "decides" which stimuli are emotionally relevant and then triggers further reactions. This is what allows us to interpret his account as a traditional appraisal theory by simply shifting the emphasis a little. The question is whether the unifying function of the calibration file is needed for the explanation of emotional reactions. I take it that this is not the case. In Chapter 2 I introduced the dynamical view,

suggesting that we could think of emotions as being constituted by psychophysiological processes. These processes are not centrally governed by a cognitive or neural program but, rather, unfold in a habitual way. The different bodily reactions involved in emotions can be seen as coupled to external stimuli that cause them (where the various stimuli are united by the relational property) and, at the same time, as dialectical processes that can activate each other. I argue further for this view in the next chapter.

Consider the studies where people evaluate cartoons as funnier when they are unconsciously forced to put on a smile by holding a pen between their teeth (Laird 2007, Adelmann and Zajonc 1989). In this example, the facial expression is not only an output of the reaction. Rather, the feedback from the facial expression seems to play a role in the evaluative process, yet it does not produce the emotional reaction by itself. The facial feedback facilitates the cognitive evaluation, and the two processes together lead to an emotional reaction. According to Prinz, both the facial feedback and the cognitive evaluation are stored in the calibration file and can therefore come to trigger the bodily pattern that constitutes joy. But that is rather unsatisfying because it does not really explain why the facial reaction has any impact on the cognitive evaluation.

Another explanation is that the various bodily elements that can be part of an emotional reaction are not centrally "stored" or "governed" but rather can be elicited in several ways, sometimes together and sometimes in the absence of one another, and that many of them can also stimulate each other. A facial expression through direct feedback to the brain produces certain reactions in the nervous system and facilitates further cognitive evaluations of the stimuli in question. We do not need to presuppose that there is a calibration file in which the expression is "stored" in whatever format as a potential trigger of joy. Rather, the facial feedback directly causes the activation of the nervous system and contributes to a certain phenomenal quality, which can have an impact on how we cognitively evaluate a cartoon in front of us. The cognitive evaluation, then, can again trigger further reactions such as laughter and thereby activate even more bodily reactions that are part and parcel of joy.

Griffiths (1997) objects that such a dynamical view is unlikely to be true for humans, since the stimuli that can produce emotions are infinitely many and human environments are complex and rapidly changing. I agree that there is no unified class of emotional stimuli. But stimuli are only different kinds of proximal information whereas the distal information or intentional object of the emotion is always the same core relational theme. Core relational themes are higher-order properties that come with a diverse range of objects and situations, but they are present in the socially structured environment of the organism and strike the organism through different kinds of natural information through the organism's sensory channels. The nature of core relational themes as relational properties is the topic of the following section.

4. RELATIONAL PROPERTIES

Emotions are about core relational themes, and core relational themes can be understood as relational properties (plus the related action tendencies that I address in the next chapter). As I have argued, an embodied account needs to make this claim in order to support DEE and SEE with regard to emotions. As we have seen, core relational themes are things that matter or that are of value *to us*. Therefore, a pressing question for an embodied account is whether the values that emotions are about can reasonably be understood as objectively existing properties. They are objective in the sense that we can come to represent them, but—contrary to response-dependent properties—their existence does not depend on our representing them or in any sense responding to them.

Relational properties are properties that objects genuinely have but only in relation to another object or another organism. But even if one of the relata is a living and perceiving organism, whether the property is instantiated does not depend on the perceiving organism. My friend is bigger than I am. This is a property she has only in relation to me. But this does not depend on my perceiving that she is bigger than I am. Relational properties can be defined like this:

> *DF. RP:* A property P is a relational property if an object A has that property, not intrinsically, but only with regard to another object B.

You cannot identify a relational property without referring to the relation between the two objects that constitute it. If we assume that emotions represent relational properties, then emotions represent properties that can be instantiated and individuated regardless of whether the organism perceives them. But when an organism perceives an object with a given property, this typically causes a certain type of emotion. One could argue that objects or situations that are dangerous typically cause fear. For example, a poisonous snake is dangerous whether I perceive it or not. But when I do perceive it, I am scared. Objects that are indigestible or poisonous typically cause disgust. Spoiled eggs are poisonous whether I perceive them as such or not, but when I do perceive this property, they usually disgust me. Note that the perception itself is not an essential part of the whole process. I could also think about something's being dangerous or remember something's being dangerous. But such cases according to the present theory are more refined forms of the emotions in question that phylo- and ontogenetically develop later on. Zebrafish and fruit flies only know fear of objects that are before their eyes, as do young infants. Relational properties need not be present before our eyes for every single case of an emotion, but they need to be there so that we can develop the responses in question in the first place.

Relational properties are good candidates for properties that would have struck our ancestors over and over again, leading to the development of mechanisms to represent them. They can therefore account for DEE. Emotions' being subject to semantic norms can be explained by taking emotions to represent relational properties, where these representations can be adequate or inadequate depending on whether the property in question is instantiated or not. Relational properties can also explain SEE. If emotions are about relational properties that regularly occur in our environments in certain forms, we do not have to assume complex symbolic representations or chains of evaluations. Core relational themes are their own best models because they correlate with objects that can directly trigger emotional responses in us because they have the property in question.

But how can a naturalist account hold that the intentional objects of emotions are external, if the intentional objects of emotions are certain kinds of values? To make such a normative realism plausible I first suggest that "being good for" and "being bad for" are relational properties as well, and I then explain in what sense such properties fit into a naturalist theory. Peter Geach (1956) uses the distinction between attributive and predicative adjectives to show that good and bad are relational in nature.[7] Take the sentence "This is a red Mercedes" and compare it to the sentence "This is a big mouse." If I was wrong in uttering the first sentence because the car in question was instead a Mitsubishi, then the statement that the car is red is still true. If, on the other hand, I say that this is a big mouse and the animal in question happened to be a rat, then the predication is wrong as well, since a small rat is usually still bigger than a big mouse. Size predicates are examples of attributive predicates that we use in relation to a certain species, while color predicates are predicated additively. Redness is something that all red things have in common, while there is no such thing as the category of big things. Things are always big in relation to the category of things to which they belong.

Geach's claim is that "good" and "bad" are attributive as well. There is no such thing as the category of good things, and there is no such thing as the property of being intrinsically good. Things are good or bad with regard to the class of objects to which they belong. If "a goldfish is a good pet" is true, it does not follow that a goldfish is a good animal. While according to Geach things can be good and bad in relation to the class of things to which they belong, the relational properties that come to be represented in emotions are good or bad in relation to the organism in whom they can cause certain emotions. Things can be good in relation to the class of things they belong to, and things can be good in relation to a living organism that needs certain things for survival. Vitamin E is good for me, since I haven't had any today, but might be bad for someone else who has already had plenty. Encountering a poisonous snake might be bad for me but good for another snake that is looking for company. There is nothing that all good things have in common except their being good for something or someone; there

is no such thing as a category of good things, and no such thing as being intrinsically good. The properties that make a knife a good knife differ from the properties that make a friend a good friend. Geach's analysis concerns the predicates "good" and "bad," and concerns only cases where objects are good-of-their-kind. My own view concerns properties that are good for an organism. What both accounts share is the claim that things are not intrinsically good but only in relation to a certain standard. The aim of the present account is certainly not to offer a comprehensive account of goodness, the aim rather is to establish a view of the properties emotions refer to as properties that are good or bad for an organism with regard to external standards that are biological or social.

The claim that being good for and being bad for are relational properties is central for diachronic externalist semantics, since it enables us to understand good and bad as properties that occur in the external world independent of our representing them (but not independently of us). We have seen in the last chapter that the intentional objects of emotions can be understood as relational or response-dependent, but that response-dependent properties do not do the required work to satisfy DEE and SEE, because these properties are the results of representational abilities rather than their causes. But DEE demands that the properties in question have been there before the ability developed, and SEE demands that the properties in question are there so that we can emotionally respond to them. Relational properties that exist in relation to us, but do not depend on our representing them, suffice to account for DEE and SEE.

Also, taking things to be good or bad objectively but only in relation to other things is a promising way of introducing natural norms. We can introduce classes of things that are good and bad in relation to the organism in question and its well-being. We can introduce classes of behaviors that are good insofar as they obey a social norm that is common in a certain area. Such classes of things certainly do not answer Moore's question in a satisfying sense. If we talk about a certain organism with certain needs and an object that is poisonous for that organism, it still makes sense to ask, "Yes, but is it really bad for the organism?" since we might be talking about a type of cancer treatment that makes one sick and is good and bad for the organism at the same time. Also the notion of "being good for the organism" is pretty vague. It can mean "good for the individual," "good for the species," "good for survival," or "good for reproduction." Fear and disgust are probably good for the individual and for survival but not in any direct sense for reproduction. It is an epistemological or rather scientific question how best to determine the notion of goodness in question. The ontological claim here is that the notion of goodness is relational but not response-dependent, that is, that there are external biological and social standards that make an emotion adequate or inadequate. According to the present view one's being disgusted of broccoli is most likely a misrepresentation, which is a claim

that a response-dependent view would have a hard time to defend. Whether broccoli is really good in Moore's sense is not at issue here.

One might object that this is a poor notion of normativity. However, I would rather say that it is a comprehensive notion of normativity that includes all things that can be right or wrong, or good or bad, with regard to a certain standard. It is then a further question whether the standard in question is a biological or a social or moral standard. And it is yet a further question what differentiates biological and social from moral standards. My aim here is not to defend a general version of moral realism but to show how introducing biological and social norms and things that can be objectively good or bad with regard to these norms is compatible with naturalism and allows us to see the intentional objects of emotions as relational properties.

Ecological psychology is a school that subscribes to a kind of biological naturalism and (at least in the case of Gibson) explicitly embraces normative realism with regard to the affordances animals directly perceive. This is because the idea of direct perception (or action-oriented representation) cannot do without the claim that the environment has a certain kind of structure in relation to the organism and that certain features afford things to the animal, that is, should be approached or avoided because they are either good or bad for the organism's well-being (Gibson 1986; Millikan 2004). It is the idea of affordance perception that allows us to discard the idea of animals as building inner representations, which can then be evaluated and finally lead to behavioral responses. Instead, the features of the environment are supposed to be of value in relation to the needs of the animal. These values can be directly perceived and guide behavior. I discuss James Gibson's approach of direct perception and affordances in detail in the next chapter. Here, suffice it to say that the idea that perception directly motivates action, not because of how we judge things to be but because of how we directly perceive them, implies that there must be something about how things are that prompts certain kinds of reactions. I do not see how the view that we directly perceive things as affording certain kinds of actions, can be defended without assuming that these things are goods of a certain value for the organism in question. Claiming that our environment is full of things that have the property of being of (dis)value in relation to us is a version of normative realism that avoids open questions about whether the properties in question are really good, since they are supposed to be good only in relation to something and not intrinsically or noninstrumentally. It also avoids the implausible claim that things are intrinsically good for everyone and in every situation alike, without thereby abandoning the claim that things are objectively good or bad for an organism independently of its being able to represent things as such. As the example of ecological psychology makes clear, in the context of biologically inspired theories, the commitment to relational properties is not a makeshift move to avoid

bioconstructivism but, rather, is part and parcel of ecological psychology. As Gibson nicely puts it,

> [t]he perceiving of an affordance is not a process of perceiving a value-free physical object to which meaning is somehow added in a way that no one has been able to agree upon; it is a process of perceiving a value-rich ecological object. Physics may be value-free, but ecology is not.

<div align="right">(Gibson 1986, 140)</div>

Given that our environment is full of value-rich objects, the next question to consider in detail is whether such biological values are plausible candidates to explain what the emotions are about.

5. CORE RELATIONAL THEMES IN A BIOLOGICAL ENVIRONMENT

The main aim of this chapter is to make sense of core relational themes as part of our environment. To do so we first need to introduce types of relational properties between the individual and the environment that fundamentally concern one's well-being, such as "being dangerous." "Being dangerous" is the relational property that something has in the relation to an individual that comes to be represented in fear. If we look back at Lazarus's list of core relational themes, it is easy to see how a central part of what is expressed by them can be translated into relational properties. The core relational theme of disgust is "taking in or being too close to an indigestible object or idea (metaphorically speaking)," and the relational property in that case is "being indigestible." The relational property present in anger is "a demeaning offense against me and mine"; "irrevocable loss" is the relational property we represent in sadness; the relational property present in jealousy, "loss or threat of another's affection;" the relational property represented by guilt, "having transgressed a social norm"; and for pride, "a valued object or an achievement, either one's own or that of some group with whom one identifies." But are these properties that we can assume to objectively exist in relation to the organism?

Fright, disgust, and anger are all taken to be typical examples for evolutionarily acquired basic emotions. They seem to involve panculturally present facial expressions; as reactions of the autonomic nervous system, they can be located in the neural circuitry; and they are present in early infancy and in other animals (see Chapter 2). This is what makes Lazarus's analysis of these emotions obvious cases of overintellectualization. Infants and animals do not seem able to evaluate that a situation is relevant for—yet not congruent with—their goals, that they should enhance their social esteem, that there is another person to blame, that an attack is a viable option, and

that the response of the environment to the attack will probably be positive. This kind of reasoning is even unlikely to take place in most cases of adult anger, since emotions appear to be cognitively impenetrable to a high degree (see Chapter 1).

I therefore argued that emotions are better described as being constituted by embodied functional mechanisms that are set up to represent situations of relevance for the organism's well-being. I also suggested in the previous section that relational properties must be understood as part and parcel of a biological naturalism that admits to a certain kind of normative realism. I now argue for this claim in more detail with regard to the relational properties represented by fear and disgust.

One stone being bigger than the other is a typical example of a relational property; it is perfectly objective and measurable, yet only exists in the relation between the two objects. "Being bigger than x" cannot be an intrinsic property of an object, yet it undoubtedly exists, independently of whether we perceive it or not. That a substance is poisonous for an organism is also a relational property and can be proved independently of the organism's abilities to perceive the substance's being poisonous. The commitment to relational properties is therefore generally not restricted to biologically inspired versions of naturalism or to ecological psychology. A stone being bigger than something else and paracetamol being able to cause liver failure in certain organisms are properties that can be explained with reference to physical and chemical laws.

Yet this is rather different with the properties emotions are about, which seem to be situated on a somewhat higher level. Being dangerous, indigestible, or offensive are properties that cannot be reasonably explained with reference to physics or chemistry alone. To say that something is dangerous for something else introduces a kind of normativity into the story that can only be captured adequately in a biological framework. It makes sense to speak of "dangerousness" as a property that exists in relation to the organism only if the survival of the organism is introduced as a basic value in the first place. If survival is taken to be a basic value, then it makes sense to say that something is dangerous, indigestible, or offensive and to say that all these things are bad for the organism. Negative emotions represent something as being bad for the organism; more specifically, fear represents something as being dangerous; disgust, as indigestible; and so on. Relational properties can therefore be described as occurring on different levels. From a biological perspective, it can be said that something is bad for the organism or dangerous. In many cases it might also be possible to describe what makes something dangerous from a physical or chemical perspective. A snake is dangerous because of its ability to bite and release poison into the blood. And the snake being poisonous can be analyzed in chemical terms.

But although many instances of core relational themes refer to events that can be described in physical or chemical terms, the biological level is needed to get an adequate description of the properties *as they come to be*

represented in core relational themes: fear doesn't represent the chemical structure of the snake's poison; it represents the snake's being dangerous. The representational mechanism in question does not react to only one kind of poison; it reacts to a variety of stimuli that occur in the biological environment of the organism. What these objects and situations have in common is that they are dangerous and therefore bad for the organism. In the case of fear or disgust, it might be possible to give physical or chemical descriptions of the property in question in some cases. But these descriptions only capture the present fear reaction and do not adequately describe the content of fear in general. Furthermore, there are cases of fear, anger, and disgust reactions that cannot be described on a physical or chemical level in a fruitful way, namely, cases that involve the social world. I return to these cases later.

Assuming that emotions are not evaluations or about response-dependent properties (as enactivist accounts would have it) but about objectively existing relational properties allows for the claim that these properties can be instantiated and individuated regardless of whether the organism is able to perceive them. A token of a relational property (that is not a response-dependent property) can be instantiated without the organism actually perceiving it, but there can also be certain tokens with a relational property frequently present in the environment of a species, without any member of the species being able to represent this particular kind of token. A new predator might enter the environment of a species and would constitute a danger even if no member of the species has the ability to represent this property. The Dodo is an example of a not only flightless but also fearless bird that was endemic to the island of Mauritius where it evolved in isolation from predators. The bird became extinct briefly after sailors introduced rats and other animals on the island in the seventeenth century that plundered Dodo nests. Dodos had no naturally evolved abilities to detect these predators and defend themselves. Nevertheless, it makes sense to say that the predators were dangerous for the Dodo's. In a less extreme case than the entirely fearless Dodos being plundered by a hungry crowd of artificially introduced predators, the presence of a new predator might not lead to extinction but to a selective pressure that would cause a change in abilities. Any perceptual or emotional mechanism that would enable members of the prey species to represent predators as a danger to avoid would be of great value for survival.

While it makes sense to say that *certain kinds* of dangerous situations start to exist at some point in the history of a species, it is hard to imagine a living organism in an environment where danger is never instantiated. It seems that organisms that are alive but mortal and vulnerable are, by definition, organisms that can find themselves in danger. It is furthermore hard to imagine any historical or possible environment that entails nothing that could threaten an organism's well-being. "Being in danger" seems to be a property that deserves the name of a "core relational theme" in a very

fundamental sense. Being alive, mortal, and vulnerable implies the possibility of being in danger. Being afraid is a reaction that we find in various forms in many animals, such as zebrafish (Kalueff et al. 2012) and fruit flies (Gibson et al. 2015) that has the function to deal with dangerous situations. The same is true for disgust. As soon as there are living organisms in need of nourishing themselves who are able to absorb certain things but not others, they need to be able to distinguish the things that can be absorbed from those that cannot. Disgust is a reaction with many homologues forms in animals that enables them to avoid things that are indigestible (Kelly 2011). As with fear, given the biological context of explanation and the assumption of relational values, it seems plausible to take "being dangerous" and "being indigestible" as properties that exist in relation to all living organisms and are of central relevance for survival. Therefore, it is not surprising that many animals developed a simple embodied action-oriented representation that represents these relational properties.

One might still object that such a biological approach hardly captures what Lazarus had in mind when talking about core relational themes. It might work for giving an explanation of fear and disgust, but already anger seems to be a harder case. Lazarus argues that the core relational theme, "demeaning offense against me or mine," results from an evaluative process in which a current situation has been evaluated as relevant yet incongruent with one's goals, since someone else can be blamed for having acted harmfully against us, and an aggressive reaction seems to be appropriate and its effect promising. Some of Lazarus's intuitions can be captured by the present approach, but some aspects should be seen as possible elements of adult anger rather than essential parts of the emotion.

Anger is an example where the formulation of the core relational theme, as Lazarus coins it, does not really capture the core of the emotion's aboutness. Anger appears early in infancy. Eight-week-old infants have been shown to react angrily: they first learned that pulling on a string led to the appearance of a slide showing a baby's smiling face and the *Sesame Street* theme song being played; when this outcome was removed, the infants pulled harder on the string and showed angry faces (Lewis, Allessandri, and Sullivan 1990). Something looking very much like anger can also be observed in five-month-old children whenever their movements are physically restricted (Watson 1930, Camras et al. 1992). To capture these cases, anger could be described as perceived goal interference, where some kind of resistance prevents us from getting through (as an instance of anger where we are in a rather active mode) or where an external situation is massively threatening our own position (where we are in a rather passive mode).[8]

Both forms of anger-eliciting scenarios can be understood in a literal and a metaphorical manner. The infants in the study could not get through to an immediate goal or were physically restricted. Adult anger reactions frequently occur when we have the impression that we "cannot get through"

to somebody we are trying to communicate with or when somebody disrespects our social position. Similarly, as with disgust reactions, there is a literal meaning of the core relational theme that is expanded via associative learning in the social context. To capture this, the core relational theme of anger should be described as "being restricting" rather than "being offensive." Being offended is one way of being restricted. Blame and illegitimacy can be seen as subsequent cognitive articulations of more simple perceptions that are at the core of anger (Parkinson, Fischer, and Manstead 2005, 203). Being restricting can be described as a relational property in the same sense as being dangerous or being indigestible. In a biological environment, several objects and organisms can come to restrict the organism with regard to its goals. Detecting restricting situations and reacting aggressively is an embodied functional mechanism. We recognize restricting situations by responding to them with a bodily reaction that prepares us to "fight back." The bodily reactions involved in the emotion constitute our feeling angry, which is a nonconceptual mode of making sense of, and interacting with, the world. Recognizing the situations in question as situations where somebody offends us or can be blamed for something are cognitive refinements of this reaction that are learned later on. In the next chapter I discuss more about how we can think of having a goal, and evaluate something as being incongruent with that goal, and evaluating the current situation with regard to the possible effects of reacting angrily. I argue that these shouldn't be understood as cognitive appraisals but rather as the skillful parts of an embodied action-oriented representation.

To sum up what I have argued for, introducing relational properties into one's ontology offers a way of adding properties of value for an organism to a naturalist ontology. Relational properties concerning the organism's well-being provide the ontological grounding for a naturalist, noncognitivist theory of emotions. They make sense of the claim that our environment confronts us not as unstructured matter but includes things and situations that are dangerous, offensive, or indigestible. And, as I argue in the next chapter, because that is how our environment is and has been structured, we can come to represent it via simple mechanisms.

I have been concerned with fear, disgust, and anger so far. These emotions are usually described as basic or evolutionarily acquired reactions. I argued that the core relational themes of these emotions are relational properties that exist in the ecological environment of an organism. Very well, Lazarus might say, but in doing so you have picked out those emotions for which this kind of explanation can easily be given. Yet there are other emotions that are about social rules and norms. How is your theory supposed to work for those emotions? Do jealousy, guilt, and pride react to relational properties that occur in the biological environment as well? I discuss these emotions, what their core relational themes are, and how infants can come to grasp them in the next section.

6. WHAT IS COGNITIVE ABOUT HIGHER COGNITIVE EMOTIONS?

Core relational themes, such as "resenting a third party for loss or threat of another's affection" in jealousy or "having transgressed a moral imperative" in guilt, should not be understood as internal judgments but be translated into relational properties that can be represented by embodied action-oriented representations. Before turning to the ontological question of what kind of relational properties we have to assume for the emotions in question we face another challenge. The main challenge in defending the claim that emotions such as jealousy or guilt could be embodied action-oriented representations is that these are often labeled as "higher cognitive emotions" and are therefore described as involving certain higher cognitive abilities. I first discuss what kind of abilities are frequently assumed and then argue that we do not need to ascribe these abilities to infants or animals to which we ascribe emotions such as jealousy or guilt. I return to the ontological nature of core relational themes only afterwards.

First, it has been argued that so-called higher cognitive emotions presuppose the ability of self-referencing (Lewis et al. 1989, Lewis 2014).[9] Self-referencing can be seen as a sign of the presence of an explicit representation of the self in relation to others. That means that somebody has to be able to entertain explicit I-thoughts, as well as recursive thoughts regarding her relation to others such as "I know that she knows that I bought a new tie."[10] Empathy, jealousy, embarrassment, and perhaps envy are emotions that presuppose this form of self-consciousness as a necessary condition. These emotions appear briefly before or around the second birthday. The empirical evidence for the claim that infants cannot entertain emotions such as empathy or embarrassment before being self-conscious comes from studies that try infants on the rouge test and then bring them into situations that are supposed to trigger embarrassment or empathy. The results show that most infants who do not pass the rouge test do not show signs of embarrassment or empathy either (Lewis et al. 1989, Bischof-Kohler 1991). This seems to suggest that the capacity for self-referencing has to be in place before an infant can show signs of embarrassment or empathy.

Second, there is the understanding of social rules and norms. Lewis distinguishes self-evaluative emotions from self-conscious emotions. Self-evaluative emotions require self-consciousness as well but have a further necessary condition: they require an abstract understanding of social rules and norms. Self-evaluative emotions are guilt, shame, and pride and these emotions emerge around the age of two and a half. Following Lewis, when we feel guilty we represent that we have violated a social rule; when we are proud we represent that we have achieved a goal others find laudable. The reason that infants younger than two and a half cannot entertain such emotions is that they are unable to represent such normative contents.

Lewis's claims have been highly influential in the last decades in developmental psychology and, furthermore, he nicely captures what many—if not most—people implicitly or explicitly claim when they explain the normative assessability of those emotions that are concerned with social rules and norms, such as Descartes's definition of envy, introduced in the first chapter:

> Envy . . . is a kind of sadness mingled with hatred, which results from our seeing good coming to those we think unworthy of it. Such a thought can be justified only in the case of goods of fortune. . . . But sometimes fortune gives advantages to someone who is really unworthy of them. Then envy stirs us up only because having a natural love of justice, we are vexed that it is not upheld in the distribution of these goods.
>
> (Descartes 1649/1988, §182–183)

The evaluation of a person as unworthy of receiving certain goods is certainly a case of an evaluation of the other person with regard to certain standards, rules, or norms, combined with a desire that one had received those goods instead.

Or take Hume's famous definitions of pride and humility, where he distinguishes between the object and the cause of a passion. In the case of pride and humility the object is always the self. But different causes, like a brilliantly done essay or a failed test, can lead to either pride or humility. It is important here that both object and cause build up the emotion in question:

> Beauty, consider'd merely as such, unless plac'd upon something related to us, never produces any pride or vanity; and the strongest relation alone, without beauty, or something else in its place, has as little influence on that passion.
>
> (Hume 1739/2007, II.i.2)

Again, one could argue that to make sense out of this, you would have to accept that in order to entertain pride you need an explicit concept of the self: you need to evaluate something as beautiful or worthy, and then recognize that it is yours. This is also the reason Davidson (1976) argues that Hume's theory of associated impressions and ideas has to be translated into the inference that we experience pride for a reason and this reason always includes ascribing something of value to ourselves. The same assumptions can be found in Lazarus's definition of the core relational theme of pride as "enhancement of one's ego-identity by taking credit for a valued object or an achievement, either one's own or that of some group with whom one identifies."

Although the earlier examples fit nicely with Lewis's definition, I present evidence that suggests that the emotions in question can occur in infants before they have objective self-consciousness and can explicitly evaluate social situations with regard to certain standards, rules, or norms. This

evidence will constitute the primary motivation for pursuing a noncognitivist account of "higher cognitive emotions" describing them as embodied action-oriented representations that are embedded in a social context. What the evidence shows (in accordance with the evidence presented in Chapter 2) is that there are good reasons to believe that so-called higher cognitive emotions might be present early in infancy. This puts pressure on emotion theories—whether cognitivist or not—to come up with explanations for this kind of phenomenon. My own account puts those observations into a theoretical framework that makes sense of the claim that these emotions are present earlier than the abstract concepts of a self and social norms by taking social norms to be part of the structured environment to which emotional reactions are adapted.

The developmental psychologists Draghi-Lorenz, Reddy, and Costall (2001) argue that the evidence poorly supports the view that higher cognitive emotions only occur after infants acquire the capacity to represent abstract goals and rules and explicitly represent the self in relation to others. Even worse, the small amount of existing research seems to support the claim that "nonbasic" emotions occur much earlier than usually assumed. First, two- to four-month-old infants show behavioral patterns resembling adult expressions of embarrassment during positive interaction with a caregiver. They show a specific combination of smile and gaze aversion, sometimes accompanied by hands raising toward the face (see also Reddy 2008). This pattern is similar to embarrassment expressions that have been observed in adults across many cultures (Reddy 2000; see also Draghi-Lorenz, Reddy, and Morris [2005] for evidence that adults frequently evaluate these expressions in infants as embarrassment). In addition, five- to six-month-old infants show distress reactions that can be interpreted as jealousy when their mother is paying attention to other infants. This reaction seems to be specific to cases where the infant is afraid to lose exclusive maternal attention to another infant, since tests with control groups showed that infants don't react with distress when their mothers are paying attention to a book or an adult person (Hart forthcoming; Hart and Carrington 2002, 2004, Draghi-Lorenz 2010). And infants around twelve months show first signs of empathic concern and offer objects or gently touch the distressed other person, seemingly trying to comfort her. Notice that this is not merely a case of emotional contagion, since the infant is trying to help the distressed other person instead of just being affected by her emotions (Zahn-Waxler, Radtke-Yarrow and King 1992). Finally, from two months on, joyful reactions can be observed in infants that are praised by their mother for an achieved goal. The infant's caregivers frequently interpret these reactions as pride (Reissland 1990). Nadja Reissland herself interprets her observations not as cases of pride in infants but as pre-forms of the emotion. Yet what this shows is that there is little dispute over the early occurrence of signs of higher cognitive emotions in young infants. The dispute concerns the question of from what age on, and due to what cognitive abilities, can

we speak of genuine pride instead of pleasure or showing-off behavior as a response to being praised. While this interpretative question probably cannot be settled on empirical grounds yet, my aim in the following will be to take the evidence for genuine social emotions in infants seriously and offer a theoretical framework in which it makes sense to ascribe them.

According to Draghi-Lorenz et al., these observations of pride, embarrassment, and so on have frequently been explained away or ignored in similar ways because of the well-established theoretical assumptions that nonbasic emotions require capacities such as abstract rule and goal representations and explicit representations of the self in relation to others, which can only be found in children older than two years. Following such cognitivist suppositions, the infant's emotional reaction, for example, to its mother's praise for achieving a goal must be interpreted as a mere pleasure-response to being the center of attention. On this view such reactions do not deserve to be called genuine cases of pride, jealousy, or embarrassment (see, e.g., Lewis 2014). This theoretical paradigm is also the main reason why there is still little research on the occurrence of so-called higher cognitive emotions in infants. The interesting result of Draghi-Lorenz et al.'s meta-empirical research is that the standard cognitivist interpretations of the data are based mainly on theoretical assumptions concerning the further cognitive abilities of infants. The claim that infants cannot display nonbasic emotions is not empirically grounded.[11] The point at which the reaction of a smile to someone else's expression of admiration deserves to be called genuine pride is an open question. As long as there are no criteria to decide whether an infant shows a simple pleasure reaction or should be described as proud on empirical grounds, the plausibility of the claim that infants can show pride, embarrassment, jealousy, and the like, depends on further arguments about the abilities of infants to entertain such emotions.

To question the typical theoretical assumptions, it must be shown that nonbasic emotions do not require an explicit representation of the self in relation to others, but can make do with simpler abilities, which may well be present at an early age. This is what I do in the next section. Furthermore, it must be demonstrated that we do not need an understanding of abstract rules or goals to display nonbasic emotions. This is what I do in what follows.

7.   THE SELF IN SELF-CONSCIOUS EMOTIONS

The claim that emotions such as envy, empathy, jealousy, and embarrassment presuppose self-consciousness might seem very plausible at first. Those emotions fundamentally concern the relationship between the self and others. In empathy, one reacts to another's emotion with a similar feeling. Somebody who is jealous sees the specific relationship that she has to another person as endangered by someone else. In embarrassment, one feels

bad about a public lapse because of how others might evaluate oneself. It is hard to imagine that those emotions could occur in a subject who is not self-conscious and thereby does not have any explicit representation of herself.

But I think this intuitive affirmation rests on a misunderstanding of the notion of "self" in question. What Lewis wants in place before the onset of higher cognitive emotions is the ability to show self-referential behavior, which is seen as the sign for the presence of explicit thoughts about oneself. His hypothesis, furthermore, is that at the moment when we can observe self-referencing behavior, an infant has already acquired the ability to think about herself in relation to others or about others in relation to herself. Lewis calls this an objective self-representation. I do not doubt that emotions such as embarrassment, envy, and empathy are all fundamentally social emotions that only make sense in the interaction of oneself and others. The child has to recognize other persons, their intentions and emotions, and has to make sense of her own position and possibilities in relation to the others. However, does an infant need an objective self-concept or explicit thoughts about herself to do so?

Self-consciousness comes in many forms and degrees. As Dorothée Legrand (2007) puts it, the simplest form is the prereflective self-as-subject. Experiences differ in their content and quality but they are all from the first-person perspective; they share the quality of "mineness." This consciousness of the self-as-subject is prior to consciousness of the self as object, of which the recognition of oneself in the mirror is a case. Susan Hurley (1997) develops a theory of a certain form of prereflective self-consciousness as well, and she, too, describes the first-person perspective, the quality of "mineness," as a fundamental part of consciousness. Hurley goes further, however, in explaining the first-person perspective not only as a phenomenologically graspable phenomenon but also as an effect of the interdependence of action and perception. Having a perspective, according to Hurley, means that what is experienced and perceived depends systematically on what the subject does and vice versa. Agency is essential to this form of prereflective self-consciousness. But the form of agency Hurley has in mind is an embodied kind of motor agency that can be ascribed to animals and infants as well. For an animal to move successfully through its environment, the ability to keep track of the relations between what it perceives and what it does is required. Therefore, the perspectival interdependence of action and perception involves the ability to use information about the self or about one's own states and activities. However, this doesn't mean that the animal has to have a general concept of itself. The perspectival use of such information is context-bound.[12]

There is an ongoing debate about whether prereflective forms of self-consciousness, as they appear in proprioception and motor agency, should count as self-conscious phenomena or whether full-fledged self-consciousness requires explicit thoughts about the self (see, e.g., Rochat

2009, Musholt 2013, 2015). I tend to agree with those authors who argue that while proprioception and motor agency might be the basis for the ability to develop self-consciousness and be helpful for understanding how explicit thoughts about the self can be acquired, the notion of self-consciousness should be reserved for the ability to entertain explicit thoughts about the self, an ability that is required in an interactive social process (Rochat 2003, Musholt 2015). But my aim is to investigate whether the prereflective self, understood as the ability to use proprioceptive information for the guidance of action, together with certain social abilities of the infant, can account for the early onset of what Lewis calls self-conscious and self-evaluative emotions such as embarrassment and envy. This could be the case even if the prereflective forms of self-consciousness differ radically from full-fledged explicit thoughts about the self.

What is lacking from the idea of the prereflective self as an embodied agent is an explanation of how the prereflectively self-conscious infant can relate to and understand others since the emotions in question are fundamentally social. Embarrassment, envy, and empathy are unimaginable outside of a social context in which the infant grasps the intentions or feelings of the other person. What would embarrassment be without any consciousness of another person watching one's misbehavior? What would envy be without the impression that another person has something one would like to have for oneself? And what would empathy be without another person who shows feelings to which we then react? If we want to avoid an objective self-representation, we still need more than just a prereflexive subjective perspective; we need a connection between the prereflective self and others.

Such a connection is articulated in the reasoning on *primary intersubjectivity*. Primary intersubjectivity is a pretheoretical, nonconceptual, embodied understanding of others (Trevarthen 1979, Spaulding 2014). Many authors from the vibrant field of embodied and social cognition defend primary intersubjectivity as a set of developmentally fundamental abilities that underlies and supports higher cognitive skills for understanding others. From birth on, infants have several capacities for engaging in human interaction that are emotional, sensory-motoric, perceptual, and nonconceptual ranging from facial mimicry to eye-movement detection. These abilities have been widely investigated in recent years. Evidence ranges from research on the mirror neuron system and its role in imitation learning, in the perception of communicative actions (Gallese 2009, Ammaniti and Gallese 2014), and in sharing emotions and sensations with others (de Vignemont and Singer 2006, Bernardt and Singer 2012), to psychological studies on the affective coordination between gestures and expressions of the infant and the interacting caregiver (Gopnik and Meltzoff 1997).

The result of these studies taken together is that young infants are capable of grasping purposeful intentions of others through the perception of bodily movements, gestures, and facial expressions. This capability is not replaced

but only supplemented by explicit knowledge in adults. In most social inter-actions, we do not infer what another person's beliefs or desires might be; instead, the other person shows her intentions through her actions, behaviors, and expressions, and we directly understand them through embodied simulation. Shaun Gallagher argues in this vein that in most intersubjective situations we directly understand the intentions of other people because they express them in their embodied actions, and we mirror them with our own capabilities for action (Gallagher 2005, 2008a). The direct perception of others' intentions is based on three conditions:

1. A proprioceptive awareness of one's own body
2. A differentiation between self and other
3. A recognition that the other is of the same sort as oneself

The proprioceptive awareness of one's own body is a constitutive part of prereflective self-consciousness. The claim that infants can differentiate between self and others but still recognize other persons as a conspecific is supported by evidence on infants' ways of attending to faces more atten-tively than to other objects (Meltzhoff and Moore 1977, 1983, Johnson 2000, Meltzoff and Brooks 2001). With his model of direct perception, Gallagher argues specifically against certain theories of mind, such as the so-called theory theory and the simulation theory, which are committed to the claim that in order to understand others' minds we have to form a the-ory or simulate the other's situation.[13]

A second pragmatic phase, labeled *secondary intersubjectivity*, where the infant starts to go beyond the person-to-person immediacy and engages in triadic relationships occurs around the age of one (Trevarthen 1979, Gal-lagher and Hutto 2007, Spaulding 2014). The child learns to communi-cate with others about objects in the environment. This process starts with the so-called nine-months revolution, of which joint attention is the most important acquisition. Infants between nine and fourteen months start to alternately monitor the gaze of the other and the object the other is gazing at to verify that they are attending to the same thing. Furthermore, infants tend to look at the caregiver's gestures or gazes to find out about his intentions or about the meaning of a specific object. Infants search the face of their caregiver for an emotional reaction, before trying to crawl up the stairs, and they play less with toys to which their mother reacted with a face of disgust (Campos and Stenberg 1981).

The general conclusion from these results is that infants, from nine months on, start to understand that others' movements and expressions depend on the pragmatic context in which they occur. Instead of merely searching the caregiver's face for emotional expressions, they start to make sense of others' reactions to the environment and thereby learn about the environment, as well as about the caregiver (Gallagher and Hutto 2007). The pragmatic idea behind this approach goes back to Gibson's notion of

affordance-perception: we always see things in relation to their possible uses, and therefore never as a disembodied observer. While this idea is popular with regard to the perception of the environment (e.g., for the claims that we directly perceive the floor as walk-upon-able and the tree as climbable), it is less well known that Gibson himself explicitly states that the notion of affordances covers the perception of the social world as well:

> The richest and most elaborate affordances of the environment are provided by other animals and, for us, other people . . . [T]hey are so different from ordinary objects that infants learn almost immediately to distinguish them from plants and nonliving things. When touched they touch back, when struck they strike back; in short, they interact with the observer and with one another. Behavior affords behavior . . . Sexual behavior, nurturing behavior, fighting behavior, cooperative behavior, economic behavior, political behavior—all depend on the perceiving of what another person or other persons afford, or sometimes on the misperceiving of it.
>
> (Gibson 1986, 135)

While this gives a vivid impression of what direct perception might be, not only in biological but also in social environments, there is a heated discussion about whether embodied, enactive, or ecological accounts can really make sense of early social interactions without ascribing *any* kind of representations of others to infants (see, e.g., Ciaunica 2014, Spaulding 2014, De Bruin and Kästner 2012). But what most current authors would agree on is that (1) infants obviously come to the world well equipped for social interactions with others and (2) have various ways to make sense of the feelings and intentions of others long before they pass the classical false-belief task around the age of three. Rather, the dispute circles around the question of how to account for these abilities in infants and whether we can explain them in terms of direct perception or if we need to ascribe at least some kind of minimal representations to the infants.

　　The claim I want to establish, that self-conscious and social emotions can occur before infants develop the conceptual skills to self-reference and to explicitly understand social rules and norms, is compatible with all of these views, since they all acknowledge a preconceptual sense of oneself in relation to others. I come back to the debate about antirepresentationalism and minimal representations in the next chapter and defend a view of emotions as embodied action-oriented representations. But the main question at issue here, namely, whether the so-called higher cognitive or self-evaluative emotions can be understood in a noncognitivist framework, is open to more or less radical approaches to self-consciousness and the understanding of others, as long as the claim is not that what we need is an explicit and objective representation of the self and conceptual representations of the mental states of others.

As we have seen, it is a common cognitivist claim that higher cognitive emotions presuppose an objective self-concept. For example, jealousy presupposes an understanding of the relationship between the self and the other and what one can reasonably demand from another person; pride involves the recognition of one's being laudable. Yet empirical studies suggest that five-month-old infants show distress when their caregiver pays attention to other infants. From two months on, infants praised by their caregivers react with joy and show various kinds of behaviors that can be labeled as showing off (Reddy 2008). These results are not controversial. On the contrary, cognitivists also accept them. What is controversial is whether we can interpret such early reactions as actual cases of jealousy and pride, or whether jealousy and pride can only be present once objective knowledge of the self and the relations to other persons are present. Lewis suggests the latter and argues that early occurrences of jealousy- or pride-like expressions ought to be interpreted as precursors of these emotions (see, e.g., Lewis 2014). I argue for the former in the following.

If we ascribe an embodied sense of self to infants and the ability to successfully understand others' intentions and feelings via direct perception or action-oriented representations, we must also assume that infants can show empathy, jealousy, and so on without having conceptual knowledge of themselves and their relation to others. As far as we know, empathy, embarrassment, and jealousy all occur as expressive forms in situations where it makes sense to show the emotion in question. It is thereby not an over-interpretation of a random bodily reaction, as it would be if we counted the smiles that newborns show in the first weeks of their life as joy. Instead, only after some weeks do we observe that infants start to react with smiles in interactive situations where it does make sense to smile. Therefore, these early occurrences of smiles can be interpreted as joy. In the same vein we can distinguish jealousy from other kinds of stress reactions, because jealousy reliably occurs in situations where the caregiver turns her attention to other infants. Emotions should be described as reactions that are embedded within a social context and that social contexts are structured insofar as certain patterns of interaction between infant and caregiver become established immediately after birth and can be recognized by the infant from early on.

The relationship between infant and caregiver is first of all characterized by mutual attachment (Rochat 2009). Second, the infant is needy and depends on the caregiver to take care of her needs. Given that an infant is in a closely attached relationship with her caregiver, and given that she can reidentify the caregiver as the source that steadily supplies all kinds of needs and desires, it is not surprising that an infant is sensitive to the lost attention of the caregiver and that she is able to distinguish between situations where the caregiver is paying attention to a book and situations where the caregiver is paying attention to another infant. Therefore, an infant is able to represent situations in which its well-being is threatened because of the lost attention of the caregiver.

Yet one might argue that jealousy involves more than just that. Lazarus calls jealousy a reaction that has "resenting a third party for loss or threat of another's affection" as its core relational theme. "Resenting" could involve a kind of moral blame: when we are jealous, we think of the other person as deceiving us. This is even clearer in Prinz's formulation that jealousy represents "infidelity." Infidelity is certainly a morally loaded notion and to detect something as a case of infidelity requires an explicit understanding of social rules and norms. How we can think of emotions as being about social rules and norms without overintellectualizing their intentional objects is the topic of the next section.

8. CORE RELATIONAL THEMES IN A SOCIAL ENVIRONMENT

I have argued so far that emotions such as envy, embarrassment, guilt, and jealousy do not presuppose an explicit or objective sense of self, but rather an embodied sense of agency that appears to be present from early on. I now return to the question of how to think of core relational themes as relational properties that are given in the environment of an organism (and that motivate skillful responses). Before addressing the question of how to make sense of the relational properties in question from an ontological point of view, I first clarify what I take the core relational themes in question to be, since the description of core relational themes themselves is often overintellectualizing them. Yet an overintellectualization of the content of jealousy sets different demands for a theory of the representational abilities we need to assume to explain the emotion in question. You might find it plausible that "being the center of attention" is something that we can grasp on a nonconceptual level whereas "being the one who spoiled the party" is something that we can only understand conceptually. In this sense it is a central question what precisely social emotions are about.

Many people would have it that jealousy is about another's infidelity. While I do not doubt that scenarios and considerations regarding "infidelity" can be part of jealousy reactions, they are not essential or necessary parts of being jealous. If I am in love with somebody who happens to be in love with somebody else, I can be jealous, without representing any sort of infidelity. Jealousy is usually applied to a three-person situation in which another person is involved with someone who is or was involved with you. Parkinson, Fischer, and Manstead (2005) suggest that jealousy is an emotion that functions as a means of communicating being left out, and thereby soliciting reintegration into interpersonal interaction. In accordance with that hypothesis, I suggest that "being left out" or "withdrawn attention" is what jealousy is about (and, as I argue in the next chapter, "reintegration-to-call-for" is the action tendency that comes with the emotion). Being left out is a relational property the emotional person herself has in relation to others who interact without integrating the person into their interaction. The

scenario that jealousy tracks and motivates to change is this. It is a difficult question what the precise ontological ingredients for the description of such a scenario are. I will come back to that question later.

Here I want to further clarify how the core relational theme of jealousy should be described in order to understand what it takes to be able to represent it. Lewis suggests that those emotions that he calls self-evaluative emotions presuppose an explicit understanding of social rules and norms because this is what they are about. But with regard to jealousy it is not entirely clear whether this emotion is about a rule violation or not. If jealousy were about cases of infidelity, it would be about particular cases of rule violations, such as those rules that are violated when monogamous partners betray each other. Yet we have seen that this is an overintellectualization that does not capture many instances of jealousy and is also inadequate to capture early cases of jealousy. If, on the other hand, we call the relational property that jealousy represents "being left out," it is not even clear whether there is anything normative about this property. In what sense, biological or social, is it of disvalue to be left out?

The answer certainly has to start with the fact that an infant is in need of her caregiver, is attached to the caregiver, and is accustomed to a certain amount of caring behavior from the caregiver. Given the usual closely attached relationship, it is a matter of fact that when a caregiver turns her attention to another infant she is leaving her own child out of the interaction. Since the infant depends on the caregiver's attention, it can be said that jealousy concerns something of value for the infant's well-being. Being left out of a social context has immediate consequences for the infant's well-being and, therefore, jealousy can be categorized as an early-occurring mechanism that reacts to an urgent social scenario.[14] Yet one way of describing the present scenario would be to say that by leaving the infant out of the interaction the caregiver frustrates a certain desire the infant has. In that case the infant is not responding to the violation of a rule and so there is no need to explain the cognitive capacity to do so. Another way of putting the case would be that in the interaction between infant and caregiver certain patterns of expressing attachment and attention have been established and have set up an implicit standard about how much attention to expect. This norm is violated in this case, and this is what the infant is emotionally responding to.

As far as I can see, there is no available evidence to clarify whether early jealousy reactions respond to a frustrated expectation that has its origins in the sheer neediness of the infant or whether the infant is rather detecting a rule that was established in the early interaction. If anything, only very careful comparative studies on the question of whether different amounts of caregiver attention lead to different kinds of jealousy behavior could answer that question. But an educated guess I have is that infants of caregivers that pay almost no attention to them and establish no reliable patterns of mutual attention will react with even more distress to the caregiver's paying

attention to another infant, simply because infants heavily depend on the caregivers attention. That would speak in favor of the claim that, at the very bottom of it, jealousy is about being left out, and being left out is a situation we detect and respond to with a certain urgency because we are beings in need of attachment, not because we learned certain rules of when it is appropriate to get attention. Still I do not doubt that from the earliest social interactions on rules of appropriateness govern who gets how much attention and also shape our expectations and embodied strategies involving how to ask for attention.[15] But these rules do not have to be explicitly represented by the infant (and, as I argue later, not even by adults) to be able to represent being left out and respond to it with the urge to be reintegrated or get back the lost attention. A similar case is embarrassment, which occurs very early in infants and seems to respond to public attention with the urge to avoid it rather than as a response to the violation of a social norm.

Yet there are emotions that obviously refer to social rules and norms or violations of these rules and norms like guilt and shame. For these emotions it is obvious that their intentional objects cannot be understood in any kind of biological framework alone that is concerned with bodily well-being instead of with the norms that govern our relations to others.[16] One could try to come up with a possible explanation of guilt without an understanding of social rules by thinking about possible early scenarios where guilt might occur in infants. Early examples of guilt include reactions to accusations made by close relatives. An infant could start showing recompensive behavior on seeing the stern and angry face of her caregiver. This could be explained with the already developed notion of the prereflective self and the direct perception of others' feelings. But this story is unsatisfying. A stern face might *trigger* a guilt reaction, but if we take guilt to be an emotion with an intentional content, it would be rather odd to say that guilt is *about* the stern face of the caregiver. In guilt we represent that we did something wrong with regard to a social norm. We might do so by first detecting another person's stern face, but the distal content, the core relational theme detected by the emotion, is the violation of a rule (and the need to make amends). The problem that a radically noncognitivist account faces here is to give an explanation of emotions as embodied action-oriented representations (or something similar), without ending up with a radically reductionist account that denies that emotions are concerned with values, social rules, and norms.

So how then, to echo Brooks, can we come to see core relational themes as their own best models when they involve the violation of social rules? How can we think of emotional objects as being out there, as a part of a structured environment to which we are well adapted? How, in other words, can we introduce social objects of emotions that satisfy both diachronic and synchronic environmental externalism? The challenge of an embodied, embedded view of emotions is to come up with an ontological explanation of our *social* environment and the way we are adapted to it. To say that the world is its own best model with regard to emotions means that

we negotiate our needs and intentions with others by understanding their emotional reactions and reacting emotionally ourselves, without forming complex inner representations about something being offensive or a violation of a rule, since these things are out there and we directly react to them. I have already argued that we can respond to things as being dangerous or offensive, since these properties reliably occur in the biological environment. To expand this theory to include social emotions means that we need to develop a similar theory with regard to the social world. Such a view includes realism with regard to social rules, norms, and values. The claim must be that the social rules and norms in question do not exist independently of us but do exist independently of our representing them as such.

What I want to sketch in the following is a social ontology that can satisfy DEE and SEE with regard to social emotions. According to my view the social world we are surrounded by is a human construction that is a historically developed, objectively existing structure that is permanently, causally influenced by us and influences us (causally and otherwise).[17] The entities that belong to the social world are (at least to a huge degree) things that are reproduced because they serve a certain function. These entities include, among others, artifacts, institutions, contracts, people with social statuses, forms of behavior, and communication. Nowadays common sense in social ontology has it that these things come into being via collective intentionality and prescriptive speech acts (Searle 1995, 2010). Humans impose functions on objects and people where objects and people cannot perform the functions solely by virtue of their physical structure. I call this the *intentionality-relative view*.

According to the intentionality-relative view, we declare things to be "such and such," and we collectively view and treat things as being "such and such" in certain contexts. These things thereby come to perform a certain function. We regard slices of metal and pieces of paper as money so that they can perform the function of a currency, and we declare a contract to be valid so that it can perform its particular function. Such collective ascription of a status to a certain object in a certain context is what Searle calls the constitutive rule, the rule according to which x counts as y in context c: the bills issued by the Bureau of Engraving and Printing count as dollars in the United States, the stonewall counts as the border of that village in that state, and so on. For the members of a community to collectively accept something is for each of them to have an acceptance attitude toward it. A constitutive rule therefore is not just to give a set of happenstance conditions for something to be a dollar or a border. It is to give the conditions for grounding a fact about a dollar or about a border (i.e., the constitutive rule spells out the metaphysical reason that something is a dollar or a border). A typical constitutive rule articulates the link between a set of grounding conditions x and a grounded fact of type y (Epstein 2015).

Searle speaks of social institutions that owe their existence to our collectively accepting them as objects that are epistemically objective: if I tell you

what the money in my pocket is worth, my claim can be proved true or false by objective standards. Social institutions are, however, ontologically subjective; like pain and tickles and response-dependent properties, they exist only as experienced by human beings.

The view I sketch here and apply to core relational themes is more radical than the intentionality-relative view. It claims ontological objectivity for artifacts, institutions, social interactions, and social rules and norms in the sense that these entities, interactions, rules, and norms are established through a history of common practices and can be causally efficient, even if nobody notices, experiences, or fully understands them. Their ontological status as a social object in many cases does not depend on collective acceptance at all and in other cases not on collective acceptance alone.[18]

Consider artifacts as an example. A screwdriver is not a screwdriver because somebody sees it as such but because the screwdriver belongs to a category of things that are reproduced because they serve a certain function. Consequently, a screwdriver can have the property of being useful to people in certain circumstances, independently of whether I notice that the screwdriver has this property or not. Artifacts owe their identity to the fact that they belong to a category, members of which are reproduced because they serve a certain function. I call this the *function realist view*. The function realist view has its roots in Ruth Millikan's (1984) work on biological categories and their function. Millikan explicitly uses the notion of functional categories to capture social categories, particularly language, as well. Yet she does not develop a full-fledged ontology of the social world herself.[19]

Note that the function realist view differs from the intentionality-relative view, insofar as Searle, the main supporter of the intentionality-relative view, explicitly states that the social world is constructed by the agent's intentional states with regard to it. Searle holds that from a god's-eye view, a screwdriver is not a screwdriver but just a piece of metal (Searle 1995, 12). While probably everyone (me included) would agree that without human beings, who are able to intentionally represent the world, the social world (including screwdrivers) could have never been constructed, I disagree that social objects owe their identity to observers and their collective intentional or declarative acts.

Possible god's-eye views are difficult to evaluate, but I take it that a screwdriver owes its identity to the reproductive history that makes it a member of a certain category of artifacts. This category exists independently of agents that see things as certain things. Objects that belong to this category have a certain function, like screwing screws, and if they cannot fulfill their function, say, because they are edgeless, this does not mean that they do not belong to the category, but that they are malfunctioning or broken members of the category. Functional categories are not only historical and objective but also normative. They give a criterion for separating well-functioning objects from objects that are broken or damaged. With regard to Geach, one could say that an artifact can be good or bad with regard to what the standard is for objects that belong to the category in question. This kind

of realism allows for the claim that a screwdriver can have the property of being useful or useless, independently of whether I represent the screwdriver as such. Such a realist account can ascribe a central role to material conditions, all kinds of social practices and the interrelations that hold between social objects, thereby reducing the role of collective intentionality and speech acts in the metaphysical grounding of an artifact.

Brian Epstein (2015) has recently developed a comprehensive critique of Searle's intentionality-relative view and what is known as individualism in the social sciences. Epstein—put roughly—suggests that facts about people and what they accept figure into anchoring frame principles and grounding social facts. Yet the social world does not revolve around us in the sense that a group is nothing but the set of its members or that the Supreme Court is what it is because we collectively accept it as such. According to Epstein, there can be many different frame principles expressing the grounding conditions for particular social facts, while according to Searle the constitutive rule is the one and only frame principle that explains how social kinds are glued together. What Epstein rather mentions than develops is the idea that we need to understand how functional roles and environmental facts can partly anchor a set of frame principles.

I think the function realist view contributes to that project from a slightly different angle. While Epstein himself focuses on rather complex social institutions like the Supreme Court, I am interested in the objects of social emotions, and to explain them I rely on Millikan's account with regard to social categories. Typical objects of social emotions are rule violations (in guilt and shame), achievements (in pride), and being left out (in jealousy). To say that we can perceive rule violations and achievements directly, we not only have to explain the ontological status of rules, their violations and achievements but also the form in which we are in permanent causal contact with them, because they recur in our environments.

The central difference between the function realist view and the intentionality-relative view of socially reproduced categories is that the function realist view can account for relational properties that have an impact on us whether or not we perceive them. According to Millikan, a screwdriver can have the property of being useful for us independent of our perceiving the screwdriver as such. In the same sense, a rule violation can take place even if nobody notices that it does. Social rules and norms do not (always and only) depend on individuals accepting that they are in place, as Searle would have it. There are many rules that we make up as we go along in our social interactions. Rules can be established as conventions that the members of a social group follow without being explicitly represented beforehand. It is not a necessary criterion for a rule to be in place that somebody represents it and then purposefully establishes it. It is a criterion for a rule that it is followed and can be violated by more than one member of a social group. But people can establish these rules without the intention of doing so and might or might not come to explicitly represent these rules later on.

Rules are visible in social contexts in the form of recurring patterns of behavior and in people's sanctioning behavior when a rule is not followed. Recurring behavior patterns are grounded in the dispositions of people to do similar things under similar conditions. Recurring greeting behaviors are grounded in the ability to wave hands, shake hands, and say "Hi" and "Bye." They get reproduced because they serve a certain function in social coordination and communication. Teaching somebody to follow an etiquette rule like "you should greet people you are familiar with" can be understood as a second-order behavior, that is, a behavior that refers to other behavior with the intention to correct it and thereby establish or maintain patterns of behavior. We teach our kids to smile, wave, and say "Hi" and "Bye." We tell them that it is not polite to hide their faces and hands, when parting from people they know. We establish the ability to greet in situations where it is appropriate in the infant and thereby maintain greeting conventions. While the rule itself is present in the regular patterns of behavior and in our utterances of it, rule violations become visible in the sanctioning behavior.

Such an account of rules and rule violations allows us to claim that animals can represent rule violations when they show shame, and infants can come to represent rule violations in guilt without having an explicit understanding of the rules in question. Some animals, like baboons, can establish complex social rank hierarchies in their interactions and can come to behave accordingly in their social environments. The social structures in question can be understood as networks of social relations between the animals. Being a parent of, being higher ranked than, and being a permanent grooming partner of another are the places or nodes that animals can come to occupy in these relations. Social relations are constituted through recurring patterns of behavior. Practices relate the animals to each other and the material world and situate them in nodes in a structure.[20]

Recent research suggests that chimpanzees have a social norm that forbids severe aggression toward infants (von Rohr et al. 2011). This rule is not only widely followed by chimpanzees; it also results in irritation on the part of an adult chimpanzee if she witnesses an infanticide, even if the infant in question is completely unrelated to her. Yet to explain the behaviors of baboons and chimpanzees we do not need to ascribe to them complex internal representations of rank hierarchies or abstract rules of how much violence is tolerable toward whom. Instead, we can think of the animals as being in constant causal contact with a structured social environment where they can follow simple cues that guide their behavior in accordance with certain established social rules. Showing a shame response is therefore an immediate reaction to perceiving that one was about to violate a social rule concerning one's own status in the rank hierarchy. That such a rule was violated can be communicated by the aggressive reaction of a higher-ranked animal. The shame reaction of the lower ranked animal can have a communicative function with regard to the other animal. The complex arrangement

of rules of behavior and the rank hierarchy established through such behavioral reactions do not need to be represented as a whole by any of the animals participating in the social system. Instead the animals make up the rules in the interaction and implicitly represent parts of the system through the others' behavior. A similar story can be told about infants' emotional reactions in social contexts.

Jealousy represents cases of being left out. In a setting where people live in certain constant relations with each other, interact with each other, and depend on each other, those relations are governed by a huge number of explicit and implicit rules. Marriage is a case of an explicit rule according to which something is a case of infidelity or not, and you can ascribe not only the relational property "being left out" to a betrayed spouse but also the property "having been unfaithful" to somebody who did not follow the rule in question. In the case of other relationships, the rules in question are more implicit than that. I have suggested earlier that jealousy is driven by the need to get a certain amount of attachment and attention in social contexts. Yet from early on caregivers establish rules of when and how getting attachment is an appropriate demand through attachment and recurring patterns of behavior. An infant depends on its caregiver and is accustomed to getting a certain amount of attention from her. The withdrawal of this attention is an objective fact that can violate a rule established by antecedent behavior and can be represented by the infant without any explicit understanding of the social rules and norms in question. Early occurrences of jealousy can be seen in cases where an infant represents the violation of a rule on the side of the caregiver. If attention and care are types of behavior that get reproduced in the relationship between infants and caregivers because they serve a certain function, then the infant's reaction is set up to represent cases where these behaviors fail to be performed by the caregiver.

Such an account captures the results from developmental psychology, which suggest that jealousy occurs early in infancy, without denying that jealousy involves a sense of the self and the ability to understand others' intentions and feelings. Jealousy can still be seen as being about the violation of a rule, since being left out can be seen as a rule violation of the caregiver that is of high urgency for the infant's well-being. Yet as I suggested earlier, it can also be seen as a response to a frustration of desire that is not governed by rules. If anything, further studies might solve the question what jealousy really is about. In any case we need not ascribe an explicit understanding of social rules and norms to the infant. The infant is situated in a social environment and constantly experiences itself in relations with others, where it quickly learns to skillfully react to all kinds of social scenarios where it needs to negotiate its needs and intentions.

Such an account is open with regard to the question of whether jealousy is innate or learned and whether it has a biological or a social function. The present account takes emotions to be set up to represent a certain core relational theme. But this leaves open (at least to a certain degree) whether

the emotion acquired its function evolutionarily or got set up only during infant development in the social context. Given that I take emotions to be embodied, every emotion certainly involves bodily reactions or basic abilities that have their origins in evolution. But these basic reactions or abilities can be combined and directed to represent a core relational theme during development. The final answers to questions about which parts of emotions are evolutionary acquired, innate, and/or hardwired can only be given by further empirical studies, if they can be given at all. Yet I demonstrate in the following section how to think about the development of an embodied action-oriented representation that steadily develops in most cultures in infancy, though it might not have direct roots in evolution.

9. THE GUILT-ACQUISITION MYTH

What I have said so far about prereflective self-consciousness and the rules and norms that can be established through social practices suffices to develop a sketch of how guilt can be explained along the lines of my account. The example of guilt helps to further clarify the commitments of the theory at hand and its explanatory power. What I develop can be labeled as the "guilt-acquisition myth," since it is more an educated guess then a hypothesis that follows directly from empirical studies. As mentioned earlier, there are very few empirical studies on the onset of so-called self-evaluative or higher cognitive emotions. Yet consensus has it that guilt can only be present after the infant has acquired an explicit sense of self and an explicit understanding of social rules and norms (Lewis 2014). This is a theoretical assumption that is not well grounded, since guilt can be thought of as an embodied action-oriented representation that detects rule violations in a familiar social setting. Such an account replaces cognitivist assumptions about the onset of guilt without reducing guilt to a mere behavioral disposition.

Guilt is an emotion that cannot be observed in newborns' expressions, as, for example, joy and anger can. It is furthermore certainly not too speculative to say that signs of a guilt reaction in infants younger than nine months are rare. Joint attention seems to be a necessary condition for its onset, since the scenarios in which guilt is triggered usually involve the caretaker somehow referring to something in the world that the infant should or shouldn't have done. It is also a reasonable assumption that guilt requires the ability to show empathy or other-relatedness. The first signs of empathy, that seem to be more than mere emotional contagion, can be observed by the age of twelve months when infants start to, for example, gently touch an apparently distressed caregiver.

What has to be presupposed to explain the early onset of guilt is that there are stable social rules in the social environment of the infant that sanction unwanted behaviors, like "Cruel behavior is wrong!" While shame in

its most basic forms seems to be associated with inappropriate behaviors in rank hierarchies, guilt seems to be more concerned with rule violations that harm other people's well-being. What needs to be in place are recurring ways to make the infant experience that it violated such a harm-rule. Parents usually raise their voices in these cases, they put on a stern or angry face, and they might also react with love-withdrawal or rejection. The infant, on the other hand, is well equipped from early on to make sense of the bodily postures, gestures, and facial reactions that her caregiver shows in the inter-action. What she has to learn is to steadily detect her own rule violations through these reactions and react appropriately herself.

Given the theoretical framework developed earlier, these assumptions can be put together to form the guilt-acquisition myth. Suppose that there is a recurring scenario where a caregiver reliably punishes a child when she shows cruel or harmful behavior:

1. The child steadily experiences distress, fear, sadness, or other more basic bodily or emotional reactions as a response to the punishments.
2. The child learns to reidentify these scenarios as cases in which she violated a rule.
3. The child learns via association or trial and error that showing recom-pensive behavior aids reintegration.
4. As a consequence, the child in the future will not (or not only) feel distress, sadness or fear when being punished but will feel motivated to make amends for what she has done.

My point is that at the moment where the child reliably reacts emotion-ally to situations where it is accused of having transgressed a social rule not only with distress, fear, or sadness but is also motivated to make amends, there is a full-fledged guilt reaction. We do not need to think of guilt as being a fully innate program or an evolutionarily hardwired reaction (which we have no concrete evidence for), and we do not need to deny that guilt is con-cerned with normative affairs (which would be inadequately reductionist).

What I suggest instead is that guilt is an emotion that steadily devel-ops in a certain social context because it serves a social function. Infants can acquire a kind of social "know-how" via emotional learning in social interactions as they learn to deal with their own needs and intentions in relation to others' feelings and concerns. These processes of interpersonal negotiation are governed by social rules that the infant can internalize with-out conceptually understanding them. Note that this account is open to the question of whether guilt is evolutionarily hardwired or develops out of more basic abilities in a certain social setting. Guilt is set up to represent that we have violated a social rule by harming others. But whether it owes its function to evolutionary or to constantly reoccurring demands of a given social context remains an open empirical question. Guilt is an embodied action-oriented representation through which the infant detects that she did

something wrong and should make amends. Guilt is embodied, since an infant represents her own wrongdoing, not through abstract thinking, but through the bodily arousal that constitutes the emotion in question; and it is embedded, since what allows for the claim that infants can represent social rules and norms on a nonconceptual level is that infants permanently interact with a socially structured environment through which normative rules can have a causal impact on the infant's behavior.

Given that guilt is such an embodied action-oriented representation, it can certainly be said to be subject to semantic norms. It can be appropriate or inappropriate depending on whether the represented core relational theme is present or not. Guilt is also subject to social norms, since guilt represents nothing else but the violation of social rules and norms. I come back to the question of how to account for rational norms in the case of guilt and other emotions in the next chapter.

10. CONCLUSION

In this chapter, I have argued that emotional beings are embedded into a structured environment that is full of emotion-relevant information through which we represent core relational themes. Relational properties occur in the biological and social environment of the organism. This claim depends on an ontological commitment to biological values that exist in relation to the organism and a commitment to the function realist view that takes social rules and norms to be established and reinforced through social practices that steadily recur and can easily be grasped by an individual without conceptually understanding them. Given a conception of a prereflective embodied sense of the self and the various abilities of infants to understand and interact with their social environment, such a view can describe social emotions such as guilt and shame as embodied action-oriented representations.

There are three main questions that have to be answered in what follows. First, relational properties, as they have been introduced by now, appear to be fairly cheap in that the world is full of relational properties, such as "being dangerous" or "being indigestible," yet emotions respond only to some instances of them. The table in front of me is indigestible, yet I do not react with disgust; the kitchen knife is dangerous, yet it doesn't frighten me, nor does the hot plate, the mixer, or other items in my household that could cause harm. Does that prove that the extension of disgust is determined by the intension, that is, by the organism's mental capacities to grasp these properties? The answer is no. What is true is that emotions include biologically established mechanisms and that these mechanisms are not only extremely fallible but also *restricted* in the ways in which they can access a property in the world. People can die peacefully eating sweet-tasting poison without ever feeling disgusted. But that doesn't mean that the extension of emotions is determined internally. Externalism claims that the extension of

Embedded Emotions and the Ontology of Core Relational Themes   149

a representation is determined externally and to decide whether something is a misrepresentation, you must look into the world and not into a person's head. By making that claim, an externalist need not deny that an organism's capacity to grasp certain properties is restricted in many interesting ways, which determine when and how the organism can represent the property in question. To claim differently would be to assume a crude version of the "myth of the given." It is not my intention to argue that properties such as "being dangerous" or "being indigestible" are simply "given" to us via emotions whenever they are present. Of course, we can only represent them under certain circumstances: they have to strike us in a way we are acquainted with. Rather, the claim is that the properties exist independently of our representing them and that it doesn't take complex cognitive machinery to represent them. To come up with a more detailed explanation of how we grasp core relational themes in certain situations and not in others, I develop an action-oriented view of emotions in the next chapter that holds that we have the skills to respond to certain instances of relational properties but not to others.

Second, I argued, contrary to Prinz, that emotions directly motivate action and that the motivating potential of emotions stems from the bodily arousal they involve. Yet how can the present account integrate that motivating potential? I answer both questions in the following chapter, arguing that emotions are not only embedded and embodied but also action-oriented. I do so by suggesting that emotions are about affordances. Introducing affordances into the theory clarifies the role that bodily abilities play in determining to which relational properties we react emotionally. Emotional affordances, as I describe them, capture the motivating potential of emotions.

The final question is how to account for the emotions being subject to rational norms without treating them as part of a holistic conceptual reasoning system. I answer this question by sketching a theory of how emotion, as a faculty, constitutes a skillful ability that enables us to interact with the social world on a nonconceptual level.

NOTES

1   To avoid terminological confusion: Most traditional positions that go under the label of "moral realism" can also be labeled as "cognitivism" about moral facts insofar as they claim that there are moral judgments that refer to facts (there is a strong trend to hybrid theories though). Cognitivist positions are distinguished from noncognitivist positions such as "expressivism" claiming that (1) there are no moral facts and (2) therefore moral judgments cannot be true or false. They might express approval or disapproval but do not refer to any facts. The cognitivist–noncognitivist distinction in the emotion debate is certainly not completely unrelated to this debate (cognitivists about emotions as well hold that emotions are cognitive states with truth conditions while noncognitivists claim that this is not the case), still one can be a cognitivist

with regard to moral judgments and a noncognitivist with regard to emotions and the other way around. When I talk about (non)cognitivism throughout the book I usually refer to (non)cognitivism with regard to emotions.

2   See Chemero (2009) for an overview of studies on direct perception and natural information.

3   The same problem holds for the ecological account of Turvey, Shaw, and Mace (referred to as the Turvey–Shaw–Mace account), although they assume laws that hold only in an ecological niche. Though these laws are not universal, they must still be general (Chemero 2009).

4   See Neander (1996) for a modest interpretation of Dretske and Godfrey-Smith (1989), and Wild (2010) for further critique of Dretske's notion of information. I'm grateful to Chrisine Sievers and Markus Wild for clarifying discussions of Dretske's and Millikan's different notions of information.

5   Millikan also departs from Dretske in claiming that not all true or adequate representations carry information, because some are just true by accident. Only those that are acquired on the usual and reliable ways really carry natural information.

6   This objection has already been raised by Stephan, Walter, and Wilutzky (2013).

7   See Foot (2001) for an application of Geach's idea to virtue ethics and Wild (2010) for an explanation of natural norms that relies on Geach.

8   For similar suggestions as to what anger is about see Parkinson, Fischer and Manstead (2005), Parkinson (2001), and Frijda (1993).

9   Using the famous rouge test, where an infant is placed in front of a mirror with rouge on her nose, tests the ability of self-referencing. No child younger than fifteen months touches her own nose in response to remove the color, while at fifteen months approximately 20% do so. Given normal development 100% will touch their own nose at the age of twenty-four months (Lewis and Brooks-Gunn 1979).

10 Lewis also explicitly equates self-consciousness, that is, the capacity to consciously think about the self, with the capacity to feel or experience (e.g., Lewis 2006, 200f.). I take this to be a highly implausible equation since it suggests that infants only start to feel their emotions once they are able to think about themselves, but I won't argue against this view here any further.

11 As noted previously, the (only) empirical hint that Lewis and some of his colleagues established in several studies was the rouge test combined with a test of whether infants show certain emotions in certain situations. That this evidence is questionable is shown later.

12 Similar claims can already be found in the work of Gibson, who points out that perception always contains proprioceptive information and therefore information about the self: "All the perceptual systems are propriosensitive as well as exterosensitive, for they all provide information . . . about the observer's activities. The observers movements usually produce sights and sounds and impressions on the skin along with stimulation of the muscles, the joints, and the inner ear" (Gibson 1986, 115). Several authors who argue for a pre-reflective self have used Gibson's theory (see, e.g., Bermudez 1998, Neisser 1988).

13 These theories traditionally assumed the false-belief task to be a crucial test for the infants' ability to understand others' minds. They therefore claim that the infant's ability to understand others' minds develops around the age of three. Whether actual versions of these theories are really opposed to Gallagher's views and not just complementary is a question I want to leave open here, since it would only be helpful for my argument if there were wide agreement on Gallagher's claims.

14 It does not matter for present purposes whether jealousy is hardwired or set up through learning. It is also not central to decide whether jealousy has a biological or a social function and because of which of these it becomes reproduced.

15 See Welpinghus (2015) about the social contexts in which jealousy can be appropriate or not.

16 This is not to say that guilt might not have fitness-enhancing qualities. It only says that to explain them we need to look at groups and social interaction, because this is obviously what guilt is essentially concerned with in contrast with, for example, disgust, which is an emotion that in principle can occur outside of a social context or in asocial animals.

17 For an argument as to how social structures impose constraints on our actions without directly causally influencing us, see Haslanger (1995, 2015).

18 Latour objects against the intentionality-relative view that the distinction between the natural and the social world is entirely artificial and breaks down if we only think for a minute about the necessity of certain material conditions for maintaining something like money (Latour 2005). The function realist view can put more weight on these material conditions.

19 See Wild (2010) for a developed biosemantic account of normative categories in social contexts.

20 I follow Haslanger's (2015) description of social structure here. Haslanger does not apply the notion of social structure to animals. But as far as I can see, there is no in-principle reason why this should not be possible.

5   Loving Is as Loving Does
Embodied Action-Oriented Representations

1. INTRODUCTION

Emotions are embodied, action-oriented representations set up to represent affordances. In the last chapter I translated Lazarus's notion of core relational themes into relational properties that recur in the organism's biological and social environment. In Chapter 3 I suggested that an embodied view makes sense of the idea that emotions are intrinsically motivating because emotions are realized by patterns of bodily arousal that prepare for action. In this chapter I develop the view that emotions are action-oriented. The action-oriented view takes emotion to be a skillful ability, that is, a faculty that is set up evolutionarily but only fully develops its skillful ways of representing core relational themes through an ongoing process of interaction with the social environment. When emotional, we not only represent a core relational theme by passively receiving information about it, but we are prepared for a reaction at the same time, because the bodily reactions that constitute emotions are skillful. These skillful reactions constitute the emotions being intensional. When being afraid we do not simply represent a danger but a *danger-to-be-avoided*. Being afraid is a fundamental mode of access to the world; we primarily grasp what is dangerous through the bodily reactions that prepare for flight.

In what follows, I sketch the action-oriented view, and the critique of "the sandwich model" of the mind. I argue that emotions are a prime example of mental processes that cannot be adequately explained by the sandwich model and therefore fit better into an action-oriented framework. I then develop a theory of emotions' intentional objects by relying on the results presented in the last chapter about natural information and relational properties. Emotions, I take it, are about affordances. Affordances are relational properties that we grasp via skillful abilities, which is why representing them is intrinsically motivating. I further discuss what role the notions of skill, sensorimotor knowledge, and entrainment play with regard to emotions in contrast with perception. To conclude I explain how an embodied action-oriented approach can account for the normative structure of emotions.

2. NOT A SANDWICH

Traditional accounts of perception argue that perception begins with a stimulation of the sense organs. In the case of vision, such stimulation results in a retinal image. For perception to occur, various information-processing operations have to follow. The action-oriented view instead claims that perception is an act of the whole organism that cannot be reduced to passive sensation. John Dewey, in his essay *The Reflex Arc Concept in Psychology* (1896), argues that the coordination between eye and hand forms a well-organized system. The *act of looking*, executed by the whole organism, is what is primary and both sensation (i.e., the processing of the input) and movement are parts of this act, which constitutes perception. Action and sensation are closely intertwined; they "have been so often bound together to reinforce each other, to help each other out, that each may be considered practically a subordinate member of a bigger coordination" (Dewey 1896, 359). While the hand permanently depends on the control and stimulation of the visual information, vision in turn must be stimulated and controlled by the body's movements. If the eye is not kept fixed on the goal of grasping, if there is no proprioceptive feedback about the body's posture, the task cannot be fulfilled.

Dewey's description of an ongoing sensorimotor circuit is an alternative to the classical reflex arc idea that takes sensory stimulus and motor response to be distinct physical existences that have to be "somehow adjusted to each other, whether through the intervention of an extra-experimental soul, or by mechanical push and pull" (Dewey 1896, 365). Dewey's critique is directed against what would today be labeled the "sandwich model of cognition" (Hurley 1998, 2001). Classical models of cognition tend to see the mind as a "sandwich" with action and perception being two separate and peripheral slices, and cognition being the hearty filling. The mind, according to this model, is divided into *vertical* modules where each module performs a broad function and then passes the resulting representations on to the next level. First, information about location, color, motion, and so on is extracted from inputs by different streams of domain-specific perceptual processing. Different input processing streams produce representations that converge and are combined by perception. Second, the unified information proceeds to cognition, the central module that interfaces between perception and action. This central module is the host of all rational thought and deliberation, where rational thinking is understood as a processing of symbols in a unified code that are syntactically structured and show the features of compositionality, systematicity, and so on. Third, based on current and stored input and cognitive processing, a motor plan is developed and passed on to motor programming processes to be executed. The relevant causal processing is one-way linear. External information stimulates the sensory systems and the resulting perceptions can be the subject of higher cognitive processes and cognition can then trigger action.

The sandwich model and Dewey's critique of it are usually applied to perception. An example of a theory that adopts a passive picture theory of perception to explain emotions is Prinz's account developed in *Gut Reactions* (2004, 232f.).[1] Prinz adopts Dretske's theory of perception and compares emotions to Fodor's "modules." He intends to thereby clearly separate emotions from higher cognitive processes and provide evidence for his noncognitivist approach. Yet at the same time Prinz not only separates emotions and the bodily arousal that constitutes them from higher cognitive processes but also from action tendencies. Prinz's approach has the highly implausible consequence that emotions are constituted by bodily arousal yet this bodily arousal has no motivating function in itself: it is not what pushes us to avoid or approach certain situations. Prinz introduces further neural evaluations to explain the valence of emotions and these evaluations then cause action tendencies.

Contrary to Prinz I argue that emotions are a prime example of why the classical sandwich model of the mind is inadequate. With regard to perception, Hurley argues that adherents of the sandwich model of the mind confuse the personal and the subpersonal level: they project their own experience of perception and action as separated faculties onto subpersonal processing. Yet even if it is true that we experience action and *perception* as separated faculties, with regard to emotions, the sandwich model is implausible even from a first-person point of view. In the first place, it is counterintuitive that an emotion should be a mere perception of a dangerous or offensive situation that, only after further processing, is evaluated as being good or bad, which then triggers a certain pattern of bodily arousal and behavioral reactions. Emotions have an intrinsically motivating character that structures their phenomenology (see Chapter 3). In shame we feel that we want to vanish, in anger we feel like exploding, and when in love we feel the urge to be near to the beloved. Taking emotions to be action-oriented representations is the most straightforward explanation for this phenomenon.

Arguments for an action-oriented approach, however, are not restricted to the first-person perspective; they can be complemented with empirical data from various sources. From a psychophysiological point of view it can be doubted whether the input and output side of an emotion can be clearly separated. Several studies suggest that emotional behavior (e.g., bodily postures, facial and vocal expressions), which is commonly regarded as a mere output, can modulate emotions and sometimes even trigger them (see Niedenthal et al. (2005) for an overview). To test the influence of facial feedback on emotions and emotional feelings, for example, participants were given nonemotional tasks to guide the production of facial expressions without cueing the emotional meaning of the expression. Participants then had to evaluate certain stimuli, like cartoons, with regard to whether they were funny. Findings suggest that the intensity and quality of the participants' manipulated facial expression affected the intensity of their self-reported emotional feelings as well as their autonomic

responses. Facial expressions modulate emotions and sometimes even trigger them (Laird 2007). When facial feedback is reduced through Botox injections, finer notes of emotional feelings tend to be reduced (Davis, Senghas, Brandt Ochsner 2010, Chapter 2). Bodily postures seem to have an impact on emotional experience in a similar way. In a study conducted by Stepper and Strack (1993), participants' bodily postures were manipulated by asking them to adopt one of two conventional working positions, an upright or a slumped posture. Participants then received positive feedback concerning their performance on an achievement task. Those who received success feedback in the slumped posture felt less proud and reported being in a worse mood than participants in the upright position (for similar results see Duclos et al. 1989, Laird 1974). A different study found that an emotion-specific tone of voice has been found to amplify emotional experience as well (Hatfield et al. 1995).

Taken together, expressive behaviors facilitate, modulate, and produce corresponding emotional reactions and feelings. This suggests that in emotional processing input- and output-processing pathways are not clearly separated. Rather, emotional processing works in feedback loops: the organism produces output that affects its input in a systematic way. It not only receives stimuli from the world but also is a source of stimulation to itself. Getting angry in a certain situation might not depend on the input received from the external world alone. The organism's entire situation including bodily posture and current facial expression plays a role as well.

There is also neuroscientific evidence that Hurley's view of the mind's organization fits emotions better than Fodor's and Prinz's views. To apply Fodor's notion of modularity to emotions, Prinz argues that emotions can be seen as domain-specific, if only the initiation and response pathways are understood in the right sense. The amygdala receives input from various regions of the brain and initiates patterns of bodily outputs. The amygdala is part of an emotion's initiation pathway while the bodily responses and their perception, according to Prinz, form the response pathway. While the initiation pathway is not domain-specific, since it can receive all kinds of inputs, the response pathway is domain-specific: the only thing that contributes to the emotional perception and its feeling is bodily arousal.

> Inputs to the emotion initiation pathway can vary dramatically, and they can change through learning and experience. The amygdala, and related structures, is not domain specific. But the emotion response pathways are domain specific. They respond to bodily changes and core relational themes.
>
> (Prinz 2004, 234)

Unfortunately, Prinz doesn't say what or where these response pathways are, which would make it easier to verify or falsify his hypothesis. Support for Prinz's argument is found in Damasio's claim that emotion-feeling cycles

begin in the brain with the appraisal of a stimulus, triggering a process that spreads over brain and body, building up the emotional state. In the case of fear, the nuclei in the amygdala dispatch commands to the hypothalamus and brain stem that result in several parallel actions. Heart rate and blood pressure change, as do the respiration pattern and the state of contraction of the gut. What Damasio calls the "body loop" is completed through humoral signals (chemical messages conveyed via the bloodstream) and neural signals that constitute bodily feedback and lead to a change in the neural representation of the body from the brain stem up (Damasio 1999, 2010, 111).

Yet while many researchers agree that the amygdala is necessary to trigger fear and that the bodily feedback involved in fear is processed in other parts of the brain, it is rather unlikely that the role of the amygdala is restricted to input processing and that it plays *no* role in registering bodily changes, in producing feelings, or in the reappraisal of ongoing emotions. There are studies that suggest that the amygdala can alter attention and awareness by modulating the hippocampal memory system (Phelps 2004). Furthermore, the amygdala not only receives crude sensory input quickly via a subcortical route, but it can also provide feedback to perceptual systems in the presence of emotional stimuli; the connection to the sensory cortex is reciprocal. Through projections back to the sensory cortical regions, the amygdala may enhance further perceptual processing in the presence of emotional stimuli, resulting in an overall heightened perceptual vigilance (Phelps 2004; see also Sneddon 2006). Prinz's claim about emotions' being processed in a one-way linear fashion is not very convincing. Also the idea that emotional feelings should be generated by bodily feedback *alone* seems rather unlikely. It is more plausible that, while bodily feedback might constitute the core of emotional feelings, activation of attention in fear through the amygdala is probably mirrored on the phenomenological level as well. Being highly focused on the situation that is dangerous is part of how fear feels.

To sum up, the data from psychophysiology and neuroscience discussed so far suggest that emotions cannot be adequately described using a passive picture model of perception, as suggested by Prinz, where the processing is taken to be one-way linear (see also Colombetti 2014). The claim that emotional processing starts with a neural appraisal (e.g., in the amygdala) and triggers bodily arousal, which is then felt, ignores the multiple feedback sources in an emotional process. The sensory apparatus and the amygdala stimulate each other, bodily feedback can shape the appraisal, and an emotional feeling is not only constituted by bodily arousal but also by e.g. altered attention.

In the following section I develop an approach that takes emotional content to be external but simultaneously shaped by proprioceptive feedback from the organism. This account relies on Gibson's theory of affordance perception and defines emotional content as descriptive and motivating at the same time.

3. EMOTIONAL AFFORDANCES

I argued in previous chapters that core relational themes should be translated into external properties that exist in relation to the organism. The claim that emotions represent relational properties and the claim that emotions are action-oriented can be brought together if emotions are taken to represent affordances. The relational properties that are represented in emotions can be described as *affordances* in a Gibsonian sense, although Gibson only applied the term to perceptions, and emotions differ in several ways from perception. Gibson assumes that perception always involves proprioception[2] and thereby is fundamentally observer-relative, although the external information that is picked up through perception is assumed to be real. This also gives a good *prima facie* understanding of the relation between world-directedness and motivating potential found in emotions. Gibson's concept of affordances is therefore introduced and applied to emotions in the following.

Gibson introduced the term of affordances to account for what human and nonhuman animals perceive. He explains the term as follows:

> The affordances of the environment are what it *offers* the animal, what it *provides* or *furnishes*, either for good or ill. The verb *to afford* is found in the dictionary, but the noun *affordance* is not. I have made it up. I mean by it something that refers to both the environment and the animal in a way that no existing term does. It implies the complementarity of the animal and the environment.
>
> (Gibson 1986, 127)

Affordances are properties of the environment that have to be seen as relative to the animal. For example, for a particular organism, certain fruits look edible; for persons of a certain size and shape, certain objects look sit-upon-able; others, stand-upon-able; and so on. An affordance is an objective external feature of the environment, yet it only exists with its specific value *for* a certain animal. A certain mushroom might be poisonous for one animal but nutritious for another. A lion might be a possible mate for a lioness, but a predator for a zebra. As Lazarus remarks, emotions are always concerned with the relations between the individual and the environment: they have a relational meaning. Yet Lazarus further argues that the individual can only grasp these relations through a complex process of cognitive evaluation, while Gibson argues that affordances themselves have a value for the organism and that that value can be directly perceived:

> The perceiving of an affordance is not a process of perceiving a value-free physical object to which meaning is somehow added in a way that no one has been able to agree upon; it is a process of perceiving a value-rich ecological object.
>
> (Gibson 1986, 140)

The perception of the environment is always laden with the perception of values concerning the organism's goals and its well-being. Applying this line of reasoning to emotions forces us to see core relational themes as entailing motivational power or an action-orientation. Fear is not simply about something's "being dangerous" but rather about a *danger-to-be-avoided*. The emotion refers to a relational property, the mode of presentation consists in an action orientation that makes the property appear a certain way, in the case of fear something dangerous appears as something to be avoided. The ontological background for the claim that emotions represent properties that are of value for the organism I developed in the previous chapter. Yet it doesn't explain in which sense emotions are directly motivating and why we only react to some instances of core relational themes and not to others. This is what the notion of affordance adds to the story.

Relational properties exist independently of whether a single organism actually perceives them and independently of whether the whole species in question is able to detect them at all. While the former is true for affordances as well, the latter is not. Affordances cannot be specified independently of the abilities of the organism to detect and react to a feature. While something can be dangerous for an organism even if the organism is not able to detect this property, something can only be a "danger-to-be-avoided" if the organism has the ability to represent and react to the affordance. Affordances belong to animal–environment systems and in order to exist depend on the abilities of the organism to detect and interact with the environment.

Gibson's original term, however, has been interpreted in several ways. Some argue that affordances are relational properties that establish selection pressure to which the organism then learns to respond (Reed 1996). Others take a different stance and argue that affordances are dispositional properties of the environment complemented by dispositions of the animal. Something can be walk-upon-able only if the animal already has the ability to walk. Therefore, affordances cannot be seen as establishing selection pressure. The abilities they might bring about must be already in place for the affordance to exist (Turvey 1992).

I agree that it is misleading to describe affordances simply as relational properties that establish a selection pressure, since they must be the result of a selection process. Affordances are properties of the environment as they can come to be perceived by the organism. Yet it is still true that affordances are constituted by relational properties that can be detected by the animal because it is valuable for the organism to have the ability to detect these properties. The relational properties must be in place before the organism acquires the ability to detect them. Relational properties can *become* affordances when the animal acquires the ability to detect them because of a selection pressure.

Some argue that the environment of an organism should be understood as the set of affordances the organism is able to perceive. Organism and environment are then described as complementary: what the environment

contains and what the animal can perceive are interdependent. In accordance with what I have said so far, the term *environment* is defined more broadly. The environment of an organism should be described in relation to the animal and in relation to what is of value for its survival. For example, the environment of a tick and a human are therefore very different. Yet our interaction with the environment can have a causal impact on the tick's environment. Such an interaction can create new selection pressures for the tick that it is not able to detect but that still have a causal impact on it. If the tick were to develop a new detector, for example, smells that would help it to avoid humans who can wield pesticides, then that smell would be an affordance. Yet pesticides are dangerous features in the environment of the tick even if no tick ever learns to detect them.

Unlike relational properties, affordances can only be specified with regard to the abilities of the animal. This is the fundamental insight of Turvey's interpretation of the term. Yet it is important to highlight (contrary to Turvey) that affordances correspond to *abilities* in the animal, not *dispositions*. When coupled with the right enabling conditions, dispositions are guaranteed to become manifest. The soluble solid sugar will always dissolve in water in suitable conditions. This is not true of abilities. Contrary to dispositions, abilities are properties that can only be ascribed to agents. An agent can be able to walk and the walking conditions can be ideal but that does not imply that the agent will walk or that she will not fall down. Abilities appear to have a function that they can fail to perform. These functions depend on an individual animal's developmental history or the evolutionary history of the species, both of which occur in the context of the environment (Chemero 2009).

The notion of affordances adds something fundamental to the story of how we represent the world when being emotional. Thus far, I have argued that we can represent relational properties because they directly strike us in our biological and social environment. Additionally, I have argued that we represent the properties in question as being a certain way because of our own abilities that determine the way the representation is given to us. The bodily reactions involved in an emotional reaction determine how we grasp the property in question. When being afraid we not only perceive a danger but a danger-to-be-avoided due to the fact that our body prepares for flight and this bodily reaction constitutes the emotion.

Emotions represent affordances through the bodily reactions they involve and are therefore about normative properties and motivating at the same time. Emotions being about things of value and their capacity to represent social rules and norms can best be explained by the idea that core relational themes are relational properties that are normative in themselves. Core relational themes are of value for us whether we represent them or not. Yet in an emotional reaction, their value is directly translated into a bodily reaction that motivates further behavior. Emotions are set up to represent core relational themes by directly preparing the organism for action.

The notion of affordance also offers an explanation as to why we do not react to all the relational properties around us that belong to the core relational theme in question. As I argued in the last chapter, relational properties are "cheap" in the sense that they occur everywhere around us and we only react to some instances emotionally. Emotions might be set up to represent certain properties, such as being dangerous, but they cannot enable us to represent *all* things that are dangerous since natural information about dangers can occur in an infinite number of different forms. Gibson explains why animals sometimes do react to affordances and sometimes do not with reference to "potential." What affordances offer to the animal is contingent upon the particular circumstances. The lioness affords a danger to the zebra. This doesn't mean that the lioness will attack the zebra, but an attack is possible and under certain circumstances (e.g., the lioness being hungry) highly likely. For the zebra to perceive the lioness as something that should be avoided does not mean to, first, perceive it, then draw the inference that it is dangerous, and then prepare to flee. Instead, the zebra directly perceives the lioness as dangerous and is immediately motivated to flee (Scarantino 2003). According to Gibson, this sort of perception consists of picking up certain abstract patterns in the ambient energies that strike the organism; those patterns then guide the organism's behavior directly.

Gibson's radical claim is that certain variants in the energies impinging on the active organism carry information about the relations of significant distal affairs and guide motions to make use of these affairs. Natural information is external; according to Gibson we can directly perceive it with no representations needed, not even inner processing of information. I argued in Chapter 3 that this radical denial of information processing and representation is already hard to defend on the level of perception. When it comes to emotions, it makes no sense to deny that they are representations, since emotions have a normative dimension and, at least in human adults, they play a complex role in a larger cognitive architecture. Still, emotions are embodied action-oriented representations that represent affordances.

This is a minimal commitment to representations that does not think of representations as context-independent inner states that are realized by the brain alone. Such an account is largely compatible with Clark (1997) but also with Millikan's notion of pushmi-pullyu representations (see also Scarantino 2014). Both Millikan and Clark argue that there are simple representations that directly motivate for action and are nonconceptual, yet these reactions should still be described as representations with a distal content. While Clark highlights that the vehicles of a representation can be spread over brain and body, Millikan highlights that all simple representations involve a descriptive and an imperative side. Millikan also points out that affordance representations can be anticipative because of their history and function (Millikan 1995, 2004).

With regard to the claims proposed by cognitivists that were discussed in Chapter 1, seeing emotions as affordance representations is a promising

account. It allows us to explain emotions' intentionality as well as their intensionality. As I argue further in the following section, categorizing emotions as intensional doesn't presuppose conceptual background knowledge. Instead, it presupposes the sensorimotor learning history of a species as well as of an individual organism. Emotions are not simply constituted by hardwired bodily reactions; they are rather constituted by sensorimotor skills that prepare an organism for action. Emotional objects are given to us in a certain mode of presentation: we do not see something as being dangerous but as a danger-to-be-avoided. But that mode of presentation is not constituted by conceptual background knowledge but rather by sensorimotor skills that prepare for action.

4. REPLACING APPRAISALS WITH SKILLFUL ABILITIES

The obvious remaining question is whether the theory, outlined above, of emotions as embodied action-oriented representations of affordances can account for the seeming cognitive complexity of emotions present in features such as their normative assessability. In this section, I further develop the notion of emotions as skillful engagements with the environment and argue that such an approach can capture the main cognitivist claims in noncognitivist terms. To begin, I clarify the notion of skill and apply it to emotions.

A skill is a kind of embodied practical knowledge, such as knowing how to ride a bike. Some argue that speaking of "skillful knowledge" is an over-intellectualization of perception (Hutto 2006). I think this critique rests on a misunderstanding of what a skill or skillful knowledge is. As abilities skills are properties of agents. They enable living organisms to show certain reactions in certain conditions. The notion of ability is broader than the notion of a skill. While abilities can include hardwired responses such as the blinking reflex, the notion of a skill assumes learning and training. Yet to train or to learn something doesn't necessarily imply that any kind of conceptual or representational guidance is involved in acquiring a skill. Learning to hold one's own head up requires the acquisition of a sensorimotor skill and it takes a lot of training, so does learning how to crawl or walk. Yet training here doesn't mean that the individual consciously aims at learning something or that it learns something for a reason. Infants usually learn to hold their own heads up before the age of one. Yet infants are not able to hold their heads up at birth, as they are able to breathe, blink, and digest. Breathing, blinking, and digesting are abilities, not skills. Although breathing is not a skill itself, there are skillful ways of breathing ranging from panic induction to the breathing techniques mastered by divers and opera singers.

Enactivism makes the interesting claim that perception is not only an automatic input-generating process but is a skillful way of gaining nonconceptual knowledge. To make sense of sensory information we have to know how things look from a certain angle and how their size might change when

we move closer. This is not a kind of propositional knowledge that we apply to raw sense data; it is a type of sensorimotor, or skillful, knowledge that guides the way we act during perception and is involved in the constitution of perceptual experience.[3] To simply call perception a skill ignores the fact that perception is an innate ability with the biological purpose of accessing information in a nonconceptual format. I therefore call perception a "skillful ability," that is, an ability that is set up to gain action-oriented representations with nonconceptual content and that is shaped by learning in ways that are fundamental to its proper functioning.

In this sense, emotion is a skillful ability, too. Emotion entails biologically established bodily reactions with a certain purpose, yet before single emotions can successfully fulfill their purpose (e.g., to prepare the organism for flight in the case of fear) many things have to be learned, and can be shaped in different ways by different environments (being able to flee for example has to be learned, as do the reactions to most triggers that signal danger; both can vary in different environments). In emotion, just as in perception, the boundary between skill and ability is blurry. Apart from the terminological confusion that this causes, the problem of drawing a boundary is an interesting phenomenon in itself. Being able to have emotions is an innate ability that begins to function at birth. Yet this ability needs to be unfolded in social interaction, and it gains an individual shape over a learning period. People get angry at rather different occasions, and the set of bodily reactions involved in an anger reaction can be established in different ways as well.

It should be noted, however, that emotions are skillful in a somewhat clumsy way. We should not think of the skillfulness of emotions in comparison with a skillful piano player that masters the movements of her fingers on a microlevel. Emotions essentially bear with them a primitive core of deeply embodied reactions that makes emotional reactions in modern societies often look somewhat heavy-handed or even totally inappropriate. Fear before an upcoming exam is usually more a stressful distraction from one's rational goals than a functional danger-warning mechanism. Jealousy can damage relationships rather than secure attachment, and uncontrolled anger often does more damage than it helps to overcome restrictions. Yet the consequence of this is neither that emotions are not subject to biological and social norms nor that they are not skillful. It rather shows that in emotions we find a complex entanglement of responsiveness to biological and social norms that stands in complicated and often paradoxical relations to our rationally held goals.

Emotions are biologically established abilities with a certain purpose that are highly plastic, meaning that the input and the output side of an emotion can be shaped differently among individuals and cultures. Basic emotional reactions are embedded in a social context, where they can change their function and content in response to the pressures and options this social context is offering. The notion of skill is meant to account for the way in which this shaping takes place. Consider the examples of anger and guilt

from the last chapter. Anger represents restrictions or, rather, restrictions-to-be-fought. What the affordance concept adds to this view is that the content of an actual anger representation is constituted by the arousal and the sensorimotor reactions that prepare the organism for a fight. An anger reaction might be present early on, yet it can only be present at the moment where the infant is able to struggle against restrictions. While infants that are only a few weeks old show such aggressive reactions in appropriate situations, guilt emerges significantly later. As I argued, guilt can be present in infants before they have an explicit understanding of social rules and norms. But, according to my account, guilt represents a rule-violation-to-make-amends-for. Only infants that are able to empathize and show recompensive behavior can grasp such content. This behavior is first observed in infants that are about a year old. In the following section I discuss in detail how an embodied action-oriented account can replace cognitivist claims, such as Lazarus's appraisals, without thereby reducing emotions to meaningless feelings or reflex-like reactions.

I have argued that anger represents restrictions. On a basic level, infants struggle against bodily restrictions, such as too tight embraces, and show aggressive reactions when they cannot get through to a goal. The action-oriented view adds to this description that the bodily arousal involved in anger constitutes the content of the reaction: anger does not simply detect restrictions but represents a restriction-to-be-fought. The content of this representation is determined by the relational property in the world *and* the proprioceptive feedback stemming from the bodily arousal that prepares for a fight. Anger has the function of representing restrictive properties and preparing for an appropriate response at the same time.

This approach accounts for the main cognitivist claims in noncognitivist terms: anger is intentional and it has an external content, namely, a restriction. Whether a restriction is present or not determines whether it is semantically appropriate. It is intensional, insofar as it presents restrictions in a certain way, namely, as something against which to fight. An emotion has intensional content when it provides a way of seeing an object and there are different ways the object can be seen. A good example of different ways of seeing one object is found in fear reactions. According to Lazarus, fear differs from anxiety. Fear represents a concrete danger in a certain situation, while anxiety represents an uncertain existential threat. According to the action-oriented approach, these two different emotions respond to the same relational property, namely, being dangerous. While fear represents a danger-to-be-fled, anxiety represents a danger to hide from. Fear motivates to flee a situation, while anxiety is associated with freezing reactions. The two emotions therefore differ on the level of affordances even though the relational properties they represent appear to be the same. Identical relational properties need not have identical affordances. Danger is an abstract property that can concern the organism in various ways: a-danger-to-be-fled is a kind of danger that can be avoided by flight, while

a-danger-to-hide-from is not a danger that can be avoided by fleeing and would therefore cause a freeze reaction. Different skillful reactions to relational properties thereby account for the way emotional objects appear to us in emotions; they account for the emotions' intensionality.

With the theory I have developed so far, Lazarus's appraisal account can be replaced. First, we need not assume that cognitive appraisals evaluate whether a present situation is relevant to our goals and, if so, whether it is congruent with our goals. To be incongruent with one's goals is an abstract description of what one represents when angry or undergoing any other negative emotion. Representing something as a restriction-to-be-fought implies representing it as being relevant to, yet incongruent with, one's goals. This can be explained with regard to what has been said about the nature of relational properties in the last chapter. The relational properties that come to be represented in emotions all concern properties of relevance according to biological and social norms. They all react to particular types of situations where a particular property, like "being dangerous" or "being indigestible," is present. Therefore, whenever the relational property "being dangerous" is instantiated, the relational property "being bad for the organism" or "being bad with regard to a social norm" is instantiated as well. Lazarus's first and second appraisals are abstract expressions of what an emotion is about, not appraisals that have to take place in order to come up with the representation of a core relational theme.

Lazarus further argues that a third cognitive appraisal concerns oneself and/or other people. Therefore, one must have a concept of oneself and others and must be able to understand the relations between self and others in order to entertain emotions such as anger that appear to evaluate another's behavior with regard to oneself. I argued in the last chapter that infants have an embodied sense of themselves as agents and various capacities to detect others' feelings and intentions. This suffices to explain an anger episode as a case where the infant struggles against something that is represented as a restriction. Later on, people gain explicit knowledge about themselves and their relations to others, which allows for anger reactions triggered by explicit and conceptual reasoning about others' behavior. Cognitive appraisals concerning the self and its relations to others are not necessary elements of anger but possible later cognitive refinements of the emotion.

Furthermore, Lazarus assumes that in anger we blame others and that, therefore, we must have a concept of responsibility for the actions one commits. One can only reasonably blame another person if one takes her to be responsible for her actions. I do not doubt that in some cases of anger we do blame others because we think that they are aware of and responsible for their wrongdoings. Yet this is not a necessary element of anger. Infants that struggle against restrictions are not alone in lacking this kind of appraisal. A caregiver can get angry because of the misbehavior of her child, while knowing perfectly well that she simply doesn't know any better.

A further appraisal concerns the kind of personal need, value, or rule that is violated by the person one reacts angrily to. What Lazarus requires for this appraisal is an explicit understanding of social rules, norms, and values. Yet I argued in the last chapter that even social norms can come to be represented on a nonconceptual level. In the case of anger, it is more likely that anger only responds to a direct threat to one's bodily well-being. Anger reacts to very direct restrictions and only gets directed onto social scenarios involving blame for certain rule-violations later on. Yet, even if this weren't the case, anger reactions would not need to involve cognitive appraisals involving social rules and norms.

Finally, Lazarus assumes that before we react with anger we evaluate the coping potential of the present situation and the future expectancies that aggressive behavior could have. The coping potential is Lazarus's notion for the kinds of action a situation affords. The action-oriented approach allows for a less demanding explanation of coping potential. The claim that emotions represent affordances already entails a coping potential. When angry, we not only represent a restriction but a restriction-to-be-fought. Coping potential can be directly perceived, since the bodily reactions involved in emotions are skillful; that is, they prepare for a certain kind of action. Just as the squirrel that perceives a tree as a climbing affordance perceives the coping potential that the tree affords, the infant that struggles against a too-tight embrace represents the embrace as a restriction-to-be-fought. On a cognitive level, coping potential and future expectancies would demand modal reasoning about possible future scenarios. Again, such reasoning certainly does play a significant role in adult emotion regulation but is not necessary.

Lazarus assumes that emotions presuppose a great deal of background knowledge that differs in individuals with different learning histories and accounts for individual emotion styles. Some people are afraid of dogs, while others are not; some people get angry quite often and easily, while others tend to remain calm. One way to account for this is to say that people appraise situations differently due to the background knowledge they have. This might have an impact on several of the appraisal dimensions that Lazarus lists. People might have different goals according to their background knowledge, they might have different moral values so that they blame themselves and others for different kinds of rule violations, and they might evaluate the expectancies of their own reactions differently. Some people might get angry in front of an aggressive dog, while others might be afraid because they evaluate their own ability to fight the animal differently. Again, I do not doubt that such background knowledge can come to play a role in the evaluations that are sometimes involved in adult emotions. Yet I think that the notion of skill allows us to assume an embodied sensorimotor background knowledge that is already present in infants and animals and remains central in adult emotions.

We have seen in previous chapters that bodily postures and facial expressions not only are outputs of emotions but can also trigger or facilitate these emotions. Apart from postures and expressions that people perform briefly when reacting emotionally in a certain situation, there are many kinds of bodily postures and nervous reactions that are involved in emotions and other reactions that can come to be habitual or chronic. Some people tend to look sad almost all the time because they usually trudge around with hanging head and shoulders and drooping corners of the mouth. Other people appear to be permanently tensed and "under pressure" and tend to react with aggression or panic for no reason. Of course, one could take these attitudes or character traits to be the results of how these people evaluate their life and situation in general, but one can also think of them as embodied habits constituting a sensorimotor background knowledge that facilitates some emotions and hampers others.[4]

That some forms of bodily posture or motoric activity may subtly influence a person's attitude is nicely demonstrated in a study done by Cacioppo, Priester, and Bernston (1993). Cacioppo et al.'s hypothesis is that arm flexion and extension is coupled with approach and withdrawal behavior, and therefore should have an impact on evaluative attitudes. To test this, Cacioppo et al. asked subjects to sit in front of a table and, in the extension condition, place their palms on the top of the table and press slightly so that they felt a light tension in their arms. In the flexion condition, subjects were instructed to place their palms on the bottom of the table and press slightly as well. An experimenter showed ideographs to the participants and they were asked to indicate whether they liked or disliked each of the figures. Results indicated that ideographs viewed during arm flexion were rated more positively than the ideographs to which subjects were exposed during extension. Cacioppo et al. concluded that arm flexion and extension activate different evaluative pleasant–unpleasant orientations. The implication is that arm flexion and extension serve as embodied valence markers that influence how we perceive and judge external stimuli.

Such bodily background knowledge not only facilitates the elicitation of emotions in different ways in different people, but also has an impact on how emotions unfold. Emotions are not mental states but psychophysiological episodes in which the organism negotiates its needs and interests with the environment. How an organism evaluates future expectancies can have an impact on the intensity and endurance of the emotion. For example, picture yourself in a conversation where you find the other person offensive. You might get angry yourself, yet there are different strategies for dealing with this feeling: you could try to remain calm thereby calm the other person as well, or turn aggressive yourself to make your point. While this is an example of a situation that appears to be governed by cognitive appraisals in Lazarus's sense, it is also a process governed by simple signs and bodily reactions. The offensive person showing a contempt-expression, for example, could heighten your own

anger significantly. People tend to unconsciously mimic the other person's posture in a conversation. Mimicking the posture of an aggressive person in a conversation could be another embodied factor with an impact on the unfolding of one's own emotion.

The question whether one reacts aggressively or frightenedly toward another's anger could also be influenced by an embodied sense of agency understood in a rather long-term sense. People perceive their own capabilities for action differently according to their size, weight, strength, and so on. This is certainly true for one's affordance perception of the stairs as climbable or of a stone as being liftable. Yet such a sense of agency could come to play a role in social scenarios as well. Reacting aggressively in front of a much larger opponent for example might be less promising than in front of a small person. Again, one need not assume that the person in question evaluates her size on a cognitive level. Instead, the embodied sense of agency and the direct perception of the other as an affordance to interact with can have an impact on how an emotion unfolds. In social contexts other factors are probably more central than the mere size of a person. In social hierarchies that we grow into, we can come to see people in Gibsonian terms as somebody-to-submit-to or somebody-to-control, because of the social nodes that people occupy in a social network and because of our embodied knowledge of our own social position. This way of social seeing and emoting is of course also shaped by education, where in Western cultures boys are, from early on, more strongly encouraged to show aggressive and dominant behavior than are girls.[5]

What this discussion shows is that, with regard to anger, Lazarus's list of cognitive appraisals that trigger a core relational theme can be replaced by embodied action-oriented representations occurring in a skillful organism and a structured environment. Cognitive appraisals do play an important role in adults later on, but they are add-ons rather than essential parts of emotions. An emotion is a skillful ability or embodied, normative know-how. Infants can acquire a personal way of skillfully reacting to the world via emotions that is highly influenced by the culture in which they grow up and yet does not involve any conceptual understanding of the rules and norms that become implemented in embodied habits.

So it seems that the embodied, embedded, action-oriented approach can account for all central cognitivist claims in noncognitivist terms. What cognitivists think of as a complex cognitive machine that must be in place before emotions can be realized is replaced by the interaction of the embodied agent with a structured social environment. I therefore suggest that Lazarus's taxonomy of emotions, along with their different appraisal structures, can be completely replaced by an account that takes emotions to be affordance representations, set up to represent the affordances listed in Table 5.1:

Though I do not discuss each emotion type in detail, the general strategy behind the approach should be clear from the discussed examples. Instead of detailed descriptions, I highlight some cases that could be seen as problematic for the present account.

Table 5.1   Core Relational Themes as Affordances

Emotion	Affordance
Fear	danger-to-be-fled
Anxiety	danger-to-hide-from
Anger	restriction-to-be-fought
Disgust	indigestible-thing-to-be-rejected
Surprise	unknown/unexpected-thing-to-be-explored
Joy	good-to-be-maintained
Sadness	loss-to-be-dealt-with
Love	attachment-to-be-maintained
Guilt	rule-violation-to-make-amends-for
Shame	rule-violation-to-be-hidden
Pride	achievement-to-show-off-with
Jealousy	Seeking-reintegration-from-being-left-out
Envy	another's-good-to-get
Embarrassment	faux-pas-to-hide

A difficult case is surprise. Surprise represents the affordance of an unknown- or unexpected-something-to-explore. Surprise is traditionally seen as an emotion that is special, insofar as it is neither positive nor negative. Descartes therefore suggests seeing surprise as a unique emotion; while all other emotions represent things as being good or bad, surprise is the only emotion with the function of detecting new or exceptional things (Descartes 1649/1988, §75). Furthermore, Descartes argues that, because it does not represent anything as being good or bad for the organism, surprise does not entail changes in heart rate and blood pressure but only in the brain (Descartes, 1649/1988 §71). One could therefore think that surprise has no valence because it lacks any underlying bodily profile. In the context of the present account this would lead to the conclusion that surprise is not an emotion.

Yet Darwin (1872/2009) and, later, Ekman (1971, 2003) point out that surprise appears to be associated with a panculturally present facial expression with eyes and mouth wide open. The expression has an obvious similarity to fear expressions. Prinz (2004, 153) speculates that surprise is a mild version of fear that lacks the distress that accompanies fear. I add that surprise is an emotion that not only shares a certain kind of excitement with fear but with several emotions, including positive ones such as joy. Instead, what surprise lacks are reactions such as tensed muscles and a release of adrenaline that prepare the body for fight or flight, as in fear and anger. Surprise is therefore restricted to a kind of intellectual excitement; its function is to prepare for further exploration of something in a way so similar to

fear that we should expect surprise to be accompanied by focused attention. Though surprise might not be valent in the sense that it always is positive or negative, surprise tends to be either positive or negative, and can quickly switch from surprise into fear or joy. I speculate that this is precisely because surprise already shares certain bodily reactions with these emotions and therefore facilitates their elicitation.

But if surprise is not valent in itself, can it be an emotion? Depending on what further studies on the bodily profile of surprise reveal, one possibility is to call surprise an embodied action-oriented representation that is closely related to emotions, yet not directly concerned with the organism's well-being, although it has the function to prepare for exploring new things. Another, intuitively more convincing option is to follow Descartes in his suggestion that surprise detects special things, without thereby accepting the claim that this is because surprise lacks bodily excitement. Instead, I suggest that among the embodied appraisals that represent good and bad things, we have a third kind of bodily appraisal that is set up to represent weird things.

I have suggested that fear and anxiety differ on the intensional but not on the intentional level, since they come with different sensorimotor patterns: one that prepares for flight and the other, for freezing. What might appear problematic is that the occurrence of the two emotions in adults motivates us to responses that are not reasonable. When we are afraid, it seldom makes any sense to freeze or flee the situation. In particular, freezing is a reaction that makes sense in animals that try to hide from a predator, but it is hard to come up with any scenario in our urban societies where freezing is of any help in mastering a dangerous situation.

Does it make sense to talk about an embodied skill here at all? I think it does. Talking about a skillful embodied reaction doesn't imply that the reaction in question leads to successful behavior with a statistically good outcome; it suffices that the ability in question was selected because it happened to serve a certain function in the circumstances as they were. Being able to deal with core relational themes such as danger and offenses is certainly still just as relevant for a living organism as it used to be. But the embodied responses that have their origins far back in our ancestors' history adapted to current environments in the most optimal way. The emotions plasticity is what still makes them skillful abilities. Yet acquiring skills doesn't mean coming up with the best possible solution but learning to deal with the circumstances that recur in one's environment. Different people might react with anger, fear, or anxiety in the same situation due to the experiences they have had and the skillful knowledge they have acquired. Freezing is kind of a "last resort" solution and might therefore be most common in people who have experienced many situations where they couldn't fight or flee. Acquiring skills in a social scenario is as much a process of tinkering as evolutionary development. We are not the engineers of our own early development. Rather we bring with us a package of evolutionarily acquired equipment and have to learn how to use it through experience in our social

environment. This is also the reason why emotional skills are clumsy skills. In current social environments not only freezing but also flight is a response that is seldom adequate, but nevertheless, we are bodily motivated to fly.

During development we come up with many complex forms of emotion-regulation some of which have the form of second-order bodily skills, others entail cognitive reappraisal, and, of course, all of these are closely intertwined and hard to disentangle. Think of Martha Nussbaum's example of how she usually experiences anger. Nussbaum states that she experiences anger not in the form of Aristotle's famous "boiling blood around the heart" but, rather, as a "tension at the back of her head and sometimes a headache the next day." Such a form of anger can be interpreted as a strongly regulated form of anger. Social education that can differ with regard to culture, gender, and class certainly plays a central role here. This can make the individuation of emotion types a fuzzy affair. Does a strongly regulated form of anger still motivate to fight a restriction? These are vivid problems that the present account faces, but I do not think that they speak against it. Individuating emotions is a fuzzy affair because emotions are fairly fuzzy psychophysiological processes that are intertwined with other bodily processes, skills, and cognitive evaluations. That strongly regulated forms of emotions challenge our categories does not show that we have the wrong categories but simply that emotions are difficult to categorize.

5. SENSORIMOTOR HABITS

As stated earlier, we need not think of emotions as being triggered by a mental file or governed by a neural program, nor do we have to assume that emotional stimuli are stored in one single "file" in the brain, or that the way in which emotions unfold is centrally governed by a neural program or a system of cognitive appraisals. Instead, emotion is a highly plastic and flexible ability; certain emotions can be triggered by what is usually their output, which can unfold in different ways depending on the sensorimotor skills of the organism and the structure of the environment with which the organism interacts. This claim depends on what I argue in Chapter 2, namely, that it is not the bodily profile alone that allows us to individuate emotion types but rather the functional mechanism that is realized (at least in large part) by the bodily reactions involved in emotions. Emotions can be individuated by their function, which is to represent certain affordances and prepare for action.

Given what we know about skills, what is the relationship between function and skill in the individuation of emotions? Sensorimotor enactivists have suggested a way to differentiate different perceptual modalities from each other: what differentiates vision from touch or smell is the structure of the rules that govern the sensory changes produced by various motor actions (O'Regan and Noë 2001). O'Regan and Noë label these rules

"sensorimotor contingencies" and argue that each modality is specified by a certain set of such rules, which enable an organism to master the modality. Visual information changes in a lawful way when we move our eyes or our head, when we blink or go around an object in front of us. Auditory information, by contrast, does not depend on eye movements or blinks but on rotations of the head, which change the temporal asynchrony between left and right ear, and approaching the source of a sound, which influences its amplitude but not its frequency. The laws of sensorimotor contingencies are determined by the fact that perception in a certain modality is a kind of action executed by a certain kind of apparatus. It is not determined by the structure of the apparatus alone.

Is it possible to assume sensorimotor contingencies in the case of emotions? Can different types of emotions be individuated with regard to their sensorimotor contingencies rather than with regard to their function? Imagine someone who is madly in love with a barkeeper and goes to the bar where she is working to ask for a date. The bodily reactions she will show in this situation are highly predictable. Her heart will beat faster when she enters the bar and sees the barkeeper; when going closer, her legs will turn to jelly and she might even tremble—her erratic movements reveal her nervousness. After finally asking for a date, what will happen depends on the reaction of the barkeeper. If she happily accepts, the excitement might switch into relief mixed with joy. The person madly in love might leave the bar and jump around, might have a feeling of warmth and relaxation, might run around with a huge smile for the rest of the day, or might go home to write dozens of cheesy, happy love songs. On the contrary, if the barkeeper rejects her, she will certainly feel sad and disappointed, maybe even embarrassed or ashamed. Her whole posture will change from the tensed and excited stance to a depressive mode with heavy limbs and a hanging head and jaw. She might feel like she would rather hide somewhere, or lie around doing nothing at all, or listening to sad and depressing love songs.

Emotional reactions, in certain situations, are highly predictable concerning the kind of bodily arousal they cause and the behavioral effects they produce. Sensation and bodily reactions are tied together in a way that produces typical emotional reactions in certain types of social scenarios that unfold in the interaction between the organism and its social environment. It therefore makes sense to compare enactive approaches of perception with emotions. Yet I have already argued in the third chapter that emotions differ from perceptions with regard to several respects, and this can be illustrated here again. What becomes clear from the example of the girl and the barkeeper is that emotions pick up on the valent aspects from social scenarios we are engaging in; they do not constitute access to the world in the way that perception does. Emotions rather highlight what matters.

Yet the usual argument against perception theory of emotion says that emotions are not direct, since they are never directly elicited through an external stimulus because they need to be preceded by a mental state of a

different kind. I reject the former claim but accept the latter. Emotions must be preceded by other mental states in the sense that an empty mind without perceptions, thoughts or memories could not be emotional, since there would be nothing to be emotional about. But the claim that emotions must be preceded by other mental states does not mean that emotions are caused by other mental states and do not stand in any direct relation to the world themselves. Emotions can be directly elicited through external stimuli. They are set up to represent relational properties and the organism receives information about the presence of relational properties via the bodily reactions that reliably co-occur with them.

Finally, on closer scrutiny, it is questionable whether the example of the girl and the barkeeper really demonstrates that emotions operate according to sensorimotor contingencies in the same way perception does. O'Regan and Noë suggest that visual exploration obeys certain *laws* of sensorimotor contingencies. They distinguish between sensorimotor contingencies that are induced by the visual apparatus (e.g., when we move our eyes, contours shift and the curvature of lines changes) and sensorimotor contingencies that are determined by visual attributes (e.g., the retinal image only presents the front of an object, the size of the object depends on the distance, and so on).

What could comparable examples for emotions be? Are there any features of the object such that at a certain distance, when opening the door to the bar where the adored barkeeper works, for example, our heart starts to pound? Does her appearance, her style, or her startling sense of humor stand in any kind of lawful relation to us such that we cannot stop staring at her, adoring her, and wishing to talk to her forever? Or is it that when we perceive her smiling to us in a friendly way, we are flushed with joy, but when she rejects our invitation our body temperature lowers, our posture changes, and our heartbeat decreases? Is it the objects of emotions or our emotional apparatus that brings such lawful relations into the story? Are our hearts hardwired to beat faster during social contacts every once in a while, or do our facial expressions permanently "match" the social setting around us? There might be *typical* reactions for somebody in love who encounters her beloved, but they are certainly not *lawful* in the same way as O' Regan and Noë suggest that perception is. Emotions, I argue, are not so much organized by sensorimotor *laws* as by sensorimotor *habits*.

In Chapter 2 I suggested that the corrugator muscle between the eyebrows can be seen as an embodied valence marker: the muscle appears to be tensed in the presence of unpleasant stimuli and relaxed in the presence of pleasant stimuli. This comes close to a sensorimotor contingency. The same can be said about increases of heart rate that appear to accompany all emotions that are associated with positive or negative excitement. Yet the things that we find pleasant or exciting vary to a significantly higher degree than do the things we perceive as approaching quickly or being far away. There are some stable triggers for certain emotions, like unexpected loud

noises and loss of balance in fear, but these are rare. Emotions are shaped by personal experiences and the cultural surroundings to a degree that is much higher than is the case for perception. Even those triggers that appear to be hardwired can lose their efficacy in certain contexts. Film directors certainly rely on bodily knowledge when creating situations that are supposed to elicit certain emotions; sudden loud noises and camera angles that elicit vertigo are among the most well-known tools to achieve fear responses in the audience. Yet if you watch Western movies all night long, they might start to bore you, and the bangs and booms might not have any emotional effect on you.

Therefore, bodily reactions involved in emotions are habit-like rather than law-like. Almost none of the connections between stimuli in the world and bodily reactions is hardwired. Yet emotional responses are shaped early in infancy in a way that is difficult to change later, and people seem to be disposed to acquire certain reactions more easily than other reactions. What follows from that is that emotion types cannot be individuated in the way that O'Regan and Noë suggest for individuating modalities. Sensorimotor habits appear very similar in many people, particularly if they grew up in similar social contexts. But emotions can be better individuated with regard to their function and the affordances they are set up to represent. The skillful knowledge that realizes the emotional reaction has to be sufficiently similar among the various instances of the emotion type, yet it can vary to some degree. As I argued in Chapter 2, the bodily profiles underlying various emotion types do not share any essence or necessary elements.

6. MEETING THE NORMATIVE CHALLENGE

Cognitivists argue that emotions are subject to semantic, rational, and social norms, and that they should be described as judgments. I claim that while the former is true the latter is wrong. Emotions do have a normative dimension, but this normative dimension can be understood in terms of a skillful organism interacting with a structured environment. I argue that emotions are embodied, action-oriented representations set up to represent core relational themes. The bodily reactions that realize emotions are evolutionarily established abilities that become skillful during a learning and interaction process. Emotions respond to relational properties that occur in the organism's environment and are of value for the organism. While an emotion's extension is defined by the relational property the emotion is set up to represent, an emotion's intension is defined by the skillful bodily reactions that constitute the action-oriented representation.

Such a theory can account for the emotions' being subject to semantic norms. It suggests that emotions can be appropriate or inappropriate because they are set up to respond to relational properties. If a relational property is really present, the emotion is adequate. If the relational property

is not present, the emotion is inadequate. It should furthermore be clear by now that emotions can be about things that are of value in a biological sense and about social rules and norms without thereby necessarily entailing conceptual content or explicit knowledge of what is represented. Rather, emotions represent core relational themes, such as a "danger-to-be-avoided" or a "rule-violation-to-make-amends-for." Because the adequate relational properties are given in the environment, they reliably strike the infant through locally recurring natural information, which reliably triggers certain bodily responses that realize an action-oriented representation. No cognitive evaluation, nothing judgment-like or containing conceptual knowledge, is needed to account for this. This also explains that while emotions respond to things of value, on a rational level, we often disagree with our pounding hearts and shaking limbs. That emotions are clumsy skills also means that it is always hard to know whether to go with the flow or suppress them.

Emotions are subject to rational norms insofar as their formal objects are logically and not causally restricted. I cannot envy myself, not because of causal laws but because it simply is incoherent to do so. Emotions appear to be connected in reasonable ways to each other. If I am afraid that I have lost my wallet, I should be relieved when finding it in my bag or angry and frustrated if I find out that it was stolen. Finally, emotions are connected to other mental states. A state of fear can vanish because I judge myself not to be in danger anymore.

The reason that the formal objects of emotions, that is, the affordances they are about, are logically and not causally restrained is that affordances are value-properties that motivate us to act in a certain way and not in another. Affective affordances constitute the structure of our environment, and motivate us to behave a certain way with regard to our needs and concerns. These emotional motivations can fit with our rational reasons for action or stand in conflict with them, because the norms that guide our reasoning can fit or conflict with the norms that are given to us through affective affordances. We are surrounded by things that are dangerous or restricting and by situations in which we could violate a norm. It does not make sense to envy oneself because the relational property that envy responds to is "another's good." It is not only a misrepresentation to envy oneself, it means representing oneself as having a property one cannot possibly have. This impossibility is of an ontological kind. It is in the first place a metaphysical not a psychological or conceptual law according to which "another's good" is a relational property that cannot be instantiated in me. It is the structure of the world in relation to our abilities and us that makes certain kinds of actions rational with regard to certain norms and others not. This allows us to say that envying oneself does not make sense, not because it violates our concept of what envy is but because it ascribes a property to an object, which cannot possibly have this kind of property. It is first of all a metaphysical and not a conceptual truth about envy that we cannot envy ourselves.

The same is true for the emotions standing in certain relations to each other. When I am afraid to lose something, I usually feel sad when I really do lose it. This is not primarily a conceptual law. Affordances stand in certain relations to us and to each other. Things that I value and do not want to lose can make me afraid if I am in danger of losing them and sad when I have lost them. This is, again, not because of conceptual laws or a certain kind of appraisal structure but because of relations between core relational themes in the world.

The ontological commitments needed for such realism were explained in the last chapter. We need to assume that things are of value in relation to the organism and that they thereby have the relational properties of being good or bad for the organism. What good or bad means can be further specified according to the organism's needs and the structure of the object. Some things are good because they are nutritious and others bad because they are poisonous. Core relational themes are value properties that are of central importance for the organism's well-being and therefore give a basic relevance structure to the environment of the organism.

This is why emotions can be subject to rational norms even though the organism having the emotion might be unable to understand the norms in question. Baboons might be ashamed of themselves because they violated rules of the rank hierarchy. It is a rational norm that one cannot be ashamed of just anybody but only of oneself and maybe of close relatives. Apes act in accordance with this norm not because they happen to understand it, but because shame is set up to represent rule violations committed by oneself not by others. If they observe rule violations of the same sort in others, they might react with anger, because the behavior of the others poses a restriction or offense in relation to oneself and not a rule violation committed by oneself.

7. CONCLUSION

Emotions are realized by sensorimotor skills in a way that does not fit into the sandwich model of the mind. To capture the motivating component and the skillful way in which emotions represent situations that matter for the organism, I therefore suggest explaining emotions in an action-oriented framework. The action-oriented view bridges the claim that emotions are embodied and the claim that emotions are embedded insofar as it explains that the bodily reactions involved in emotions are skillful ways to detect core relational themes that occur in the socially structured environment of the organism. The kind of content such embodied action-oriented representations carry can best be explained by taking emotions to be affordance representations. Emotions not only represent relational properties; they also directly motivate for action. The content of emotions is therefore descriptive and imperative at the same time. While the extension of an emotion is

determined by the relational properties it can come to represent, an emotional representation gets its intensional shape from the bodily abilities involved in it.

With this account in the background, I argue that by describing a skillful agent that interacts with a structured environment we can replace Lazarus's appraisal dimensions. The upshot is that all kinds of emotions can occur on a nonconceptual level. The way in which emotions represent their core relational themes is different from the way in which perception represents affordances insofar as emotional representations are shaped by the social context to a significantly higher degree. Emotions involve sensorimotor habits rather than sensorimotor contingencies.

Emotions' normative structure can be explained according to the normative structure that the environment has in relation to the organism. That core relational themes are properties in this environment accounts for the fact that emotions are subject to semantic norms. Emotions can be about values and norms because these properties are of biological value or instantiate a rule violation, regardless of whether the organism in question has the means to conceptually understand such rules and norms. The environment that contains these properties relevant for well-being is, finally, an environment that gives reason for action by highlighting situations of urgency with regard to biological and social norms. Emotions are subject to rational norms insofar as they can happen in accordance with these norms or violate them.

Given that the normative structure of emotions can fully be accounted for according to a story about skillful organisms interacting with a structured environment, we do not need overintellectualizing theories of emotions. We can account for emotions in a naturalist framework without inadequately reducing them to meaningless phenomena. It is an individual's environment that is meaningful, and it is the whole organism that is well equipped to respond to the affective affordances in its environment.

NOTES

1 Prinz criticizes Fodor's concept of modularity in later works, though, and also adjusts his view of emotions as being modular (2006b, 2006c).
2 "Egoreception accompanies exteroception, like the other side of a coin" (Gibson, 1986, 126).
3 For an account that develops the importance of learning time scales for enactivism, see Fingerhut (2011).
4 See Krueger (2014) and Colombetti and Krueger (forthcoming) for examples of entrainment or the establishment of embodied habits. See also Rietveld (2008) and Rietveld and Kieverstein (2014) for the claim that social affordances are entangled with social norms and forms of life.
5 I am grateful to Saray Ayala for having pointed out to me the relation of social affordances and power structures (see also Ayala, forthcoming).

References

Adelmann, P.; Zajonc, R. (1989): Facial Efference and the Experience of Emotion. *Annual Review of Psychology*, vol. 40. 249–280.

Ammaniti, M.; Gallese, V. (2014): *The Birth of Intersubjectivity. Psychodynamics, Neurobiology and the Self*. New York: W. W. Norton & Company.

Averill, J. (1997): The Emotions. An Integrative Approach. In R. Hogan, J. Johnson (eds.), *Handbook of Personality Psychology*. San Diego: Academic Press. 513–541.

Ayala, S. (forthcoming): A Structural Account to Speech Capacity. Affordances and Solicitations to Speak.

Barrett, L. F. et al. (2004): Interoceptive Sensitivity and Self-Reports of Emotional Experience. *Journal of Personality and Social Psychology*, vol. 87, no. 5. 684–697.

Beer, J. (2007): Neural Systems for Self-Conscious Emotions and Their Underlying Appraisals. In J. Tracy; R. Robbins; J. Tangney (eds.), *The Self-Conscious Emotions. Theory and Research*. New York: Guilford Press. 53–68.

Bermúdez, J. (1998): *The Paradox of Self-Consciousness*. Cambridge, MA: MIT Press.

Bernhardt, B.; Singer, T. (2012): The Neural Basis of Empathy. *Annual Review of Neuroscience*, vol. 5. 1–23.

Bischof-Kohler, D. (1991): The Development of Empathy in Infants. In M. Lamb; H. Keller (eds.), *Infant Development. Perspectives from German Speaking Countries*. Hillsdale: Erlbaum. 245–275.

Boyd, R. (1989): What Realism Implies and What It Does Not. *Dialectica*, vol. 43, no. 1. 5–29.

——— (1991): Realism, Anti-Foundationalism, and the Enthusiasm for Natural Kinds. *Philosophical Studies*, vol. 61, no. 1–2. 127–148.

Bradley, M. (2000/2007): Emotion and Motivation. In J. Cacioppo; G. Tassinary; G. Bernston (eds.), *Handbook of Psychophysiology*. Cambridge: Cambridge University Press. 602–642.

Brandon, R. (1997): Does Biology Have Laws? The Experimental Evidence. *Philosophy of Science*, vol. 64, suppl. 444–457.

Brooks, R. (1991): Intelligence Without Representation. *Artificial Intelligence*, vol. 47. 139–159.

Cacioppo, J.; Priester, J.; Bernston, G. (1993): Rudimentary Determinants of Attitudes. II: Arm Flexion and Extension Have Differential Effects on Attitudes. *Journal of Personality and Social Psychology*, vol. 65, no. 1. 5–17.

Cacioppo, J.; Tassinary, G.; Bernston, G. (2000): *Handbook of Psychophysiology*. Cambridge: Cambridge University Press.

Campos, J.; Stenberg, R. (1981): Perception, Appraisal, and Emotion: The Onset of Social Referencing. In M. Lamb; L. Sherrod (eds.), *Infant Social Cognition: Empirical and Theoretical Considerations*. Hillsdale: Erlbaum. 273–314.

Camras, L. et al. (1992): Japanese and American Infants' Responses to Arm-Restraint. *Developmental Psychology*, vol. 28, no. 4. 578–583.

Cannon, W. (1927): The James-Lange Theory of Emotion: A Critical Examination and an Alternative Theory. *American Journal of Philosophy*, vol. 39. 106–124.

——— (1936): *Bodily Changes in Pain, Hunger, Fear, and Rage*. New York: Appleton-Century.

Chemero, A. (2009): *Radical Embodied Cognitive Science*. Cambridge, MA: MIT Press.

Churchland, Patricia. (2011): *Braintrust. What Neuroscience Tells Us About Morality*. Princeton: Princeton University Press.

Chwalisz, K.; Diener, E.; Gallagher, D. (1988): Autonomic Arousal Feedback and Emotional Experience: Evidence from the Spinal Cord Injured. *Journal of Personality and Social Psychology*, vol. 54, no. 5. 820–828.

Ciaunica, A. (2014): Under Pressure. Processing Representational Decoupling in False-Belief-Tasks. *Review of Psychology and Philosophy*, vol. 5. 527–542.

Clark, A. (1997): *Being There. Putting Brain, Body and World Together Again*. Cambridge, MA: MIT Press.

——— (2008a): *Supersizing the Mind. Embodiment, Action, and Cognitive Extension*. Oxford: Oxford University Press.

——— (2008b): Pressing the Flesh: A Tension in the Study of the Embodied, Embedded Mind? *Philosophy and Phenomenological Research*, vol. 76, no. 1. 37–59.

——— (2014): Mindware. An Introduction to the Philosophy of Cognitive Science. Oxford: Oxford University Press.

Clark, J. (2010): Relations of Homology Between Higher Cognitive Emotions and Basic Emotions. *Biology and Philosophy*, vol. 25, no. 1. 75–94.

Colombetti, G. (2010): Enaction, Sense-Making and Emotion. In Steward, J.; Gapenne, O.; Di Paolo, E. (eds.), *Enaction: Toward a New Paradigm for Cognitive Science*, Cambridge MA, MIT Press. 145–164.

——— (2007): Enactive Appraisal. *Phenomenology and the Cognitive Sciences*, vol. 6. 527–546.

——— (2014): *The Feeling Body. Affective Science Meets the Enactive Mind*. Cambridge, MA: MIT Press.

Colombetti, G.; Krueger, J. (forthcoming): Scaffoldings of the Affective Mind. *Philosophical Psychology*.

Cosmides, L.; Tooby, J. (2010): The Evolutionary Psychology of the Emotions and Their Relationship to Internal Regulatory Variables. In M. Lewis; J. Haviland-Jones; L. Feldman-Barrett (eds.), *Handbook of Emotions*. New York: The Guilford Press. 114–138.

Crane, T. (1988): The Waterfall Illusion. *Analysis*, vol. 48, no. 3. 142–147.

Craver, C. (2007): Explaining the Brain. Mechanisms and The Mosaic Unity of Neuroscience. Oxford: Oxford University Press

Cussins, A. (2003): Content, Conceptual Content, and Nonconceptual Content. In Y. Gunther (ed.), *Essays on Nonconceptual Content*. Cambridge, MA: MIT Press. 133–163.

Damasio, A. (1994): *Descartes' Error. Emotion, Reason and the Human Brain*. London: Vintage Books.

——— (1999): *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York: The Ecco Press.

——— (2010): *Self Comes to Mind: Constructing the Conscious Brain*. New York: Pantheon Books.

Darwin, C. (1872/2009): *The Expression of the Emotions in Man and Animals*. London: Penguin Books.

Davidson, D. (1976): Hume's Cognitive Theory of Pride. *The Journal of Philosophy*, vol. 73, no. 19. 744–757.

Davis, J.; Senghas, A.; Brandt, F.; Ochsner, K. (2010): The Effects of BOTOX-Injections on Emotional Experience. *Emotion*, vol. 10, no. 3. 433–440.

De Bruin, L.; Kästner, L. (2012): Dynamic Embodied Cognition. *Phenomenology and the Cognitive Sciences*, vol. 11, no. 4. 541–563.

De Lancey, C. (2002): *Passionate Engines: What Emotions Reveal About Mind and Artificial Intelligence*. New York: Oxford University Press.

Deonna, J.; Teroni, F. (2012): The Emotions. A Philosophical Introduction. New York: Routledge.

Descartes, R. (1649/1988): The Passions of the Soul. J. Cottingham; R. Stoothoff; D. Murdoch (trans. and eds.), Cambridge: Cambridge University Press.

De Sousa, R. (1987): *The Rationality of Emotion*. Cambridge, MA: MIT Press.

De Vignemont F.; Singer, T. (2006): The Empathic Brain: How, When, and Why? *Trends in Cognitive Sciences*, vol. 10, no. 10. 435–441.

Dewey, J. (1896): The Reflex Arc Concept in Psychology. *The Psychological Review*, vol. 3 no. 4. 357–370.

Dickerson S.; Kemeny, M. (2004): Acute Stressors and Cortisol Responses: A Theoretical Integration and Synthesis of Laboratory Research. *Psychological Bulletin*, vol. 66, no. 1. 124–131.

Di Paolo, E. (2005): Autopoiesis, Adaptivity, Teleology, Agency. *Phenomenology and the Cognitive Sciences*, vol. 4. 97–125.

Di Paolo, E; Thompson, E. (2014): The Enactive Approach. In L. Shapiro (ed.), *Routledge Handbook of Embodied Cognition*. New York: Routledge. 68–78.

Döring, S. (2007): Seeing What to Do: Affective Perception and Rational Motivation. *Dialectica*, vol. 61, no. 3. 363–394.

Draghi-Lorenz, R. (2010): Parental Reports of Jealousy in Early Infancy: Growing Tensions Between Evidence and Theory. In S. Hart; M. Legerstee (eds.), *Handbook of Jealousy: Theory, Research and Multidisciplinary Approaches*. Oxford: Blackwell. 235–266.

Draghi-Lorenz, R.; Reddy, V.; Costall, A. (2001): Rethinking the Development of 'Nonbasic' Emotions: A Critical Review of Existing Theories. *Developmental Review*, vol. 21, no. 3. 263–304.

Draghi-Lorenz, R; Reddy, V.; Morris, P. (2005): Young Infants Can Be Perceived as Shy, Coy, Bashful, Embarrassed. *Infant and Child Development*, vol. 14, no. 1. 63–83.

Dretske, F. (1981): *Knowledge and the Flow of Information*. Cambridge, MA: MIT Press.

——— (1986): Misrepresentation. In R. Bogdan (ed.), *Belief: Form, Content and Function*. Oxford: Oxford University Press. 17–36.

——— (1995): *Naturalizing the Mind*. Cambridge, MA: MIT Press.

Dreyfus, H. (1991): *Being-in-the-World: A Commentary on Heidegger's Being and Time*. Cambridge, MA: MIT Press.

Duclos, S. et al. (1989): Emotion-Specific Effects of Facial Expressions and Postures on Emotional Experience. *Journal of Personality and Social Psychology*, vol. 57, no. 1. 100–108.

Ekman, P. (1971): Universals and Cultural Differences in Facial Expressions of Emotion. In J. Cole (ed.), *Nebraska Symposium on Motivation 4*. Lincoln: University Press of Nebraska. 207–282.

——— (2003): *Emotions Revealed. Understanding Faces and Feelings*. London: Orion Books.

Ekman, P.; Rosenberg, E. (1997): *What the Face Reveals. Basic and Applied Studies of Spontaneous Expressions Using the Facial Action Coding System*. New York: Oxford University Press.

Ekman, P.; Levenson, W.; Friesen, W. (1983): Autonomic Nervous System Activity Distinguishes among Emotions. *Science*, vol. 221, no. 4616. 1208–1210.

180   *References*

Epstein, B. (2015): *The Ant Trap, Rebuilding the Foundations of the Social Sciences.* Oxford: Oxford University Press.

Evans, G. (1982): The Varieties of Reference. McDowell, J. (ed.), Oxford: Oxford University Press.

Feldman-Barret, L. (2006): Are Emotions Natural Kinds? *Perspectives on Psychological Science*, vol. 1. 28–58.

Fessler, D. (2004): Shame in Two Cultures: Implications for Evolutionary Approaches. *Journal of Cognition and Culture*, vol. 4, no. 2. 207–262.

Fingerhut, J. (2011): Sensorimotor Signature, Skill, and Synaesthesia. Two Challenges for an Enactive View of Perception. In J. Fingerhut; S. Flach; J. Söffner (eds.), *Habitus in Habitat III, Synaesthesia and Kinaesthetics*. Berlin: Peter Lang Verlag. 101–120.

Fodor, J. (1983): *The Modularity of Mind. An Essay on Faculty Psychology.* Cambridge, MA: MIT Press.

——— (1990): *A Theory of Content and Other Essays*. Cambridge, MA: MIT Press.

Foot, P. (2001): *Natural Goodness*. Oxford: Clarendon Press.

Frijda, N. (1986): *The Emotions*. Cambridge: Cambridge University Press.

Frijda, N. (1993): The Place of Appraisal in Emotion. *Cognition and Emotion*, vol. 7, no. 3–4. 357–387.

Gallagher, S. (2005): *How the Body Shapes the Mind*. Oxford: Clarendon Press.

——— (2008a): Direct Perception in the Intersubjective Context. *Consciousness and Cognition*, vol. 17, no. 2. 535–543.

——— (2008b): Are Minimal Representations Still Representations? *International Journal of Philosophical Studies*, vol. 16, no. 3. 351–369.

Gallagher, S.; Hutto, D. (2007): Understanding Others Through Primary Interaction and Narrative Practice. In J. Zlatev et al. (eds.), *The Shared Mind. Perspectives on Intersubjectivity*. Amsterdam: John Benjamins. 17–38.

Gallese, V. (2009): Mirror Neurons, Embodied Simulation, and the Neural Basis of Social Identification. *Psychoanalytic Dialogues,* vol. 19. 519–536.

Geach, P. (1956): Good and Evil. *Analysis*, vol. 17, no. 2. 32–42.

Gibson, J. (1986): *The Ecological Approach to Visual Perception*. New York: Psychology Press.

Gibson, W. et al. (2015): Behavioral Responses to a Repetitive Visual Threat Stimulus Express a Persistent State of Defensive Arousal in Drosophila. *Current Biology*, vol. 25, no. 11. 1401–1415.

Godfrey-Smith, P. (1989): Misinformation. *Canadian Journal of Philosophy*, vol. 19. 533–550.

Goldie, P. (2000): *The Emotions: A Philosophical Exploration*. Oxford: Oxford University Press.

Gopnik, A.; Meltzoff, A. (1997): *Words, Thoughts, and Theories*. Cambridge: MIT Press.

Greenspan, R. (2001): The Flexible Genome. *Nature Reviews Genetics*, vol. 2, no. 5. 383–387.

Griffiths, P. (1997): *What Emotions Really Are. The Problem of Psychological Categories.* Chicago: University of Chicago Press.

Griffiths, P.; Scarantino, A. (2009): Emotions in the Wild, the Situated Perspective on Emotion. In P. Robbins; M. Ayede (eds.), *Handbook of Situated Cognition*. Cambridge: Cambridge University Press. 437–454.

Harrison, N.; Kreibig, S.; Critchley, H. (2013): A Two-Way Road: Efferent and Afferent Pathways of Autonomic Activity in Emotion. In J. Armony; P. Vuilleumier (eds.), *The Cambridge Handbook of Human Affective Neuroscience*. Cambridge: Cambridge University Press. 82–106.

Hart, S. (2015): *Jealousy in Infants. Laboratory Research on Differential Treatment.* New York: Springer.

Hart, S.; Carrington, H. (2002): Jealousy in 6-Month-Old Infants. *Infancy,* vol. 3, no. 3. 395–402.

Hart, S.; Carrington, H.; Tronick, E.; Caroll, S. (2004): When Infants Lose Exclusive Maternal Attention: Is It Jealousy? *Infancy,* vol. 6, no. 1. 57–78.

Haslanger, S. (1995): Ontology and Social Construction. *Philosophical Topics*, vol. 23, no. 2. 95–125.

——— (2015): What Is a (Social) Structural Explanation? *Philosophical Studies* (online first).

Hatfield, E. et al. (1995): The Impact of Vocal Feedback on Emotional Experience and Expression. *Journal of Social Behavior and Personality*, vol. 10, no. 2. 293–312.

Helm, B. (2001): Emotional Reason. Deliberation, Motivation, And the Nature of Value. Cambridge: Cambridge University Press.

Henley, N. (1995): Body Politics Revisited. What Do We Know Today? In P. Kalbfleisch; C. Cody (eds.), *Gender, Power and Communication in Human Relationships*. Hillsdale: Erlbaum. 27–61.

Hohmann, G. (1966): Some Effects of Spinal Cord Lesions on Experienced Emotional Feelings. *Psychophysiology,* vol. 3, no. 2. 143–156.

Hufendiek, R. (2012): Das Muttertier am Ursprung der Moral. Neue Naturalistische Ansätze in der Metaethik. *Zeitschrift für philosophische Forschung*, vol. 66, no. 2. 270–291.

Hume, D. (1739/2007): *A Treatise of Human Nature*. D. Norton; M. Norton (eds.), Oxford: Oxford University Press.

Hurley, S. (1997): Nonconceptual Self-Consciousness and Agency: Perspective and Access. *Communication and Cognition*, vol. 30, no. 3/4. 207–248.

——— (1998): *Consciousness in Action*. Cambridge, MA: Harvard University Press.

——— (2001): Perception and Action. Alternative Views. *Synthese*, vol. 129, no. 1. 3–40.

Hutto, D. (2006): Impossible Problems and Careful Expositions. In R. Menary (ed.), *Radical Enactivism. Intentionality, Phenomenology and Narrative. Focus on the Philosophy of Dan Hutto*. Amsterdam: John Benjamins Publishing Company. 45–65.

——— (2005): Knowing What? Radical versus Conservative Enactivism. *Phenomenology and the Cognitive Sciences*, vol 4, no. 4. 389–405.

——— (2012): Truly Enactive Emotion. Emotion Review, vol. 4, no. 2. 176–181.

Hutto, D.; Myin, E. (2013) Radicalizing Enactivism. Basic Minds Without Contents. Cambridge, MA: MIT Press.

James, W. (1890/1980): *The Principles of Psychology*. New York: Dover Publications.

——— (1884): What Is an Emotion? *Mind*, vol. 9. 188–205.

Johnson, S. (2000): The Recognition of Mentalistic Agents in Infancy. *Trends in Cognitive Science*, vol. 4. 22–28.

Kalueff, A. et al. (2012): Time to Recognize Zebrafish Affective Behavior. *Behaviour*, vol. 149, no. 10–12. 1019–1036.

Kelly, D. (2011): *Yuck! The Nature and Moral Significance of Disgust*. Cambridge, MA: MIT Press.

Kenny, A. (1963): *Action, Emotion, and Will*. London: Routledge.

Kreibig, S. (2010): Autonomic Nervous System Activity in Emotion: A Review. Biological *Psychology*, vol. 84, no. 3. 394–421.

Krueger, J. (2014): Emotions and the Social Niche. In M. Salmela; C. Scheeve (eds.), *Collective Emotions*. Oxford: Oxford University Press. 156–171.

——— (forthcoming): Varieties of Extended Emotions. *Phenomenology and Cognitive Sciences*.

Labouvie-Vief, G. et al. (2003): Age and Gender Differences in Cardiac Reactivity and Subjective Emotion Response to Emotional Autobiographical Memories. *Emotion*, vol. 3, no. 1. 115–126.

Ladyman, J.; Ross, D. (2009): *Everything Must Go. Metaphysics Naturalized.* Oxford: Oxford University Press.

Laird, J. (2007): *Feelings. The Perception of Self.* Oxford: Oxford University Press.

——— (1974): Self-Attribution of Emotion: The Effects of Expressive Behavior on the Quality of Emotional Experience. *Journal of Personality and Social Psychology*, vol. 29, no. 4. 475–486.

Laird, J.; Lacasse, K. (2014): Bodily Influences on Emotional Feelings: Accumulating Evidence and Extensions of William James's Theory of Emotions. *Emotional Review*, vol. 6. 24–37.

Lang, P.; Bradley, M.; Cuthbert, B. (1990): Emotion, Attention and the Startle Reflex. *Psychological Review*, vol. 97, no. 3. 377–395.

Latour, B. (2005): *Reassembling the Social. An Introduction to Actor-Network-Theory.* Oxford: Oxford University Press.

Lazarus, R. (1991): *Emotion and Adaptation.* New York: Oxford University Press.

Legrand, D. (2007): Pre-Reflective Self-as-Subject from Experiential and Empirical Perspectives. *Consciousness and Cognition,* vol. 16, no. 3. 583–599.

Lerner, J.S. et al. (2005): Facial Expressions of Emotions Reveal Neuroendocrine and Cardiovascular Stress Responses. *Biological Psychiatry*, vol. 58. 743–750.

Lewis, M. (1992): The Self in Self-Conscious Emotions. *Research in Child Development*, vol. 57, no. 1. 85–94.

——— (2006): Early Emotional Development. In A. Slater; M. Lewis (eds.), *Introduction to Infant Development.* Oxford: Oxford University Press. 192–209.

——— (2014): *The Rise of Consciousness and the Development of Emotional Life.* New York: Guilford Press.

Lewis, M.; Brooks-Gunn, J. (1979): *Social Cognition and the Acquisition of Self.* New York: Plenum.

Lewis, M.; Allessandri, S.; Sullivan, M. (1990): Violation of Expectancy, Loss of Control, and Anger in Young Infants. *Developmental Psychology*, vol. 26. 745–751.

Lewis, M.; Sullivan, M.; Stanger, C.; Weiss, M. (1989): Self Development and Self-Conscious Emotions. *Child Development*, vol. 60, no. 1. 146–156.

Lyon, M. (1999): Emotion and Embodiment: The Respiratory Mediation of Somatic and Social Processes. In A. Hinton (ed.), *Biocultural Approaches to the Emotions*. Cambridge: Cambridge University Press. 182–213.

Maiese, M. (2011): *Embodiment, Emotion and Cognition.* London: Palgrave Macmillan.

Manstead, A. (2005): The Social Dimension of Emotion. *The Psychologist,* vol. 18, no. 8. 484–487.

Marr, D. (1982): *Vision. A Computational Investigation into Human Representation and Processing of Visual Information.* New York: Freeman & Co.

Matsumoto, D.; Keltner, D.; Shiota, M.; O'Sullivan, M.; Frank, M. (2010): Facial Expressions of Emotion. In M. Lewis; J. Haviland-Jones; L. Feldman Barrett (eds.), *Handbook of Emotions*. New York: The Guilford Press. 211–235.

Mauss, B.; Robinson, M. (2009): Measures of Emotion: A Review. *Cognition and Emotion*, vol. 23, no. 2. 209–237.

Mauss, B.; Levenson, R.; McCarter, L.; Wilhelm, F., Gross, J. (2005): The Tie That Binds? Coherence Among Emotion Experience, Behavior and Physiology. *Emotion*, vol. 5, no. 2. 175–190.

McNaughton, N. (1989): *Biology and Emotion.* Cambridge: Cambridge University Press.

Meltzoff, A.; Brooks R. (2001) 'Like Me' as a Building Block for Understanding Other Minds: Bodily Acts, Attention, and Intention. In B. Malle; L. Moses; D. Baldwin (eds.), *Intentions and Intentionality: Foundations of Social Cognition.* Cambridge, MA: MIT Press. 171–191.

Meltzoff, A; Moore, M. (1977): Imitation of Facial and Manual Gestures by Human Neonates. *Science*, vol. 198, no. 4312. 75–78.

——— (1983): Newborn Infants Imitate Adult Facial Gestures. *Child Development*, vol. 54, no. 3. 702–709.

——— (1994): Imitation, Memory and the Representation of Persons. *Infant Behavior and Development*, vol. 17, no. 1. 83–99.

Millikan, R. (1984): *Language, Thought, and Other Biological Categories. New Foundations for Realism.* Cambridge, MA: MIT Press.

——— (1993): *White Queen Psychology and Other Essays for Alice.* Cambridge, MA: MIT Press.

——— (1995): Pushmi-Pullyu Representations. *Philosophical Perspectives*, vol. 9. 185–200.

——— (2004): *Varieties of Meaning.* Cambridge, MA: MIT press.

Moore, G.E. (1903): *Principia Ethica.* Cambridge: Cambridge University Press.

Musholt, K. (2013): Self-Consciousness and Nonconceptual Content. *Philosophical Studies*, vol. 163, no. 3. 649–672.

——— (2015): Thinking About Oneself: From Nonconceptual Content to the Concept of a Self. Cambridge, MA: MIT Press.

Nakayama, K.; Goto, S.; Kuraoka, K.; Nakamura, K. (2005): Decrease in Nasal Temperature of Rhesus Monkeys (Macaca mulatta) in Negative Emotional State. J. Physiological Behavior vol. 84, 783–790

Neander, K. (1996): Dretske's Innate Modesty. *Australian Journal of Philosophy,* vol. 74. 258–274.

——— (2002): Types of Traits. The Importance of Functional Homologues. In A. Ariew; R. Cummins; M. Perlman (eds.), *Functions: New Readings in the Philosophy of Psychology and Biology.* Oxford: Oxford University Press. 390–415.

Neisser, U. (1988): Five Kinds of Self-Knowledge. *Philosophical Psychology,* vo1.1, no. 1. 35–59.

Newmark, C. (2008): *Passion—Affekt—Gefühl. Philosophische Theorien der Emotionen zwischen Aristoteles und Kant.* Hamburg: Meiner Verlag.

Niedenthal, P.; Barsalou, L.; Ric, F.; Krauth-Gruber, S. (2005): Embodiment in the Acquisition and Use of Emotion Knowledge. In L. Feldman Barrett; P. Niedenthal; P. Winkielman (eds.), *Emotion and Consciousness.* New York: Guilford Press. 21–51.

Noë, A. (2004): *Action in Perception.* Cambridge, MA: MIT Press.

Nussbaum, M. (2001): *Upheavals of Thought. The Intelligence of Emotions.* Cambridge: Cambridge University Press.

O'Regan, K.; Noë, A. (2001): A Sensorimotor Account of Vision and Visual Consciousness. *Behavioral and Brain Sciences*, vol. 24, no. 5. 939–1031.

Panksepp, J. (1998): *Affective Neuroscience. The Foundations of Human and Animal Emotions.* Oxford: Oxford University Press.

Papineau, D. (1993): *Philosophical Naturalism.* Oxford: Blackwell.

Parkinson, B. (1996): Emotions Are Social. *British Journal of Psychology*, vol. 87, no. 4. 663–683.

——— (2001): Anger On and Off the Road. *British Journal of Psychology*, vol. 92. 507–526.

Parkinson, B.; Fischer, A.; Manstead, A. (2005): *Emotions in Social Relations, Cultural, Group, and Interpersonal Processes.* New York: Psychology Press.

Peacocke, C. (2001): Does Perception Have a Nonconceptual Content? *Journal of Philosophy*, vol. 98, no. 5. 239–264.

Phelps, E. (2004): The Human Amygdala and Awareness: Interactions Between Emotion and Cognition. In M. Gazzaniga (ed.), *The Cognitive Neurosciences III.* Cambridge, MA: MIT Press. 1005–1015.

Place, U. (1956): Is Consciousness a Brain Process? *British Journal of Psychology*, vol. 47. 44–50.

Porges, S.; Carter, S. (2011): Neurobiology and Evolution: Mechanisms, Mediators, and Adaptive Consequences of Caregiving. In: S. Brown; R. Brown; A. Penner (eds.),

Toward a New Understanding of Human Caregiving. Oxford: Oxford University Press. 53–71.

Prinz, J. (2004): *Gut Reactions. A Perceptual Theory of Emotions*. Oxford: Oxford University Press.

——— (2006b): Is Emotion a Form of Perception? In L. Faucher; C. Tappolet (eds.), *The Modularity of Emotions. Canadian Journal of Philosophy*. Calgary: University of Calgary Press. 137–161.

——— (2006c): Is the Mind Really Modular? In R. Stainton (ed.), *Contemporary Debates in Cognitive Sciences*. Oxford: Blackwell. 22–36.

——— (2007): *The Emotional Construction of Morals*. Oxford: Oxford University Press.

——— (2010): For Valence. *Emotion Review*, vol. 2, no. 1. 5–13.

Putnam, H. (1960): Minds and Machines. In S. Hook (ed.), *Dimensions of Mind*. New York: New York University Press. 148–179.

Ratcliffe, M. (2005): William James on Emotion and Intentionality. *International Journal of Philosophical Studies*, vol. 13, no. 2. 179–202.

Reddy, V. (2000): Coyness in Early Infancy. *Developmental Science*, vol. 3, no. 2. 186–192.

——— (2008): *How Infants Know Minds*. Cambridge, MA: Harvard University Press.

Reed, E. (1996): *Encountering the World. Toward an Ecological Psychology*. New York: Oxford University Press.

Reisenzein, R.; Stephan, A. (2014): More on James and the Physical Basis of Emotions. *Emotion Review*, no. 6. 35–46.

Reissland, N. (1990): Parental Frameworks of Pleasure and Pride. *Infant Behavior and Development*, vol. 13, no. 2. 249–256.

Rietveld, E. (2008): Situated Normativity: The Normative Aspect of Embodied Cognition in Unreflected Action. *Mind*, vol. 468, no. 117. 973–1001.

Rietveld, E.; Kieverstein, J. (2014): A Rich Landscape of Affordances. *Ecological Psychology*, vol. 26, no. 4. 325–352.

Robbins, P.; Ayede, M. (2009): A Short Primer on Situated Cognition. In P. Robbins; M. Ayede (eds.), *Handbook of Situated Cognition*. Cambridge: Cambridge University Press. 3–10.

Roberts, R. (2003): *Emotions: An Essay in Aid of Moral Psychology*. Cambridge: Cambridge University Press.

Rochat, P. (2003): Five Levels of Self-Awareness as They Unfold Early in Life. *Consciousness and Cognition*, vol. 12, no. 4. 717–731.

——— (2009): *Others in Mind. Social Origins of Self-Consciousness*. Cambridge: Cambridge University Press.

Roeser, S. (2010): Moral Emotions and Intuitions. London: Palgrave Macmillan.

Von Rohr, R.; Burkart, J.; Van Schaik, C. (2011): Evolutionary Precursors of Social Norms in Chimpanzees. A New Approach. *Philosophy and Biology*, vol. 26. 1–30.

Rowlands, M. (2012): *Can Animals Be Moral?* Oxford: Oxford University Press.

Salmela, M. (2011). Can Emotions Be Modeled on Perception? *Dialectica* vol. 65, no. 1. 1–29.

Scarantino, A. (2003): Affordances Explained. *Philosophy of Science*, vol. 70, no. 5. 949–962.

——— (2010): Insights and Blindspots of the Cognitivist Theory of Emotions. *The British Journal for the Philosophy of Science*, vol. 64, no. 4. 729–768.

——— (2014): The Motivational Theory of Emotions. In D. Jacobson; J. D'Arms (eds.), *Moral Psychology and Human Agency*. Oxford: Oxford University Press. 156–185.

Schachter, S.; Singer, J. (1962): Cognitive, Social, and Physiological Determinants of Emotional State. *Psychological Review*, vol. 69, no. 5. 379–399.

Schwartz, G.; Ahern, G.; Brown, S. (1979): Lateralized Facial Muscle Response to Positive and Negative Emotional Stimuli. *Psychophysiology*, vol. 16. 561–571.

Searle, J. (1995): *The Construction of Social Reality*. New York: Free Press.

——— (2010): *Making the Social World. The Structure of Civilization*. Oxford: Oxford University Press.

Shapiro, Lawrence (2004): *The Mind Incarnate*. Cambridge, MA: MIT Press.

——— (2010): *Embodied Cognition*. New York: Routledge.

——— (2014): *The Routledge Handbook of Embodied Cognition*. New York: Routledge.

Slaby, J. (2014): Emotions and the Extended Mind. In M. Salmela; C. Scheeve (eds.), *Collective Emotions*. Oxford: Oxford University Press. 32–46.

Smart, J. (1959): Sensations and Brain Processes. *Philosophical Review*, vol. 68. 141–156.

Sneddon, A. (2006): Two Views of Emotional Perception: Some Empirical Suggestions. *Canadian Journal of Philosophy*, vol. 32, suppl. 36. 161–183.

Solomon, R. (2003): Not Passions Slave. Emotions and Choice. Oxford: Oxford University Press

Spaulding, S. (2014): Embodied Cognition and Theory of Mind. In L. Shapiro (ed.), *The Routledge Handbook of Embodied Cognition*. New York: Routledge. 197–206.

Stemmler, G; Heldmann, M.; Pauls, C.; Scherer, T. (2001): Constraints for Emotions Specificity in Fear and Anger: The Context Counts. *Psychophysiology* vol. 38, 275-291.

Stephan, A.; Walter, S.; Wilutzky, W. (2013): Emotions Beyond Brain and Body. *Philosophical Psychology*, vol. 27, no. 1. 1–17.

Stepper, S.; Strack, F. (1993): Proprioceptive Determinants of Emotional and Nonemotional Feelings. *Journal of Personality and Social Psychology*, vol. 64, no. 2. 211–220.

Sterelny, K. (2010): Minds Extended or Scaffolded? *Phenomenology and the Cognitive Sciences*, vol. 9, no. 4. 465–481.

Strack, F.; Martin, L.; Stepper, S. (1988): Inhibiting and Facilitating Conditions of the Human Smile: A Nonobtrusive Test of the Facial Feedback Hypothesis. *Journal of Personality and Social Psychology*, vol. 54, no. 5. 768–777.

Tappolet, C. (2000): *Émotions et valeurs*. Paris: Presses universitaires de France.

Taylor, S. (1991): Asymmetrical Effects of Positive and Negative Events : The Mobilization-Minimization Hypothesis. *Psychological Bulletin*, vol. 110. 67–85.

Thelen, E.; Smith, L. (1994): *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge, MA: MIT Press.

Thompson, E. (2007): *Mind in Life. Biology, Phenomenology and the Sciences of Life*. Cambridge, MA: Harvard University Press.

Tracy, J.; Robbins R. (2004): Show Your Pride. Evidence for a Discrete Emotion Expression. *American Psychological Society*, vol. 15, no. 3. 194–196.

——— (2007a): The Prototypical Pride-Expression: Development of a Nonverbal Behavior Coding Scheme. *Emotion*, vol. 7, no. 4. 789–801.

——— (2007b): The Nature of Pride. In J. Tracy; R. Robbins; J. Tangney (eds.), *The Self-Conscious Emotions. Theory and Research*. New York: The Guilford Press. 263–282.

Tracy, J.; Shariff, A.O., Zhao, W.; Henrich, J. (2013): Cross-cultural Evidence That the Pride-expression Is a Universal Automatic Status-sign. *Journal of Experimental Psychology,* vol. 142. 163–180.

Trevarthen, C.B. (1979): Communication and Cooperation in Early Infancy: A Description of Primary Intersubjectivity. In M. Bullowa (ed.), *Before Speech*. Cambridge: Cambridge University Press. 321–348.

Turvey, M. (1992): Affordances and Prospective Control: An Outline of the Ontology. *Ecological Psychology*, vol. 4. 173–187.

Varela, F.; Thompson, E.; Rosch, E. (1991): *The Embodied Mind. Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.

Watson, J. (1930/1970): *Behaviorism*. New York: Norton.

Welpinghus, A. (2015): Emotions as Natural and Social Kinds. An Integrative Approach to Evolutionary and Social Perspectives on Emotions. Paderborn: Mentis.

Wheeler, M. (2005): *Reconstructing the Cognitive World: The Next Step*. Cambridge, MA: MIT Press.

Wild, M. (2010): *Biosemantik. Repräsentation, Intentionalität, Norm*. Unpublished Habilitation.

Wilson, M. (2002): Six Views of Embodied Cognition. *Psychonomic Bulletin & Review*, vol. 9, no. 4. 625–636.

Zahn-Waxler, C.; Radtke-Yarrow, M.; King, R. (1992): Development of Concern for Others. *Developmental Psychology*, vol. 28, no. 1. 126–136.

Zajonc, R. (1980): Feeling Before Thinking, Preferences Need No Inferences. *American Psychologist*, no. 35. 151–175.

Zinck, A.; Newen, A. (2008): Classifying Emotion. A Developmental Account. *Synthese*, vol. 161, no. 1. 1–25.

Index