# A SOMATIC-PERCEPTUAL THEORY OF THE EMOTIONS

# A SOMATIC-PERCEPTUAL THEORY OF THE EMOTIONS

By JENNIFER-WRAE PRIMMER, B.A. [H], M.A.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the Requirements for

The Degree Doctorate of Philosophy

McMaster University DOCTORATE OF PHILOSOPHY (2015) Hamilton, Ontario (Philosophy)

TITLE: A Somatic-Perceptual Theory of the Emotions

AUTHOR: Jennifer-Wrae Primmer, B.A. [H] (Dalhousie University), M.A. (University of Saskatchewan)

SUPERVISOR: Professor Mark Johnstone

NUMBER OF PAGES: v, 162

# ABSTRACT

In this dissertation, I develop and defend a kind of somatic theory of the emotions; namely, a somatic-perceptual theory of the emotions. On this account, emotions are perceptions of physiological changes. The majority of emotion theorists, however, hold some kind of a cognitive theory of the emotions. I argue, in opposition to these theories, that cognition is never necessary for emotion. Somatic theories of the emotions have never been well-received in philosophy and psychology. This is mainly because they are often perceived as being ill-equipped to explain many of the things that a theory of the emotions ought to account for. In particular, it is argued that somatic theories of the emotions fail to take into account the fact that emotions are typically directed toward an intentional object. Somatic theories, it is argued, are also unable to explain how to distinguish between different emotions associated with identical physiological responses. Moreover, since on my view emotions are a form of perception, my view would seem to allow for the bodily perceptions constituting emotions to occur unconsciously. However, in philosophy, the notion of unconscious emotions is problematic, because in ordinary language, emotions just are feelings – and feelings *are*, by definition, conscious. Using philosophical arguments and empirical evidence from neuroscience and psychology, I argue that my somatic-perceptual theory of the emotions is able to account both for the intentional nature of the emotions and the distinctiveness of different emotions just as well as leading cognitive theories of the emotions. This is significant because these objections have not yet been adequately met by other somatic theories of the emotions. I also embrace the implication that on my view, emotions *can* be unconscious, and show that my somatic-perceptual theory provides a framework for thinking about poorly understood psychological conditions, such as alexithymia.

# ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to those who have supported me in one way or the other throughout the process of completing this dissertation. Much has changed in my life since I started this project five years ago. It has not always been an easy journey, but it certainly has been a remarkable one.

First, I am extremely grateful to my supervisor, Mark Johnstone, for his valuable guidance, insightful comments, and consistent encouragement throughout the writing process. Mark's incredible patience, his enthusiasm for philosophical discourse, and his amicable and positive disposition will always be a source of inspiration to me. Mark always cleared time from his busy schedule to meet with me, was always willing to clarify and help to resolve any issues that I may have had, and provided comments on my work faster than any supervisor I have known! I am truly indebted to him for making my research a priority, and for his unwavering support.

I am also sincerely grateful to my second reader, John Connolly, for his valuable insights, his enthusiasm for my project, and his unconditional support. I am incredibly thankful for John inviting me to participate in his lab, and for the opportunity not only to teach other lab members about the importance of philosophy, but to also learn about emotions and consciousness from a different perspective. I would also like to thank Bruce Milliken for his helpful comments, his willingness to discuss my work, and for his continued support. I am also grateful to Nancy Doubleday for her encouragement and support throughout the research and writing process.

I am deeply indebted to my mother, Judy. Not only did she give me the courage to reach for my dreams, but she also guided me and held my hand when I lost faith in myself. Her love and encouragement throughout the years has been unwavering, and has served as a constant reminder that I can achieve whatever I put my mind to. Without her, I would be lost. It is to her that I dedicate this work.

I would also like to express my sincere gratitude to my father, Robert, for his love and support – and for always reminding me that things are never as bad as I first think they are! A heart-felt thanks to my brother, Brett – his love and friendship throughout the years remains unparalleled. He will always be an inspiration and a great source of joy in my life. My deepest and most sincere thanks are reserved for my grandparents, Herman and Jean. Words cannot adequately express how much their love and support has meant to me. I would not be where I am today without their guidance and constant encouragement. I will forever be indebted to them.

My sincerest thanks go to Joanna Zaslow and Kait Pinder for their kindness and friendship during this incredible journey. Finally, I am grateful to Jarrod Hanson, who has been a great source of strength and inspiration to me. His love and encouragement has helped to make this process a much brighter and happier one.

TABLE OF CONTENTS

# LISTS OF FIGURES AND TABLES

# INTRODUCTION

Over one hundred and thirty years ago, American psychologist William James (1884) asked: "what is an emotion?" Although numerous replies have been proposed, the answer nevertheless remains unclear. Among contemporary emotion theorists there is little consensus on what emotions are. Typical emotion episodes contain a number of components – e.g., cognitions, bodily changes, action tendencies, alteration in attention, and conscious feelings – but, which component *is* the emotion? Does one single component stand out as 'the emotion'? Might 'emotion' refer to more than one component? Perhaps it refers to the emotion episode as a whole?

Folk psychology suggests that emotions are mere feelings (i.e., phenomenally conscious mental states).[1] On this view, being angry, for example, just is feeling a certain way. This theory, referred to as the *feeling theory of the emotions*, has never been popular among emotion theorists.[2] Instead, the majority of emotion theorists defend a *cognitive*

[1] Folk psychology (also referred to as common sense psychology), is a theory that psychologists developed to capture how ordinary people (i.e., individuals who lack formal training in various academic fields) attribute or predict the mental states of others.

[2] According to the feeling theory of emotions, every emotion has a distinct phenomenal character (i.e., every emotion has a distinct experiential feeling associated with it). It is these distinct experiential feelings that enable one to determine that one is experiencing joy, and not, for example, anger. Moreover, these experiential feelings can (though they need not) find expression in a person's outward behavior. For example, anger is a specific feeling which can (though it need not) lead a person to strike or yell (Bedford, 1957). If one sees another person strike or yell, one can infer, through that person's behavior, that he/she is experiencing anger. However, attributing emotions to another person based solely on his/her outward

*theory of the emotions* (e.g., Nussbaum, 2001; Ben Ze'ev, 2000; Lazarus, 1991). On this view, emotions *necessarily* contain a cognitive component (i.e., a propositional attitude, such as a belief or a judgment).

Against this majority view, I argue that cognition is not essential to emotion: while emotions *can* contain a cognitive component, they need not do so. In contrast to cognitive theories of the emotions, I defend a kind of *somatic* theory of the emotions; namely, a *somatic-perceptual* theory of the emotions. On this view, emotions are perceptions of changes in one's body. More specifically, emotions are states in one's somatosensory cortex that register certain patterns of physiological changes.

On the view I defend, one's body undergoes various perturbations (e.g., alterations in one's circulatory, respiratory, and musculoskeletal systems), and an emotion just is one's perception of that change. Physiological and neurological evidence supports this: individuals who suffer from spinal injuries (thus reducing feedback from the body) often report diminished emotional response (Bechara, 2004), damage to the cingulate cortex (i.e., an area in the cortex involved with interoception) can lead to akinetic mutism (i.e., a profound deficit in emotional response) (Damasio and Van Hoesen, 1983), and functional neuroimaging studies show activation in somatic brain centers during emotion induction (Damasio, et al., 2000).

However, somatic theories of the emotions have never been well-received in philosophy and psychology. This is mainly because they are often perceived as being ill-equipped to explain many of the things that a theory of the emotions ought to account for.

---

behavior is problematic, since it is possible for an individual to strike or yell when they are not, in fact, angry (e.g., an actor can appear to be angry without actually experiencing the feelings indicative of anger).

In particular, critics claim that existing somatic theories fail to account for the *intentional* nature of emotions; that is, they fail to take into account the fact that emotions are typically directed toward an intentional object (i.e., a particular object or event). Somatic theories, they argue, are also unable to explain how we are to distinguish between different emotions associated with identical physiological responses (e.g., sadness and guilt). Moreover, since on my view emotions are a form of perception, my view would seem to allow for the bodily perceptions constituting emotions to occur unconsciously (i.e., below the subjective threshold of awareness). This is because not all perceptions are felt (Winkielman, et al., 2005; Cheesman and Merikle, 1986). The notion of unconscious emotions, however, is heavily debated in the philosophical literature (e.g., Clore, 1994; Rosenthal, 1991; Prinz, 2004). According to critics, the notion of unconscious emotions is problematic because in ordinary language, emotions just are feelings – and feelings *are*, by definition, conscious. In order to successfully defend a somatic theory of the emotions, it is necessary that I overcome these objections.

Prominent emotion theorist, Jesse Prinz (2004), has developed the most comprehensive somatic theory currently available in the philosophical and psychological literature. However, there are problems with Prinz' theory. For example, Prinz makes inconsistent claims about the basic structure of the emotions, and fails to adequately distinguish emotions from all non-emotional interoceptive states. One of my goals in this dissertation is to draw attention to the inconsistencies and gaps in Prinz' account, and to offer viable solutions to them. Although the theory that I defend is of the same general kind as Prinz' theory (e.g., both are somatic theories which claim that emotions are

perceptions of changes in the body), it nevertheless avoids many of the problems that his view faces. By achieving clarity on points about which Prinz is unclear, and by addressing certain issues that he leaves unresolved, my somatic-perceptual theory of the emotions will be a stronger and more defensible theory than Prinz' theory of the emotions. Indeed, by resolving some of the problems that Prinz' account faces, I hope not only to strengthen the explanatory power of my own somatic-perceptual theory of the emotions, but to also show that somatic theories can account for the nature of the emotions better than leading cognitive theories.

My approach to developing a theory of the emotions (and to understanding the nature of the emotions in general) is largely interdisciplinary. I believe that the best theory of the emotions should be one that draws on the insights of various academic disciplines, and be both philosophically persuasive and consistent with the most recent empirical research. Indeed, it is important to recognize that emotion is a key topic in cognitive science. This means that the study of emotion is not the privileged domain of any one academic discipline. Instead, the study of emotion is fundamentally interdisciplinary – it encompasses philosophy, psychology, artificial intelligence, neuroscience, linguistics, and anthropology. This is significant because taking an interdisciplinary approach to the topic is, I believe, the only way we will ever come to truly understand the nature of the emotions (and the mind more generally). For this reason, in defending my somatic-perceptual theory of emotions, I will offer both philosophical arguments and empirical findings from neuroscience and psychology. I think philosophical accounts of the nature of emotion can be fruitfully informed by

empirical findings from psychology, neuroscience, and linguistics. Moreover, I believe that the philosophical analysis of empirical data can be useful for understanding how to enhance experimental designs used in the research of emotions. While analytic skills from philosophy help with interpreting empirical data, empirical data helps to narrow down possible philosophical models.

This dissertation divides into five chapters. In chapter one, I offer a sketch of the kind of somatic theory of the emotions I will be defending and distinguish it from other prominent somatic theories of the emotions. The somatic theory of the emotions that I develop and defend throughout this dissertation is closely related to the somatic theories defended by Antonio Damasio and Jesse Prinz. All contemporary somatic theories of the emotions, however, are ultimately derived from William James' somatic-feeling theory of the emotions. For this reason, I outline the somatic theories developed by James, Damasio, and Prinz, and show how they differ from one another, as well as from my own somatic theory. This is important to do because while all somatic theories of the emotions emphasize the role of bodily changes, they nevertheless contain some important differences. For example, on James' account, the bodily changes indicative of emotion are always felt, whereas Damasio argues that they are typically never felt. Prinz, on the other hand, maintains that the bodily changes indicative of emotion can either be felt or unfelt. I argue that my own somatic-perceptual theory of the emotions more closely resembles Prinz' account than it does both James' and Damasio's. I conclude by highlighting points of inconsistencies and areas in Prinz' account where he is unclear about what he takes the basic structure of the emotions to be.

In chapter two, I survey some well-known cognitive theories of the emotions and argue, in opposition to these theories, that cognition is never an essential part of any emotion. I begin by outlining some important differences between the cognitive theories that philosophers defend and the cognitive theories that psychologists defend. I claim that while philosophers typically assume that the cognitions involved in emotions are *propositional attitudes* (i.e., beliefs or judgments), psychologists assume that they are *appraisals* (i.e., evaluations). I argue that this difference may be merely terminological. Nevertheless, it is important to get clear on how theorists from different disciplines use terminology because in the emotion literature, theorists from philosophy, psychology, and neuroscience tend to talk past one another, making it difficult not only to navigate the literature, but also to create a unified theory of the emotions.

In this chapter, I then argue that despite the differences between the cognitive theories that philosophers defend and those that psychologists defend, there are some important similarities. In particular, all cognitive theories of the emotions, in both disciplines, adhere to the following two claims: 1) emotions necessarily involve concepts, and 2) emotions necessarily involve appraisals. I explore these claims and ultimately reject them both on the grounds that although concepts and appraisals are often involved in an emotion episode, they are never *necessary* for an emotion. I then argue that the perception of physiological changes is the *only* thing that is always necessary for emotion. I draw upon three bodies of evidence from neuroscience and psychology that support this: 1) emotions co-occur with bodily changes, 2) bodily changes can induce

emotions, and 3) disruption of interoceptive responses (i.e., decreased sensitivity to stimuli originating inside of the body) leads to diminution of emotions.

In chapter three, I offer a response to a prominent objection against all somatic theories of the emotions – *the intentionality objection*. According to the intentionality objection, somatic theories cannot explain the fact that emotions are typically directed toward particular objects or states of affairs. To date, Prinz is the only somatic-theorist (and non-cognitive theorist more generally) to offer a plausible response to this objection. Indeed, with some important modifications to help strengthen it, I argue that Prinz' response is adequate. Like Prinz, I argue that emotions represent core relational themes (i.e., relational properties that represent the relationship between an individual and his/her environment) by registering changes in the body. However, I argue that emotions represent a broader range of content than Prinz' account of core relational themes seems to allow. I expand on Prinz' account and show that non-cognitive theories of the emotions are able to account for various types of emotional content just as well as prevailing cognitive theories of the emotions. To show how emotions can have a representational content, I draw upon Fred Dretske's theory of psychosomatics.

In the final part of chapter three I discuss an objection that is closely related to the intentionality objection; namely, that non-cognitive theories of the emotions are unable to distinguish emotions from those interoceptive states (i.e., states of the body produced by internal stimuli) that are not emotions (e.g., ear pressure, nasal congestion, etc.) without invoking cognition. I argue that a significant problem with Prinz' account of

the emotions is that it fails to adequately meet this objection (i.e., it fails to explain how to differentiate between emotions and non-emotional interoceptive states). I then offer a solution – in particular, I argue that unlike non-emotional interoceptive states, emotions represent core relational themes. By doing this, I show how my somatic-perceptual theory of the emotions can meet the intentionality objection.

In chapter four, I offer a response to another important objection to somatic theories of the emotions, which I call the *physiology objection.* According to the physiology objection, somatic theories cannot distinguish between emotions with identical physiological responses – e.g., sadness and guilt. I begin by introducing Stanley Schachter and Jerome E. Singer's famous "adrenaline experiment," from which the physiology objection is drawn. I argue that contrary to Schachter and Singer's conclusion, the adrenaline experiment fails to establish that emotions register the same physiological changes in the body. In the second section of the chapter, I then argue that even if it is granted that emotions register the same changes in the body proper (i.e., changes in the autonomic nervous system), recent evidence from neuroscience suggests that emotions register different changes in the cortex – and thus, that emotions are neurologically distinct. Finally, I argue that even if future studies demonstrate that some emotions *do*, in fact, register the same changes in *both* the body proper *and* the cortex, my somatic-perceptual theory of the emotions still allows that they can be distinguished in virtue of their causal histories. In this way, I show that the physiology objection does not pose a problem for my somatic-perceptual theory of the emotions.

In psychology and neuroscience, it is widely agreed that perception can be unconscious. Since on my view emotions are a form of perception, my view suggests that it should be possible for emotions to be unconscious. In the fifth and final chapter, I embrace this conclusion. First, I argue that one *should* embrace the idea that emotions can be unconscious. I then show how the somatic-perceptual theory that I defend provides a framework for thinking about poorly understood psychological conditions, such as alexithymia – which seems to be best understood only on the assumption that emotions can be unconscious.[3] I argue that alexithymia is best understood as a dimensional construct; and as such, it is likely that there exist various degrees of the condition. I then explore the merits of two analogies that others have used to try to understand the nature of alexithymia: one characterizes the condition as an analogue of associative visual object agnosia (i.e., a visual condition whereby an individual is unable to recognize visually presented objects), and the other characterizes it as the emotional equivalent of blindsight (i.e., a visual condition whereby an individual who is blind in a certain area of their visual field is nevertheless able to accurately (or, at least above chance levels) respond to stimuli presented in their blind spot). I argue that the analogy with blindsight more adequately portrays the processing deficit involved in what I refer to as mild alexithymia, but that neither the analogy with blindsight nor the analogy with associative visual object agnosia is able to account for the deficit involved in what I refer to as severe alexithymia. Instead, I argue that the deficit involved in severe alexithymia is best understood through

[3] Alexithymia is a psychological condition whereby individuals are seemingly unaware of their emotional feelings. When asked to report on their emotional state, individuals characterized as alexithymic are typically unable to identify and describe what emotion they are experiencing.

an analogy with apperceptive visual agnosia (i.e., a visual condition whereby an individual is unable to access certain properties of visually presented objects).

As a whole, this dissertation presents and defends a somatic-perceptual theory of the emotions. I reject all cognitive theories of the emotions, and instead build and improve on existing somatic theories. By doing this, I present what I take to be the strongest version of a somatic theory of the emotions yet offered in the literature on the nature of the emotions. According to the somatic-perceptual theory that I defend, emotions are perceptions of physiological changes. I show why my view should be preferred to the cognitive theories that are currently dominating the literature on the emotions in both philosophy and psychology. On my view, cognition is *never* essential to any emotion – instead, I argue that the perception of physiological changes is *alone* necessary for emotion. In addition, I show how somatic theories of the emotions can accommodate both the intentional nature of the emotions and the distinctiveness of different emotions just as well as prevailing cognitive theories of the emotions. I conclude by showing how my somatic-perceptual theory of the emotions allows for unconscious emotions. By doing this, I show that my somatic-perceptual theory of the emotions can help us to better make sense of poorly understood psychological conditions, such as alexithymia. As a result, this dissertation opens the door to future studies in emotional consciousness. For example, if the conscious experience of emotions is really the perception of bodily changes, then an explanation of how emotions become conscious may be incorporated into a general account of how perceptual states become phenomenally conscious.

Although there are a number of emotion theories currently available in the literature, the majority of them are either purely philosophical or purely psychological. I believe that this divide is due to both philosophers and psychologists having a lack of awareness and appreciation of the work being done in one another's disciplines, resulting in impoverished theories of emotion. My somatic-theory of the emotions overcomes this problem by incorporating insights from these (and other) academic disciplines. The interdisciplinary nature of my account, coupled with the fact that it resolves problems associated with current leading somatic theories of the emotions, makes it the strongest somatic theory currently available in the literature.

# CHAPTER ONE:
# SOMATIC THEORIES OF THE EMOTIONS

In this chapter, I offer a sketch of the kind of somatic theory of the emotions I shall be defending in this dissertation and distinguish it from other prominent somatic theories. The somatic theory that I develop in this and subsequent chapters is closely related to other somatic theories of the emotions, especially the theory developed by Jesse Prinz (2004). It is also closely related to the somatic theory that Antonio Damasio (1994, 1999) defends. For this reason, it is important to understand the somatic theories that Prinz and Damasio defend, and to distinguish them from the somatic theory that I defend. Furthermore, since Prinz' and Damasio's somatic theories (and arguably all contemporary somatic theories) ultimately derive from William James' (1884) somatic-feeling theory of the emotions, it will also be useful to distinguish my theory from James' theory.

In 1.1, I outline James' somatic-feeling theory and argue against his view that emotions are exhausted by feelings of bodily changes. In 1.2, I outline Damasio's somatic theory and argue against his position that the only bodily change required for an emotion is when somatic centers in the brain are activated. In 1.3 I outline Prinz' somatic theory and highlight some inconsistencies found in it.

### 1.1 William James' Somatic-Feeling Theory of the Emotions

William James is perhaps the most well-known defender of the view that emotions are mere feelings (i.e., phenomenally conscious mental states). However, it is important to be clear that James (1884) defends a specific type of feeling theory: a *somatic-feeling theory*, whereby emotions are feelings of changes in the body. On this account, our bodies change (e.g., heart rate increases, blood vessels constrict, palms sweat) and the feeling of those changes as they occur *is* the emotion (e.g., fear). Shortly after James proposed his theory of the emotions, Danish psychologist Carl Lange (1885) proposed a strikingly similar theory.[4] James and Lange are believed to have independently arrived at the same conclusion – namely, that we experience emotion *in response* to changes in our bodies. For this reason, James' somatic-feeling theory is commonly referred to as the *James-Lange theory* of emotions.

Common sense suggests that James gets things backwards, and that bodily changes occur only *after* the experience of an emotion. For example, after we become afraid our heart rate increases, our stomach spasms, and we feel fear. However, James claims that this is incorrect, and argues instead that bodily changes precede our emotional experiences. He states: "…we feel sorry because we cry, angry because we strike, afraid because we tremble, and not that we cry, strike, or tremble, because we are sorry, angry, or fearful, as the case may be" (1884, p.190). In support of his view, James conducts a

[4] While Lange restricts the range of bodily states underlying emotion to changes in the vasomotor system (i.e., changes in the constriction and dilation of blood vessels), James has a more inclusive view, and suggests that the range of bodily states underlying emotion includes changes in the viscera, facial expressions, and instrumental actions.

thought experiment. He asks his readers to imagine an emotion without the corresponding bodily feelings:

> "If we fancy some strong emotion, and then try to abstract from our consciousness of it all the feelings of its characteristic bodily symptoms, we find we have nothing left behind, no "mind stuff" out of which the emotion can be constituted, and that a cold and neutral state of intellectual perception is all that remains… Can one fancy the state of rage and picture no ebullition of it in the chest, no flushing of the face, no dilation of the nostrils, no clenching of the teeth, no impulse to vigorous action, but in their stead limp muscles, calm breathing, and a placid face? The present writer, for one, certainly cannot. The rage is as completely evaporated as the sensation of its so-called manifestations, and the only thing that can possibly be supposed to take its place is some cold-blooded and dispassionate judicial sentence…" (1884, p. 193-194).

According to James, this experiment supports his view by demonstrating that without the feeling of bodily changes, there is no emotional experience. Thus, on James' account, *both* bodily changes *and* feelings (i.e., phenomenally conscious mental states) are essential components of an emotion.

While I agree with James that bodily changes are necessary for emotion, I do not agree that feelings are necessary. Instead, on the somatic theory that I defend, feelings *can* be a component of emotion, but they need not be (i.e., feelings are not an essential component of emotion). Therefore, unlike James, I do not defend a somatic-feeling theory of the emotions; instead, I defend what I will refer to as a *somatic-perceptual theory* of the emotions. On this account, there are two distinct components of an emotion: i) a bodily change, and ii) a perception of that bodily change. Thus, I argue that emotions are perceptions of bodily changes.

Perceptions, however, are not necessarily felt (i.e., there can be unconscious perceptions). There is a great deal of scientific evidence that demonstrates that perception can occur below the subjective threshold of awareness (i.e., unconsciously). The most

obvious example of this comes from subliminal perception: if a stimulus is briefly displayed, followed by a mask, we are (depending on the time interval) unaware that the stimulus was ever displayed – though the stimulus can nevertheless affect subsequent behavior. These types of experiments (called masking experiments), demonstrate that information gained through the senses can be processed on an unconscious level.[5] In this way, 'perception' is not equivalent to 'feeling', since feeling entails conscious awareness.

On my account, emotions are perceptions of changes in the body. Perceptual systems in general allow for unconscious perception – we can have unconscious visual states, unconscious auditory states, unconscious olfactory states, and unconscious tactile states. This suggests that on my account it should be possible to have unconscious (i.e., unfelt) emotions. I embrace this result and defend it against possible objections in chapter five. Thus, while on my account emotions *can* be conscious (i.e., felt), they need not be conscious, unlike on James' account.

There is one other way in which the somatic-perceptual theory that I defend differs from the somatic-feeling theory that James defends. First, James uses the term 'somatic' in a limited way. On his account, the range of bodily states underlying emotion is limited to changes in the peripheral nervous system (hereafter, PNS).[6] The PNS is divided into two major parts: the somatic nervous system (which consists of the spinal nerves that innervate the skin, the joints, and the muscles) and the visceral (or autonomic)

[5] Evidence for unconscious emotion can be seen in various other types of experiments – see, for example, Fischman and Foltin, 1992; Winkielman et al., 2005; Arntz, 1993; Strahan et al., 2002.

[6] The nervous system is divided into two parts: the central nervous system (which includes the brain and the spinal cord) and the peripheral nervous system (which includes all the nerves that lie outside of the central nervous system). The primary role of the peripheral nervous system is to connect the central nervous system to our organs and limbs.

nervous system (which consists of the neurons that innervate the internal organs, blood vessels, and glands).[7] On James' account, therefore, emotions are feelings of somatovisceral changes in the body (e.g., feelings of changes in the viscera (i.e., the internal organs –specifically the abdominal organs – of the body; e.g., chest, liver, pancreas, etc.), facial expressions, etc.).[8]

However, following other somatic theorists (e.g., Damasio, 1994, 1999; Prinz, 2004), I use the term 'somatic' more broadly. I expand the bodily states underlying emotion to include changes in both the PNS and central nervous system (hereafter, CNS) (which consists of the brain and the spinal cord).[9] On my account, therefore, the bodily states underlying emotion include changes in the respiratory system, circulatory system, digestive system, musculoskeletal system, and endocrine system. As neuroscience has demonstrated, certain structures in the limbic system (e.g., amygdala) register changes in the levels of chemicals in the brain caused by these various systems.[10] For example, when

[7] The somatic nervous system is primarily under voluntary control (i.e., its responses can be consciously controlled – with the exception of reflexive responses). In contrast, the visceral nervous system is beyond voluntary control (i.e., its responses are involuntary). The visceral nervous system is divided into three parts: the sympathetic nervous system (which is responsible for the activities that increase the body's energy expenditure – e.g., controls the nerves that increase heart rate, dilate pupils, relax the bladder, etc.), the parasympathetic nervous system (which is responsible for the activities that conserve the body's energy expenditure – e.g., controls the nerves that inhibit heart rate, constrict pupil dilation, contract the bladder, etc.), and the enteric nervous system (which is responsible for innervating the viscera – e.g., gastrointestinal tract, pancreas, gall bladder, etc.).

[8] See James (1884).

[9] The CNS is responsible for receiving information from and sending information to the PNS. The brain (which consists of three main parts: the cerebrum, the cerebellum, and the brain stem) is responsible for processing various sensory inputs (from the spinal cord and its own nerves – e.g., optic nerve) and initiating appropriate motor outputs. The spinal cord is the major conduit of information from the PNS to the brain, and vice versa.

[10] The limbic system (sometimes called the Paleomammalian brain) refers to a set of primitive brain structures involved in the regulation of autonomic and endocrine function (particularly in response to emotional elicitors – triggering emotional objects or events), motivation, and memory. The limbic system lies on both sides of the thalamus, just under the cerebrum – and contains the hippocampus, hypothalamus, amygdala, cingulate gyrus, and fornix.

the endocrine system releases an increased level of estrogen (a hormone) and serotonin (a neurotransmitter that regulates moods) in the brain, the amygdala registers these changes and subsequently sends signals to the body and various other brain regions (e.g., prefrontal cortices) thus triggering the enactment of an appropriate body state (e.g., a body state characteristic of anger). And so, unlike James, I use the term 'somatic' to encompass changes in both the brain and the body proper.[11]

James is criticized on just this point by neurologist Antonio Damasio. According to Damasio (1994), by using 'somatic' in a limited way, James does not seem to have considered the possibility that the *only* bodily change necessary for the occurrence of an emotion is when somatic centers in the brain are activated. Moreover, Damasio argues that although James' somatic-feeling theory of the emotions is able to explain the first emotions one experiences in life, it is unable to explain the often complex emotions one experiences in adulthood. In the following section, I outline Damasio's own somatic theory of the emotions, and show how it diverges from James' view.

[11] Presumably sensations (e.g., pain) also necessarily involve the perception of bodily changes. My distinction between these states will become clear in chapter three.

## 1.2 Antonio Damasio's Somatic Theory of the Emotions

In recent times, well-known neuroscientist Antonio Damasio has developed and defended a somatic theory of the emotions. Damasio's theory is heavily influenced by James' somatic-feeling theory. Like James, Damasio maintains that emotions are feelings of changes in the body. However, using recent findings from neuroscience and psychology, Damasio argues that James' somatic-feeling theory is incomplete. Expanding on James' account, Damasio argues: 1) the brain can register bodily changes in the absence of conscious awareness (i.e., emotions can be unconscious), 2) some emotions necessarily involve cognition, and 3) the only bodily change required for an emotion is when certain areas in the brain are activated. While I agree with 1), I think both 2) and 3) are problematic. Indeed, it seems to me that cognition is never an essential component of any emotion. Moreover, I think James was correct to postulate that the body proper is *always* involved in the emotion process.

According to Damasio (1994, 1999), there are two kinds of emotions: primary emotions and secondary emotions.[12] Primary (or universal) emotions are those emotions that an individual experiences early in life – for example, happiness, sadness, fear, surprise, and disgust (see Damasio, 1999, p. 50-51). These emotions, on Damasio's account, are innate (i.e., biologically predetermined) and can be explained by means of a basic mechanism such as the one postulated by James – whereby particular stimuli in the

[12] Damasio (1999) refers to a third kind of emotions – "background emotions" (e.g., enthusiasm, discouragement, edginess, etc.) – which he characterizes as providing an individual with a continuous sense of self. These emotions enable an individual to have what he refers to as "background feelings" (e.g., tension, relaxation, fatigue, malaise, anticipation, dread, etc.). Damasio notes that when a background emotion is sustained over a long period of time, it is best classified as a mood (1994, p. 151).

environment excite, via the limbic system circuitry, a specific pattern of changes in the body proper.

Secondary (or social) emotions, on the other hand, refer to those emotions that an individual experiences only as an adult – for example, embarrassment, jealousy, guilt, and pride (see Damasio, 1999, p. 51).[13] Unlike primary emotions, secondary emotions are not innate; instead, they are acquired throughout an individual's development and are dependent upon primary emotions. Damasio (1994) states: secondary emotions "occur once we begin experiencing feelings and *forming systematic connections between categories of objects and situations, on the one hand, and primary emotions on the other*" (p. 134, italics in original). Unlike primary emotions, Damasio argues that secondary emotions cannot be explained by means of a basic mechanism. Thus, the limbic system circuitry is insufficient to explain secondary emotions.[14]

According to Damasio, one important difference between secondary emotions and primary emotions is that secondary emotions necessarily involve the conscious evaluation of emotion elicitors (i.e., triggering emotional objects or events).[15] In other words,

[13] On Damasio's account, infants and young children do not experience secondary emotions. This is because secondary emotions are derived from an individual's experiences in early childhood and adolescence, when one learns to associate certain emotion eliciting stimuli with negative or positive outcomes (i.e., secondary emotions rely on acquired knowledge).

[14] According to Damasio (1994), evidence from focal brain damage demonstrates that primary emotions and secondary emotions are processed in different cortical areas. He states: "…damage to the limbic system impairs the processing of primary emotion [whereas] damage to prefrontal cortices compromises the processing of secondary emotion" (p. 140).

[15] On Damasio's account, emotion elicitors can actually exist in our external environment (e.g., seeing an injured animal) or simply be imagined (e.g., imagining being told of the death of a loved one). Similar to the way in which visual centers in the brain become active when we form a visual image, somatic centers in the brain can also become active when we imagine an event or scenario that causes us to experience an emotion. What is important to recognize, is that when we imagine an emotion eliciting event or scenario, we seem to *really* experience the emotion (e.g., sadness) as opposed to merely imagine that we are experiencing it. Indeed, when we imagine an emotion eliciting event or scenario, our bodies actually

secondary emotions depend on an individual's evaluation of certain situations. For example, Bob becomes jealous only *after* he reflects on his wife's infidelities. Damasio states: "In many circumstances of our life as social beings…we know that our emotions are triggered only after an evaluative, voluntary, nonautomatic mental process" (1994, p. 130). Since structures in the limbic system are not sufficient to produce these evaluations, Damasio proposes that secondary emotions necessarily also involve structures in the prefrontal and somatosensory cortices (e.g., ventromedial prefrontal cortex).

As we will see, Damasio defends a somatic-feeling theory of primary emotions but an impure cognitive theory of secondary emotions (whereby secondary emotions necessarily require cognition). This move, to a cognitive theory of emotion, is one that I want to resist. In chapter three, I show how all emotions (including Damasio's secondary emotions) can have an intentional structure (i.e., are directed towards particular objects or situations) *without* essentially involving cognition. For now, however, it is important to understand Damasio's use of the terms "proto-self", "feeling", and "core consciousness", as they are integral to his account of emotion.

On Damasio's account, all emotions (both primary and secondary) involve feeling. It is important to be clear, however, that he uses the term "feeling" in a unique way. For Damasio, a *feeling* is always unconscious, and only emerges when an individual perceives that the state of their body has been changed as a consequence of what he refers to as an *internal emotional state* – which he defines as an unconscious neural reaction to an emotion elicitor (i.e., a triggering emotional object or event). In response to an internal

---

undergo changes – e.g., visceral changes, changes in facial expression, changes in skin conductance, etc.) (see Damasio (1994), p. 134-138).

emotional state, one's body is subsequently changed into an *externally observable emotional state.* For example, upon seeing a bear in the woods, certain structures in one's brain (e.g., amygdala, hypothalamus, etc.) automatically send signals to the muscles in one's face and limbs, thus resulting in stiff posture, dropped jaw, raised brows, wide eyes, etc.

To be conscious of one's feeling state, or to have what Damasio refers to as *core consciousness,* an individual must perceive that their representation of their own body state (what Damasio calls the *proto-self)* has been modified by the occurrence of an emotion elicitor.[16] On this account, to have core consciousness is to know or be aware that one has a feeling (1999, p. 280).[17] Damasio's distinction between "feeling" and "core consciousness" (i.e., between having a feeling and knowing that one has a feeling) is unorthodox, since in ordinary language "feeling" implies "conscious awareness".

Damasio acknowledges that his perspective is unusual, and states (1999): "…I am suggesting that 'having a feeling' is not the same as 'knowing a feeling,' that reflection on feeling is yet another step up" (p. 284). To 'know a feeling' (or to 'feel a feeling'), therefore, is a sort of second-order feeling that comes to be known only *after* one's representations of their body state have been integrated in such a way that they give rise to a proto-self (i.e., only after one has a 'basic' feeling). Although this distinction is unusual, Damasio (1999) maintains that it is nevertheless intelligible, since "the proto-self, feelings of emotion, and the feelings of knowing feelings emerged at different points

[16] According to Damasio (1999), the proto-self is "a coherent collection of neural patterns which map, moment by moment, the state of the physical structure of the organism" (p. 177).
[17] Core consciousness involves the use of structures that lie outside of the limbic circuitry (e.g., thalamus and cingulate cortices).

in evolution and to this day emerge at different stages of individual development" (p. 281).[18]

Let us now turn to Damasio's account of how primary and secondary emotions are processed. In the case of primary emotions, an individual need not be able to identify the stimulus that initiates their internal emotional state. As Damasio (1994) states, "…in order to cause a body response, one does not even need to "recognize" the bear, or snake, or eagle, as such, or know what precisely, is causing pain" (p. 131). Instead, all that is required for a primary emotion is the detection and categorization of particular features of an emotion elicitor by early sensory cortices (e.g., primary visual cortex, primary auditory cortex, primary somatosensory cortex, etc.). Information about these particular features (e.g., size, types of motion, types of sounds, etc.,) must then be conjunctively detected by certain structures in the limbic system (e.g., the amygdala).[19]

According to Damasio (1994), the amygdala's "neuron nuclei possess [an innate] dispositional representation which triggers the enactment of a body state characteristic of [for example,] the emotion fear, and alters cognitive processing in a manner that fits [that state]" (p. 131). This process is automatic (i.e., beyond one's control) and occurs below the subjective threshold of awareness (i.e., unconsciously). It is what allows an individual, for example, to respond with alarm by fleeing from a strange sound (e.g.,

[18] Damasio (1999) notes that the "proto-self precedes basic feeling and both [i.e., the proto-self and basic feeling] precede the feeling of knowing that constitutes core consciousness" (p. 281).

[19] The amygdala is an almond shaped mass of nuclei situated in the temporal lobe, just below the cortex on the medial side. Its nuclei are commonly divided into three groups: the basolateral nuclei, the corticomedial nuclei, and the central nucleus. Information from all of the sensory systems projects to the amygdala – particularly to the basolateral nuclei (see Bear et al., 2007, p. 572-573).

growling) without first recognizing the sound as emanating from a particular animal (e.g., a bear).

Secondary emotions, on the other hand, are *not* automatic – instead, they are initiated only *after* the cognitive evaluation of an emotion elicitor. Thus, unlike primary emotions, secondary emotions necessarily contain a cognitive component. Imagine being told about the death of a beloved family member. On Damasio's account, it is only after one consciously forms and evaluates mental images about the family member (e.g., images of the family member, fond memories, one's last encounter with them, etc.) that one will experience an emotion – namely, grief.[20] More specifically, it is only after one forms mental images and judges that the content of those images will have an effect (either direct or indirect) on one's own well-being that one will experience a secondary emotion.

Certain areas in the prefrontal cortex (e.g., ventromedial prefrontal cortex) automatically respond to signals that arise from the processing of one's mental imagery. According to Damasio (1994), the response from the prefrontal cortex comes from *acquired* rather than *innate* dispositional representations – i.e., "representations that embody knowledge pertaining to how certain types of situations usually have been paired with certain emotional responses, in [one's] individual experience" (p. 136). The response elicited from these dispositional representations is then sent to structures in the

[20] As Damasio (1994) states, secondary emotions are initiated by "the conscious, deliberate considerations you entertain about a person or situation. These considerations are expressed as mental images organized in a thought process, and they concern myriad aspects of your relationship with the given person, reflections on the current situation and its consequences for you and others, in sum, a cognitive evaluation of the contents of the event of which you are a part" (p. 136).

limbic system (e.g., the amygdala) where innate dispositional representations trigger changes to both the body proper and the brain.[21]

To this point, I have shown that Damasio defends a somatic-feeling theory of primary emotions (whereby primary emotions are necessarily composed of both a bodily change and a feeling) and an impure cognitive theory of secondary emotions (whereby secondary emotions are necessarily composed of a bodily change, a feeling, *and* a cognitive evaluation). Moreover, I have shown that on Damasio's account, primary emotions are beyond voluntary control (i.e., primary emotions occur automatically) and are processed exclusively in the limbic system. Secondary emotions, on the other hand, are not automatic – this is because they necessarily involve the conscious evaluation of emotion elicitors. Unlike primary emotions, secondary emotions are processed in *both* the prefrontal cortex and the limbic system.

Damasio's conception of emotions is similar to James' insofar as both maintain that all emotions (i.e., both primary and secondary emotions) necessarily contain a bodily change and a feeling.[22] It is important to recognize, however, that unlike James, who claims that the bodily changes indicative of emotion are *always felt* (i.e., always

[21] Although the acquired dispositional representations needed for secondary emotions emerge from the prefrontal cortex, they nevertheless rely on innate dispositional representations in the limbic system for their expression (see Damasio, 1994, p. 137-138). According to Damasio (1994), innate dispositional representations in the amygdala respond to signals from prefrontal dispositional representations by activating nuclei of the autonomic nervous system, dispatching signals to the motor system, and activating the endocrine and peptide systems – all of which trigger changes in the body proper. However, they also activate, "with particular patterns, the nonspecific neurotransmitter nuclei in brain stem and basal forebrain which then release their chemical messages in varied regions of the telencephalon (e.g., basal ganglia and cerebral cortex)" (p. 138). Thus, in addition to triggering changes in the body proper, innate dispositional representations in the amygdala also respond to acquired prefrontal dispositional representations by triggering changes in the brain.

[22] Unlike James, however, Damasio maintains that secondary emotions also necessarily contain a cognitive component – specifically, a cognitive evaluation of an emotion elicitor.

conscious), Damasio maintains that the bodily changes indicative of emotion (both primary and secondary) are only felt if/when one has core consciousness – i.e., if/when one reflects on their representation of their body state (i.e., the proto-self) and recognizes that it has undergone modification because of a particular emotion eliciting stimulus. Indeed, on Damasio's account, emotions are *always unconscious* (i.e., unfelt) unless (or until) one recognizes that their proto-self has been altered.[23]

Moreover, while James argues that the body proper is *always* interposed in the feeling of an emotion, Damasio argues that the *only* bodily change *required* for a feeling is when somatic centers in the brain are activated. On Damasio's account, feeling can be invoked in two ways. First, a feeling can be invoked via what he refers to as the "body loop". In this case, an internal emotional state (i.e., an unconscious neural reaction to an emotion elicitor) leads to changes in the body proper (i.e., changes that do not include changes in the cortex – e.g., skin conductance responses, visceral responses, etc.).[24] These bodily changes are represented in one's somatosensory cortex (which is responsible for processing sensory information). Second, a feeling can be invoked via what he refers to as the "as-if loop". In this case, an internal emotional state *directly* changes the representation of one's body state in their somatosensory cortex – thus, bypassing the body proper.[25] In both cases, feeling emerges when one perceives that their body state has been modified.

[23] As we will see in chapter five, I do not think that feelings are *essentially* involved in an emotion. When they *are* involved, however, I think they can be either felt (i.e., conscious) or unfelt (i.e., unconscious).
[24] An internal emotional state leads to changes in the body proper (and thus an externally observable emotion state) via humoral signals (i.e., chemical messages conveyed through the bloodstream) and neural signals (i.e., electrochemical signals conveyed through nerve pathways) (see Damasio, 1999, p. 281).
[25] See Damasio (1994), p. 155-158, 173-185.

To understand Damasio's objection to James, let us explore these two pathways in more detail. First, consider the body loop. Imagine, for example, seeing a snake. According to Damasio, the image of the snake elicits chemical responses from several bio-regulatory areas of the brain (e.g., neurotransmitter nuclei in the brain stem, basal forebrain, amygdala, anterior cingulate cortex, and hypothalamus). Information about these chemical responses is sent to the prefrontal cortex (specifically, the ventromedial prefrontal cortex), which subsequently triggers certain changes in the body proper (e.g., increased heart rate, skin conductance responses, etc.) and at the same time, sends information about those bodily changes to the somatosensory cortex.[26]

On Damasio's account, the complicated chemical and neural responses elicited by the image of the snake result in the experience of an emotion (e.g., fear). One experiences a "feeling" when they perceive changes to their body state – changes which are initiated by the ventromedial prefrontal cortex and are represented in the somatosensory cortex. However, as we have seen, in order for one to be conscious of this feeling (i.e., to have core consciousness) one has to have an additional perception. In particular, one has to perceive that their representation of their body state (i.e., their proto-self) has undergone a change. Thus, there must be a second-order representation.[27]

According to Damasio, one's proto-self is represented in second-order neural structures (e.g., thalamus, cingulate cortices, etc.). It is one's proto-self that gives rise to an individual's sense of self. He states: "…we only know that we feel an emotion when we sense that emotion is sensed as happening in our organism…[which] comes from

[26] See Damasio (1994), p. 180-184 and Damasio (1999) p. 68-69.
[27] See Damasio (1999), p. 279-284.

representing the proto-self and its changes in second-order neural structures" (1999, p.179). In order to have core consciousness, therefore, it is not sufficient that the changes to one's body state (caused by an internal emotional state) are represented in their somatosensory cortex – it is also necessary that changes to their body state be represented in second-order neural structures. As we can see, on Damasio's account, these second-order representations "are of the relationship between the [individual] and the object (which in this case is an emotion), and of the causal effect of that object on the [individual]" (p. 280).

Let us now consider the as-if loop. In order to experience a feeling via the as-if loop, it is necessary that one first experience that feeling via the body loop. This is because the as-if loop process depends on one's repeatedly associating images of particular emotion elicitors with the body states they trigger. Damasio states:

> "The association between a certain mental image and the surrogate of a body state would have been acquired by repeatedly associating the images of given entitles or situations with the images of freshly enacted body states. To have a particular image trigger the "bypass device," it was first necessary to run the process through the body theater, to loop it through the body, as it were" (1994, p. 156).

Thus, while the body loop is hard-wired at birth, the as-if loop is developed throughout infancy and childhood, as it is dependent on one being familiar with their external environment.

Once again, imagine seeing a snake. Assuming that this is not one's first encounter with a snake, the image of it would elicit the same chemical responses from the same bio-regulatory areas of the brain that would be activated under the body loop (e.g., neurotransmitter nuclei in the brain stem, basal forebrain, amygdala, etc.). Information

about these chemical responses is sent *directly* to both the prefrontal cortex (specifically, the ventromedial prefrontal cortex) and the somatosensory cortex. Unlike the body loop where the prefrontal cortex triggers changes to the body proper, the body proper is bypassed with the as-if loop. That is, the prefrontal cortex does not trigger changes to the body proper – instead, it sends information about the emotion elicitor directly to the somatosensory cortex where the appropriate body state for that stimulus is mapped (or represented).

According to Damasio (1994), when an individual experiences a feeling via the as-if loop, it is "as if the body were being activated and modified", though in actuality this is not the case (p. 155). In other words, when one experiences a feeling elicited via the as-if loop, it *feels* as if some change is occurring in their body proper, when in fact it is not. When the as-if loop is activated, the somatosensory cortex works as if it were receiving signals about a particular body state. By bypassing the body proper, Damasio claims that the as-if loop allows us to "avoid a slow and energy-consuming process" which can be helpful in certain situations (1994, p. 155).

However, the idea that one can have an emotion in the absence of changes to the body proper seems problematic. As we have seen, for Damasio, an emotion necessarily involves a bodily change and a feeling – where a feeling involves the perception that one's body state has undergone modification. If the prefrontal cortex does not send signals to the body proper, as is the case during an as-if loop, then a feeling, on Damasio's account, would emerge from the perception that certain cortical areas (specifically within the somatosensory cortex) have been activated. While this seems

acceptable (albeit unorthodox), things fall apart when we consider having core consciousness of this "feeling". Surely it is not possible for one to be aware of their cortical activations when these activations are not accompanied by changes to the body proper? Indeed, one is certainly not aware of their neural firings in the same way that they are aware of, for example, their heart racing, or their stomach being in knots. The idea that one could have core consciousness of their cortical activations (i.e., the ability to feel the feeling of their cortical activations) is both unattainable and absurd.

It might be argued that this is not a charitable understanding of Damasio's view, since the somatosensory cortex is organized in such a way (i.e., topographically) that when certain parts of it are activated, one automatically experiences a sensation in their stomach despite the fact that receptors in the stomach are not actually being simulated. As such, to have core consciousness of a "feeling" derived from an as-if loop is not, as suggested above, to have awareness of one's cortical activations, rather, it is to have awareness of the sensation in one's stomach.

Although it is true that the somatosensory cortex is topographically organized (see Bear et al., 2007, p. 402-403), Damasio still seems to want to claim that changes in the somatosensory cortex *alone* are sufficient for an emotion. He states: "… in numerous instances the brain learns to concoct the fainter image of an "emotion" body state, without having to reenact it in the body proper…we conjure up some semblance of a feeling within the brain *alone*" (1994, p. 155-156, emphasis mine). If the activation of certain areas in the somatosensory cortex automatically result in one's having core

consciousness of, for example, their stomach being in knots, then Damasio's assertion that an emotion can occur in the absence of changes to the body proper is incorrect.

Moreover, Damasio acknowledges that a feeling yielded by an as-if loop is *not* likely to feel the same as one yielded via a body loop. He states: "I doubt…that [feelings produced within the brain alone] feel the same as the feelings freshly minted in a real body state" (1994, p. 156). On Damasio's view, it is not likely that the human brain can determine how every internal emotional state will affect the body proper. This is because the body proper responds to the chemical and neural signals released by the prefrontal cortex in a slightly different manner every time an internal emotional state is elicited – e.g., one experiences various intensities of emotion depending on one's circumstance, environment, age, etc. He states:

> "What is played out in the [body proper] is constructed anew, moment by moment, and is not an exact replica of anything that happened before. I suspect that the body states are not algorithmically predictable by the brain, but rather that the brain waits for the body to report what actually has transpired" (1994, p. 158).

Additionally, Damasio notes that there are "numerous variables within the body [proper] itself which are not fully represented neurally" (1994, p. 158).

It seems to me that the "feeling" produced by an as-if loop is best understood through an analogy with hallucination. Much as a hallucination resembles a perception (e.g., visual, auditory, tactile, etc.) but is produced in the absence of an external stimulus, so to the "feeling" produced via an as-if loop resembles a perception of a change occurring in one's body proper but is produced when no such change is in fact occurring. However, if there is no change to one's body proper, then one cannot have a perception

of a bodily change – and thus one does not *actually* experience an emotion, even though what they are experiencing may be subjectively indistinguishable from an emotion.[28] Similarly, although a hallucination *resembles* a perception – e.g., both are associated with the same cortical activations (see Weiss & Heckers, 1999) and may be subjectively indistinguishable from one another – it is not, as many philosophers maintain, a true instance of perception (e.g., Martin, 2002; Fish, 2009).[29]

Although certain structures in the limbic system (e.g., amygdala) have been shown to register changes in the levels of chemicals in the brain and subsequently trigger the enactment of appropriate body states (e.g., body states characteristic of anger, fear, joy, etc.), I do not think that Damasio is correct to assert that changes in the CNS *alone* are sufficient for an emotion. For reasons that will become clear in chapter three, on my account, changes in *both* the CNS and the PNS are necessary for an emotion. Since my account emphasizes the importance of changes to both the PNS and the CNS, it is situated in between James' somatic-feeling theory (which emphasizes changes to the PNS) and Damasio's somatic theory (which emphasizes that an emotion *can* occur via changes to the CNS alone).

[28] Recall that on Damasio's account (1994, 1999), an emotion necessarily involves the feeling (or perception) of a bodily change.

[29] An ongoing debate within the philosophical literature on perception centers around the question of how much the mental states resulting from veridical perceptions (i.e., perceptions that coincide with what is actually occurring in one's external environment) and those resulting from hallucinations have in common. While advocates of the disjunctive theory of perception (e.g., Hinton, 1967; Martin, 2002; Fish, 2009) argue that the phenomenal experiences involved in veridical perceptions are of a different nature than those involved in hallucinations (e.g., while veridical perceptions necessarily involve mind-independent objects, hallucinations do not), opponents (e.g., Sturgeon, 1998; Robinson, 1994) argue that veridical perceptions and hallucinations involve mental states of the same basic kind (see Soteriou, 2009).

Like Damasio, contemporary philosopher Jesse Prinz (2004) defends a somatic theory of the emotions that appeals to evidence from recent findings in neuroscience and psychology. Unlike Damasio, however, who defends an impure cognitive theory of secondary emotions, Prinz maintains that *all* emotions (i.e., primary emotions and secondary emotions) can be explained without appeal to cognition. In the following section, I outline Prinz' own somatic theory of the emotions, and show how it diverges from James' and Damasio's somatic theories.

### 1.3 Jesse Prinz' Somatic Theory of the Emotions

Jesse Prinz, a prominent emotion theorist in philosophy, has developed and defended a somatic theory of the emotions that is heavily influenced by James' somatic-feeling theory of the emotions. Like James, Prinz argues that changes to the body proper are essential to an emotion. However, James maintains that the bodily changes indicative of an emotion *must* be felt (i.e., conscious), whereas Prinz maintains that they *can* be felt, but that they need not be. Nevertheless, when these changes *are* felt, Prinz claims that those feelings 'just are' the emotion (see Prinz, 2005, specifically p. 9).[30] While I agree with Prinz that an emotion can be unfelt (i.e., unconscious), I do not agree that when an emotion is felt, the feeling 'just is' the emotion.

On Prinz' (2004) account, emotions are perceptions of changes in the body. Recall, however, that perceptions are not equivalent to feelings – whereas feelings necessarily involve conscious awareness, perceptions do not (i.e., perceptions can be

[30] According to Prinz (2005), "…when an emotion is felt, the feeling literally is the emotion, and there are no other components" (p. 9).

either conscious or unconscious). On Prinz' account, therefore, an emotion can be conscious, but it need not be, unlike on James' account. In saying that emotions are perceptions of changes in the body, Prinz means "only to say that they are states within our somatosensory systems that register changes in our bodies" (2004, p. 58).

It is important to notice, however, that on Prinz' account there is a difference between what an emotion registers and what it represents (see Prinz, 2004, p. 59). According to Prinz, an emotion *registers* changes in the body (e.g., an increase in heart rate), but it *represents* the relationship between an individual's well-being and their external environment. That is, emotions "represent organism-environment relations" (Prinz, 2004, p. 60). Prinz refers to these organism-environment relations as "core-relational themes". [31] For example, sadness represents the loss of something valuable, anger represents a demeaning offence against me and mine, fear represents an immediate and overwhelming physical danger, etc., (see Prinz, 2004, p. 16). Thus, on Prinz' account, the content of an emotion (i.e., what an emotion is about or what it represents) is not (at least exclusively) bodily changes. Instead, an emotion is *caused* by bodily changes, but its *content* consists of representations of core relational themes.[32]

It is important that on Prinz' theory emotions represent more than mere bodily changes. This is because Prinz is offering a somatic theory of emotions, and all such theories are commonly thought to be unable to account for the intentionality of the

[31] In his dimensional appraisal theory, Lazarus (1991) also refers to core relational themes. However, on Lazarus' account, core relational themes are not just the external conditions that elicit emotions – they also correspond to the inner judgments that we make in arriving at emotions.

[32] According to Prinz (2004), emotions must represent core relational themes, otherwise "they would confer no survival advantage, and we could not make sense of the seemingly intelligible uses to which they are put" (p. 60).

emotions. That is, it is argued that somatic theories fail to account for the fact that emotions are typically directed toward an intentional object (i.e., a particular object or event). Since my own theory is also a somatic theory, I too will have to address this objection. This objection (referred to as *the intentionality objection*) and my response to it will be discussed in chapter three.

Although I agree with Prinz that emotions are mental states within our somatosensory systems that register changes in the body, Prinz sometimes speaks about emotions in a rather peculiar way. In particular, he often makes the emotion the grammatical subject of various actions that certain cortical structures are responsible for performing. [33] For example, by claiming that 'emotions register changes in the body' (see Prinz (2004), specifically p. 58, & (2005), specifically pgs. 12, 15, 21-23), Prinz seems to suggest that emotions are themselves identical to the cortical structures responsible for registering changes in the body (e.g., the amygdala). But this seems somewhat strange.

It is important to recognize that the claim that emotions are mental states *is* consistent with the idea that emotions are identical to particular states of the brain. [34] However, one typically does not speak about mental states or brain states as though they are themselves identical to specific structures in the brain, as Prinz often does. Indeed, it is not "the emotion" that registers changes in the body, but rather specific structures in the brain (e.g., the amygdala) that register changes in the body. Presumably, Prinz

[33] See Prinz (2004) & (2005), specifically pgs. 12-15.

[34] The claim that mental states (or mental processes) are identical to brain states (or brain processes) is consistent with the mind-brain identity theory. On this account, mental states (or mental processes) are reducible to brain states (or brain processes) – that is, mental states and brain states are one and the same. This claim is an empirical claim, analogous to the claim that water is identical to $H_2O$ (see Chalmers, 2002, p. 4).

recognizes that emotions are *states* of the brain as opposed to *structures* in the brain – however, the language he uses when speaking about emotions is often ambiguous and vague. In this way, at key points in his account Prinz is unclear about what emotions actually a*re.*

Furthermore, while I agree with Prinz that 'emotions are perceptions of changes in the body', he makes a qualification to this view that I do not support (see Prinz, 2004 & 2005). Specifically, he claims that when the bodily perceptions constituting an emotion are felt, "the feeling [of the bodily change] *is* the emotion" (2005, p. 9, emphasis mine). This makes it sound as though, for Prinz, when the bodily perceptions constituting an emotion are felt, the feeling is the *only* essential component of the emotion. At the same time, however, he stresses the point that the *only* essential component of an emotion is the perception of bodily changes.[35] In this way, Prinz' view seems to be inconsistent.

To be charitable, there are two other ways in which Prinz' qualification could potentially be understood. First, perhaps what Prinz means is that when the bodily perceptions constituting emotions are felt, there are bodily changes which are isometric with the feeling, and these changes are what constitute the emotion. But this cannot be right – since he clearly wants to maintain that the bodily perceptions constituting an emotion can either be felt (i.e., conscious) or unfelt (i.e., unconscious).[36] And, an emotion cannot be a perception of a bodily change and at the same time be identical to a conscious feeling, since a perception of a bodily change is never identical to a conscious feeling.

[35] See Prinz (2004) p. 55-60 (specifically p. 58), and Prinz (2005) p.12.
[36] See Prinz (2004), chap. 9 (specifically pgs. 198-207), and Prinz (2005), section 3 (specifically pgs. 15-18).

Thus, in order for Prinz' view to be consistent, he must claim that feelings which accompany an emotion never themselves 'just are' the emotion (i.e., feeling can never be an essential component of any emotion).

Second, perhaps Prinz means only to say that when the bodily perceptions constituting an emotion are felt, the feeling is the most noteworthy feature of the emotion – since it is the feature that we typically (at least in ordinary language) attribute to an emotion (e.g., "I feel happy, sad, angry, etc.,"). However, if this is the case, then it seems to me that Prinz needs to clarify his position in such a way that this point cannot be misunderstood.

Nevertheless, it seems to me that in cases where the bodily perceptions constituting emotions are felt, both the feeling and the perception of bodily changes are components of the emotion episode. However, only the perception of bodily changes is essential to the emotion. On my account, feeling is not an essential component of any emotion, since it is possible to have the *same* emotion without feeling accompanying it (i.e., the *same* emotion can also be unconscious). In this way, the somatic-perceptual theory that I defend is different than the somatic theory that Prinz defends.

To this point, I have shown that the somatic theories defended by James, Damasio, and Prinz, emphasize the role of bodily changes in an emotion. While James maintains that the bodily changes indicative of emotion are always felt, Damasio argues that they are never felt (unless or until one reflects on their proto-self and recognizes that it has undergone modification due to an emotion eliciting stimulus). Prinz, on the other hand, maintains that the bodily changes indicative of emotion can either be felt or unfelt.

Although I am more sympathetic to Prinz' account than to both James' and Damasio's, my view will diverge from and improve upon Prinz' account by providing a clearer and more consistent account of what an emotion actually is. For example, unlike Prinz, I do not *identify* emotions with cortical structures. Furthermore, on my account, feeling is *never* an essential component of an emotion.

Despite their differences, all of the somatic theories of the emotions I have discussed in this section are *non-cognitive theories* (i.e., they are theories whereby cognition does not play an essential role in emotion). In subsequent chapters, I will develop and defend my own somatic theory of the emotions. However, the majority of emotion theorists defend some sort of *cognitiv*e theory (i.e., a theory whereby cognition is essential to emotion). Thus, it is important to say something about what these theories claim. This will be the focus of the next chapter.

# CHAPTER TWO:
# COGNITIVE THEORIES OF THE EMOTIONS

In the previous chapter, I offered a sketch of my somatic-perceptual theory of the emotions and distinguished it from other prominent somatic theories. However, the majority of emotion theorists reject *all* somatic theories of the emotions. Instead, most defend a *cognitive theory*, whereby an emotion must necessarily contain a cognitive component (i.e., a propositional attitude, such as a belief or a judgment). In this chapter, I examine cognitive theories of the emotions and argue, in opposition to these theories, that cognition is *never* essential to emotions.

This chapter divides into three sections. In 2.1, I highlight some important differences between the cognitive theories that philosophers defend and those that psychologists defend. In 2.2, I examine two claims common to all cognitive theories: i) emotions involve concepts, and ii) emotions involve appraisals. I reject both of these claims on the grounds that although concepts and appraisals are often involved in an emotion episode, they are not *necessary* for emotion. In 2.3, I argue that cognition is never an essential part of an emotion. Drawing on empirical evidence from neuroscience and psychology, I argue that the registration and perception of physiological changes *alone* is sufficient for emotion.

## 2.1 Cognitive Theories in Philosophy and Psychology

Somatic theories of emotion have never been well-received among emotion theorists. This is because it is typically thought that an emotion must be directed towards an intentional object (i.e., a particular object or event). For example, one can be angry that they failed a test, or fearful of the vicious dog next door. On this view, there is a causal link between intentional objects and emotions (i.e., particular objects and events reliably cause emotions). As such, emotions have as their content mental representations of intentional objects, and thus, they necessarily contain a cognitive component (see Nussbaum, 2001; Solomon, 1976; Greenspan, 1988). Theories that hold cognition to be essential to an emotion are called *cognitive theories of emotion.*

There are two types of cognitive theories: *pure cognitive theories* and *impure cognitive theories*. Pure cognitive theories reduce emotions to cognitions (i.e., propositional attitudes, such as beliefs or judgments). On this view, cognition is both necessary and sufficient for an emotion. Notice that since cognition *alone* is sufficient for emotion, and since it is generally agreed that cognitions (i.e., beliefs or judgments) can be unconscious, pure cognitive theorists tend to maintain that emotions can be unfelt (i.e. unconscious) (e.g., Roberts, 1995, 2003; Nussbaum, 2001; Solomon, 1976; Bedford, 1957).

Impure cognitive theories, on the other hand, identify emotions with cognitions and (at least) one other component (e.g., feelings, action tendencies, physiological changes, etc.). Unlike pure cognitive theorists, therefore, not all impure cognitive theorists maintain that emotions can be unfelt (i.e., unconscious) – i.e., some impure

cognitive theorists argue that feeling is a necessary component of emotion (e.g., Spinoza, 1677; Ben Ze'ev, 2000; Greenspan, 1988; Stout, 1899). For these theorists, an emotion consists of (at least) a cognitive component (e.g., a judgment) and a feeling component (e.g., a feeling of pleasure). Notice, however, that emotions can never be *mere* feelings on this view because they *must* include a cognitive component.[37]

It is important to recognize that the cognitive theories (both pure and impure) that psychologists defend (e.g., Lazarus, 1991, 1999; Zajonc, 1984) differ from those that philosophers defend (e.g., Nussbaum, 2001; Greenspan, 1988; Roberts, 2003) in three ways: 1) philosophers tend to assume that the cognitions involved in emotions are propositional attitudes, while psychologists tend assume that they are appraisals, 2) philosophers typically identify emotions with cognitions, while psychologists usually consider cognition to be a precondition for emotions, and 3) unlike most philosophers, psychologists tend to assume that the cognitive component of an emotion is highly structured. Let us look at each of these differences in turn.

[37] Against cognitive theories, one could turn to ordinary language to argue that emotions must be mere feelings. In ordinary language, we often use the terms "emotion" and "feeling" interchangeably (since emotions are typically thought to be paradigm cases of feelings). As Anna Wierzbicka (1999) notes, although some languages do not have a word for "emotion", all languages have a word for "feeling". For example, German, Russian, and Samoan lack a word corresponding to the English word "emotion", but all three have a word corresponding to "feeling" – "gefühl", "čuvstvo", and "lugona", respectively (Wierzbicka, p. 3). Of course, in ordinary language, not all feelings are emotions – we have some purely somatic feelings (e.g., itchiness, thirstiness, chilliness, etc.) – but all emotions are feelings. However, as Prinz (2005) notes, cognitivists respond to this argument by claiming that "ordinary language is not committed to the view that emotions are feelings, but merely to the view that emotions *can* be felt" (p.11 – emphasis mine). In this way, when one makes a statement such as "I feel happy", for example, they don't mean to imply that happiness is a feeling; rather, they mean to say that they are experiencing a feeling that they typically experience when they are happy. "I feel happy", therefore, does not entail "happiness is a feeling." As we saw above, cognitive theorists account for feelings in two ways: 1) by claiming that feeling is necessary, but not sufficient for an emotion, or 2) by claiming that feeling is neither necessary nor sufficient for an emotion. While the former allows us to make sense of statements similar to "I feel sad", the latter view allows us to make sense of statements like "I am sad, but do not feel sad."

First, while philosophers typically assume that the cognitions involved in emotions are *propositional attitudes* (i.e., beliefs or judgments), psychologists assume that they are *appraisals* (i.e., evaluations). It seems to me, however, that this difference may only be one of vocabulary. Propositional attitudes are mental states that consist in a representation of a proposition (i.e., a state of affairs), and an attitude toward that proposition. Take, for example, the proposition "Fred will marry Wilma." Attitudes toward this proposition can include believing, desiring, judging, etc. If one is to direct an attitude of believing (or desiring or judging) toward this proposition the result is a belief (or desire or judgment) – e.g., Fred *believes* that he will marry Wilma.

Emotion theorists in psychology maintain that appraisals are a *kind* of judgment. For example, American psychologist Richard Lazarus (1991) argues that the appraisal which leads to anger is encapsulated by the judgment that there has been a demeaning offence against me and mine.[38] Similarly, psychologist Magda Arnold (1960) argues that the appraisals necessary for an emotion involve judging whether a certain state of affairs is beneficial or harmful, whether the objects involved are present or absent, and whether they are difficult or easy to attain or avoid. Since appraisals take the form of judgments, they are thus reducible to propositional attitudes. This is because, as we saw above, judgments are propositional attitudes *par excellence*.

Second, while those philosophers who hold a cognitive theory of the emotions typically consider cognition to be a constituent *part* of an emotion, psychologists typically maintain that cognition is a *precondition* of an emotion. Take, for example,

[38] See Roberts (2003) for an example of an emotion theorist in philosophy who maintains that appraisals are a kind of judgment.

Lazarus' (1991) dimensional appraisal theory. On this account, an individual must make six appraisals (called "molecular appraisals") before an emotion can ensue. While the first three appraisals (called primary appraisals) establish whether a state of affairs is emotionally significant, the last three (called secondary appraisals) pertain to one's ability to cope with a particular state of affairs.[39] The summary of these molecular appraisals is called a "molar appraisal". It is important to be clear that a molar appraisal is not itself a judgment *per se*; rather, it captures the gist of the six molecular appraisals which precede it. According to Lazarus, a molar appraisal signifies a "core relational theme" (i.e., a relation between an individual and their external environment).[40] Thus, the appraisals that precede, for example, anger, can be summarized by the judgment that there has been a demeaning offense against me and mine.

Notice that on Lazarus' account, emotions are not *identical* to appraisals – rather, they are *caused by* appraisals. Instead, for Lazarus, an emotion is simply a core relational theme that is signified by one's molar appraisal. Thus, for an emotion to arise, one must first evaluate/appraise six specific dimensions of their relationship to their external environment, and subsequently arrive at a molar appraisal. In this way, *all* cognitive theories (i.e., those endorsed by both philosophers and psychologists) afford a central role

[39] Lazarus' (1991) six molecular appraisals are as follows: 1) goal relevance (i.e., an individual must determine whether the interaction with a particular state of affairs is relevant to one's goals), 2) goal congruence (i.e., one must determine whether the interaction with a particular state of affairs will facilitate or impede upon one's goals – if the former, a positive emotion (e.g., joy, elation, etc.) will ensue, but if the later, a negative emotion (e.g., anger, sorrow, etc.) will ensue), 3) type of ego-involvement (i.e., an individual must determine what is at stake during the interaction – e.g., one's life goals, moral values, another person's well-being, etc.), 4) blame or credit (i.e., one must determine the cause of the interaction and whether blame or credit should be assigned), 5) coping potential (i.e., one must estimate the extent to which they will be able to cope with the results of the interaction), and 6) future expectancy (i.e., one must determine whether the results of the transaction will lead to things becoming more or less compatible with their goals in the future).

[40] For an in depth discussion of core relational themes, see chapter three.

to cognition. More specifically, all cognitive theories require that emotions be bound in some way to cognition.

The third difference between the cognitive theories endorsed by philosophers and those endorsed by psychologists is that unlike most philosophers, psychologists typically assume that the cognitive aspect of an emotion is highly structured (i.e., it involves more than just a single appraisal).[41] Indeed, on Lazarus' (1991) account, molar appraisals (which directly precede an emotion) are composed of six different molecular appraisals. On this view, anger, for example, involves more than the judgment that there has been a demeaning offence committed against me and mine – it also involves judgments pertaining to one's coping potential, goals, and blame.

Despite these differences, there are nevertheless significant similarities between the cognitive theories that philosophers defend and those that psychologists defend. In particular, cognitive theorists (both pure and impure) in both disciplines adhere to the following two claims: 1) emotions necessarily involve concepts, and 2) emotions necessarily involve appraisals. In the following section, I discuss these similarities. I argue that we should reject both 1) and 2).

### 2.2 Common Features of all Cognitive Theories of Emotion

Although I deny that cognition plays an essential role in emotions, I nevertheless think that there is a way to reconcile my somatic-perceptual theory (and non-cognitive theories in general) with cognitive theories of emotion. As I noted at the end of the previous

[41] See Ben Ze'ev (1997) for an example of a cognitive theory that is defended by a philosopher and which assumes that the cognitive aspect of an emotion *is* highly structured.

section, all cognitive theories (i.e., both pure and impure cognitive theories) adhere to two claims: 1) emotions necessarily involve concepts, and 2) emotions necessarily involve appraisals.[42] Let us call these two requirements the *concept requirement* and the *appraisal requirement*, respectively. I deny both the concept requirement and the appraisal requirement. Instead, I argue that although emotions *can* involve both concepts and appraisals, it is not necessary that they do.

Let us start by exploring why I do not accept the concept requirement. First, the cognitivist claim that emotions 'involve' propositional attitudes is ambiguous. On one hand, it could mean that emotions *can* (though they need not) be *directed toward* a propositional object. This is relatively uncontroversial. Take, for example, the following proposition: "it is raining outside." It seems perfectly acceptable to claim that one can have an attitude (e.g., belief, desire, or judgment) toward this proposition. For example, Fred *judges* (or desires or believes) that it is raining outside. On this view, an emotion (e.g., happiness) is simply an attitude (e.g., a judgment) that is directed toward a propositional object (e.g., that it is raining outside).

On the other hand, however, the claim that emotions 'involve' propositional attitudes could be interpreted to mean that emotions are necessarily *constituted* (at least in part) by propositional attitudes. On this view, an individual that is experiencing an emotion (e.g., fear) can simultaneously have (at least) two distinct attitudes toward a proposition. Take, for example, the following proposition: "there is a bear in the woods."

[42] Although the concept requirement and the appraisal requirement are common to all cognitive theories, they should not be taken as a definite definition of what it means to be 'cognitive'. This is because cognitive science does not have a single agreed upon definition of 'cognition' (see Prinz, 2004, p. 41).

Two possible attitudes that an individual can have toward this proposition are: 1) the belief (or desire or judgment) that there is a bear in the woods, and 2) the belief (or desire or judgment) that there is an impending danger. Here, the emotion (e.g., fear) is constituted by 2). In this way, an emotion (e.g., fear) is independent of the propositional object that it happens to attach to (e.g., that there is a bear in the woods). One problem, however, is that philosophers who identify emotions with propositional attitudes in this sense do not agree about what those attitudes are – are they beliefs (e.g., Nash, 1989), desires (e.g., Warner, 1980), or judgments (e.g., Nussbaum, 2001; Solomon, 1976)?

Moreover, as Prinz (2004) notes: "if emotions are constituted, at least in part, by propositional attitudes, then having an emotion requires possession of the concepts that would be used to ascribe those propositional attitudes" (p. 23). For example, if sadness is the belief that 'x' is an irrevocable loss, then being sad requires possession of the concept 'x' (e.g., the death of a close family member), and possession of the concept 'irrevocable loss'. In this way, cognitive theorists are committed to the claim that emotions necessarily involve concepts.

The problem, however, is that if emotions necessarily involve concepts, then human infants and non-human animals would be incapable of experiencing emotion (since they are not linguistically competent). But, this does not seem right. Surely when an infant giggles, for example, it is because they are experiencing some sort of emotion (e.g., joy). One cognitivist response to this concern comes from Lazarus (1984), who claims that for all we know, infants and animals *are* making cognitive appraisals (e.g., judgments) whenever they exhibit emotional responses (e.g., crying and cooing). On this

account, continuity in nature demands that simpler creatures (i.e., human infants and non-human animals) have cognitions.

Lazarus' response is extremely unsatisfying. There is currently a heated debate in developmental psychology surrounding the question of whether infants (i.e., babies between the ages of three and eighteen months) have the developmental capacity to attribute mental states (e.g., beliefs, desires, judgments, etc.) to themselves and others (i.e., whether infants have a theory of mind). Research associated with this topic suggests that infants simply do not have the developmental capacity to form the required cognition (e.g., beliefs or judgments) that cognitive theorists claim are essential to emotion (see Frith & Frith, 2003; Baillargeon et al., 2010; Perner & Ruffman, 2005).

Nussbaum's (2001) meta-cognitive theory is particularly problematic (at least prima facie) when it comes to attributing emotions to infants (and non-human animals) because it requires the ability to form judgments about judgments (i.e., it involves metacognition). On this account, emotions are judgments that our evaluative judgments are justified. For example, fear might be the judgment that I am justified in believing that something poses a threat to my well-being. Evidence from developmental psychology, however, shows that humans do not develop the capacity for metacognition until the age of three or four (see Wimmer & Perner, 1983; Perner & Ruffman, 2005).

Nussbaum does not want to deny that infants and non-human animals can have emotions (see Nussbaum, 2001, p. 174-238). On her account, the judgments involved in emotions "always involve thought of an object combined with thought of the object's salience or importance" (2001, p. 23). However, these judgments do not necessarily

involve "the presence of elaborate calculation, of computation, or even… reflexive self-awareness" (Nussbaum, 2001, p. 23). Instead, they only need to consist of "something that we may call conscious awareness" (Nussbaum, 2001, p. 126). In this way, Nussbaum maintains that emotions need not be linguistically formable. She (2001) states:

> "There are many kinds of cognitive activity… in which ideas of salience and importance figure; there are pictorial imaginings, musical imaginings, the kinetic forms of imagining involved in the dance, and others. These are not all reducible to or straightforwardly translatable into linguistic symbolism, nor should we suppose that linguistic representing has pride of place as either the most sophisticated or the most basic mode" (pgs. 127-128).

By arguing that the judgments involved in emotions do not necessarily include reflexive self-awareness or linguistic representation, Nussbaum is seemingly able to overcome the objection that the concept requirement does not allow for infants and non-human animals to have emotions. However, it is difficult to grasp how an infant or non-human animal could make a judgment of the sort required by Nussbaum (i.e., a judgment that necessarily involves the thought of an object and the thought of that object's salience or importance) in the absence of reflexive self-awareness. Indeed, it is difficult to imagine how an emotion can be evoked by a judgment whereby an infant (or non-human animal) acknowledges that a particular object is important to them (e.g., a soother, a bone, etc.), and that they therefore ought to, for example, avoid losing it, without the presence of reflexive self-awareness. For this reason, I am not convinced that Nussbaum's meta-cognitive theory is able to adequately account for infant and non-human animal emotions.

Let us now turn to a description of the appraisal requirement. Cognitive theorists that identify emotions with propositional attitudes presuppose that those propositional

attitudes involve appraisals. An appraisal, or judgment, is a representation that informs an individual about their relationship to the world (i.e., their external environment). In this way, an appraisal is concerned with one's well-being (see Prinz, pgs. 16 & 25).[43] Take, for example, the following proposition: "there is a lost puppy in the driveway." The emotion (e.g., compassion) that one may experience is not identical to the *judgment* that there is a lost puppy in the driveway, since one could judge that there is a lost puppy in the driveway and experience a completely different emotion, such as anger. Instead, the emotion is identical to an *attitude* that one can take toward the proposition – in this case, concern that there is a lost puppy in the driveway. According to cognitive theorists, the attitude that one has (in this case, concern that there is a lost puppy in the driveway) plays the role of informing them that they want to help relieve another being's suffering (e.g., they want to help the puppy find its way home).

The appraisal requirement is common to all cognitive theories. However, it seems to me that similar to concepts, appraisals are not necessary for emotion. Although emotions *can* involve appraisals, it is not necessary that they do. By accepting that emotions *can* involve both concepts and appraisals, I accept that emotions *can* have an intentional structure (i.e., they can be directed at some particular object, event, or proposition). But, how can emotions be intentional and, at the same time, not essentially involve cognition? This will be the focus of chapter three.

[43] Appraisals and judgments are not synonymous – instead, appraisals are a *kind* of judgment. It is important to recognize that not every appraisal or judgment that an individual makes is concerned with their well-being. However, according to cognitive theorists, the judgments involved in emotions are *always* concerned with an individual's well-being (e.g., Nussbaum, 2001; Roberts, 2003).

For now, however, it is interesting to note that Prinz (2004) argues that impure cognitive theorists adhere to a third claim; namely, that the cognitive components bound to emotions are disembodied (i.e., the cognitive components bound to emotions consist of something over and above bodily changes or internal states that register bodily changes). In other words, the propositional attitudes bound to an emotion must be distinguished from the somatic component(s) of that emotion. Thus, the cognitive and somatic component(s) that are necessary for emotions are not identical.

Unlike Prinz, I do not regard cognitive theorists as necessarily adhering to this claim. This is because it seems to me that a large number of cognitive theorists would argue that the cognitive components bound to our emotions are necessarily *embodied* (i.e., the cognitive components bound to our emotions necessarily involve, or are reducible to, the somatic components of emotions) (see for example, Place, 1956; Smart, 1959). For example, perhaps the cognitive components of emotions are simply brain-state types (i.e., they are identical to internal states of the brain). On this view, the cognitive components of our emotions would *necessarily* be embodied.

Despite defending a non-cognitive theory of the emotions (i.e., a somatic-perceptual theory of the emotions) I do not claim that emotions can *never* contain a cognitive component – to make such a claim would be absurd. After all, one can undoubtedly experience an emotion that has a cognitive aspect. For example, I may become angry when reflecting upon a friend's hurtful comments. There seem to be many situations in which one's emotions include some type of evaluative mental process.

However, it nevertheless seems to me that cognition is *neither* necessary *nor* sufficient for an emotion. That is, although an emotion *can* contain a cognitive component, this component is not essential to the emotion (i.e., the cognitive component can be excluded from the emotion episode without our losing 'the emotion'). It is important to be clear that emotions *can* be complex. That is, they *can* contain various components (e.g., bodily changes, cognitions, feelings, action tendencies, etc.). But, the *only* necessary feature of *any* emotion is the perception of bodily changes. In this way, emotions *can* have a relatively simple structure. In the next section, I present three bodies of evidence from neuroscience and psychology that support this claim.

### 2.3 Emotions as Perceptions of the Body

There are three bodies of evidence that can be used to support the claim that emotions are perceptions of bodily changes: 1) emotions co-occur with bodily changes, 2) bodily changes can induce emotions, and 3) disruption of interoceptive responses (i.e., decreased sensitivity to stimuli originating inside of the body) leads to diminution of emotions. On their own, these bodies of evidence do not definitively show that bodily changes are necessary for emotion; however, taken collectively, they offer compelling support in favor of a somatic theory of the emotions (including my own somatic-perceptual theory). Let us examine each body of evidence in turn.

First, there is evidence that emotions co-occur with bodily changes. For example, Levensen, Ekman, and Friesen (1990) studied the autonomous changes associated with Ekman and Friesen's (1971) six basic emotions (i.e., anger, fear, disgust, sadness, joy,

and surprise), and found that every emotion corresponds to a unique bodily pattern. Moreover, every neuroimaging study of the emotions shows excitation in areas of the brain associated with bodily response (see Damasio et al., 2000). It has also been observed that the principal brain structures underlying our emotional states (e.g., cingulate cortex, insular, and somatosensory cortex) are all independently associated with bodily responses (see Damasio et al., 1999; and Critchley et al., 2004).[44]

In response to this evidence, cognitive theorists often argue that only *certain* emotions co-occur with bodily changes – and that more advanced emotions (e.g., guilt, jealousy, love etc.) do not have associated bodily states (see Solomon, 1976; Griffiths, 1997). However, as Prinz (2005) notes: "to date, every neuroimaging study of a more advanced emotion has shown activation in exactly the same brain regions that are implicated in our more ancient emotions" (p.5). For example, Shin et al. (2000) demonstrated cingulate and insula activation during guilt episodes, Buss et al. (1992) demonstrated galvanic skin responses (i.e., a change in the skin's ability to conduct electricity) during jealousy episodes, and Bartels and Zeki (2000) demonstrated cingulate and insula activation when subjects were viewing photographs of the people they love.

In addition to correlational evidence, there is evidence that bodily changes can induce emotions. For example, Zajonc et al. (1989) demonstrated that change in facial musculature is, on its own, sufficient to induce an emotional response – even if we are unaware that we are making an emotional expression. Moreover, there is evidence that emotions can be elicited via pathways from early visual structures – for example, the

[44] The amygdala seems to play a role in emotion elicitation, but it is not necessary for emotions themselves (see Prinz, (2004)).

pulvinar and superior colliculous to the amygdala – which instructs other structures to perturb the body (see LeDoux, 1996; Morris et al., 1999). This evidence suggests that emotions can arise in the absence of any cognitive component. However, demonstrating that there is a correlation between emotions and bodily changes is not sufficient to demonstrate that bodily changes are *necessary* for emotions (and thus, that somatic theories of the emotions are correct). To establish necessity, one has to show that disruption of interoceptive responses (i.e., decreased sensitivity to stimuli originating inside of the body) leads to diminution of emotions.

Unfortunately, evidence that the disruption of interoceptive responses leads to diminution of emotions is not conclusive – but it *is* suggestive. Hohmann (1966) demonstrated that individuals who suffer from spinal cord injuries (reducing feedback from the body), often report diminished emotional response. He further demonstrated that the higher the injury was on the spinal cord, the more severe the deficit in emotional response. This is because the higher the injury is on the spine, the less bodily feedback there is available. However, spinal cord damage alone does not eliminate emotions. This is because visceral information (i.e., information about the internal organs of the body) is still capable of travelling through the vagus nerve (i.e., a cranial nerve that runs from the brain stem to the colon).

However, as Bechara (2004) demonstrated, damage to the vagus nerve causes emotional impairment. And so, it is likely that a patient with damage to both their spinal cord and vagus nerve (depending on the location of each injury) would have greater emotional deficits than would a patient with damage to only one of these structures.

Furthermore, Damasio and Van Hoesen (1983) demonstrated that damage to the cingulate cortex (a center for interoception) can lead to akinetic mutism (i.e., a profound deficit in emotional response).

Collectively, these findings support the conclusion that the perception of physiological changes is not just correlated with emotions – but, that it is actually *necessary* for emotions. However, while this might be sufficient evidence to invalidate pure cognitive theories of emotions, it does not show impure cognitive theories to be false (since cognitive theorists can argue that emotions require cognition *and* registration of bodily changes – e.g., Greenspan, 1988). In order to invalidate impure cognitive theories, it would need to be shown that cognition is only contingently associated with emotions. That is, it would need to be shown that registration of bodily changes alone is sufficient for emotions.

There are a number of different findings that support this. First, changing one's facial expression and intensity of respiration influences self-reported emotions – even when one is unaware that they are changing their facial expressions (see Zajonc, 1989; Levensen et al., 1990; Philippot et al., 2002). Since this can occur in the absence of awareness, the process is apparently not mediated by cognitive labeling (see Strack et al., 1988). Moreover, Blood and Zatorre (2001) demonstrated that non-vocal music elicits patterned bodily changes, which leads to predictable emotional responses. Furthermore, drugs (such as adrenalin) that enervate the autonomous nervous system have emotional effects (see Marañon, 1924). Although these findings suggest that induction of bodily change results in emotional experience, it is nevertheless conceivable that cognition is

also generated. However, there is simply no good reason to think that this is the case, besides a desire to hold on to a cognitive theory. As Prinz (2005) notes: "to presume that subjects in these various conditions are all making [cognitive] judgments would be *ad hoc*" (p.7).

Although none of the evidence that I have presented in this section is conclusive, it nevertheless presents serious problems for cognitive theories of the emotions (both pure and impure). At the same time, it provides strong initial support for my somatic-perceptual theory of the emotions, on which emotions are perceptions of changes in the body.

In this chapter, I have shown that all cognitive theories (i.e., both pure and impure) involve two requirements: 1) the concept requirement – whereby emotions necessarily involve concepts, and 2) the appraisal requirement – whereby emotions necessarily involve appraisals. Although I deny both 1) and 2), I nevertheless acknowledge that emotions *can* (though they need not) involve concepts and appraisals. In this way, I accept that emotions *can* have an intentional structure. Determining how emotions be intentional, and at the same time, not essentially involve cognition will be the focus of the next chapter.

# CHAPTER THREE:
# THE INTENTIONALITY OBJECTION

Within the philosophical literature, the most persistent objection against somatic theories of emotion (and non-cognitive theories more generally) is the *intentionality objection.* This objection holds that somatic theories (and all non-cognitive theories) are unable to explain the intentional content of emotions. That is, they are unable to account for the fact that emotions are directed toward a particular object or event. For example, when someone is angry, they are typically angry *at* someone (or something) for something that they (or it) have done or failed to do. In this way, there is some object, event, or proposition that a person's anger is directed toward.[45]

As we saw in the previous chapter, all cognitive theories of emotion hold cognition to be necessary for emotion. According to these theories, emotions are reliably caused by intentional objects, and thus, have as their content mental representations of intentional objects. As such, cognitive theories are able to account for the intentional nature of emotions. However, if emotions are perceptions of changes in the body, as I argue they are, how are we to explain their intentional content? It would be absurd to claim that emotions represent particular bodily states. After all, when one is sad, for

[45] The intentionality objection against non-cognitive theories of emotions is typically advanced by cognitive theorists, including, for example: Solomon (1980); Neu (2000); Nussbaum (2001); Roberts (2003); Broad (1971); Bedford (1957); Greenspan (1988).

example, one is sad *about* something (e.g., the death of a loved one) – it would be ridiculous to think that what one's sadness is about is a particular state of one's own body. Determining how emotions can have an intentional nature without essentially involving cognition will be the focus of this chapter.

Prinz' (2004) response to the intentionality objection has been invaluable to somatic theories of emotion (and non-cognitive theories more generally). Not only is his response currently the most elaborate in the emotion literature, but it also appears to be the most viable. Indeed, with some important modifications to help strengthen it, I think Prinz' response to the intentionality objection is adequate. My objective for this chapter is to show how Prinz' basic account can be strengthened, and to argue that this strengthened account is capable of meeting the intentionality objection.

This chapter divides into six sections. In 3.1, I offer a sketch of Prinz' response to the intentionality objection. In 3.2, I highlight Fred Dretske's (1981, 1986) theory of psychosemantics and show that emotions represent core relational themes by registering changes in the body. In 3.3, I argue that emotions represent a broader range of content than Prinz' account of core relational themes seems to allow. I expand on Prinz' account and show that non-cognitive theories of the emotions are able to account for various types of emotional content just as well as prevailing cognitive theories of the emotions. In 3.4 I turn back to Dretske's theory of psychosemantics and consider a prominent objection against it. I show that Prinz' basic strategy for responding to the objection is appropriate. In 3.5, I discuss a well-known objection against all somatic theories of the emotions; namely, that they are unable to distinguish emotions from those interoceptive states (i.e.,

states of the body produced by internal stimuli) that are not emotions (e.g., ear pressure, nasal congestion, etc.). I argue that unlike non-emotional interoceptive states, emotions represent core relational themes. In 3.6., I show that Prinz fails to adequately explain how one's mental representations of particular objects link-up to emotions. I argue that the relationship between emotions and mental representations of particular objects is best understood in terms of causality.

By resolving some of the inconsistencies and gaps in Prinz' response to the intentionality objection, I hope to: 1) strengthen the explanatory power of my somatic-perceptual theory of emotions, and 2) show that somatic theories in general are able to account for intentionality just as well as prevailing cognitive theories.

### 3.1 Prinz' Response to the Intentionality Objection

Emotions are typically (indeed, perhaps always) directed towards particular objects, events, or states of affairs. For example, if John is afraid of the dog next door, his fear is directed towards the dog next door; if Jane is sad about the death of her grandfather, her sadness is directed towards the loss of her grandfather; if Mary is proud of a grade she received on a math test, her pride is directed towards her achievement on the math test, etc. In this way, emotions have an intentional nature – i.e., they have as their content particular objects. Just as a belief is a belief *about* a particular situation or state of affairs, and every desire is a desire *for something*, an emotion, it seems, is always *directed towards* someone or something. However, if emotions do not necessarily involve cognition, as somatic theorists maintain, how can we account for the intentional nature of

emotions? On the face of it, it seems that we cannot. In the philosophical literature, this is referred to as the *intentionality objection.* Traditionally, this objection has been the most difficult for somatic theorists (and non-cognitive theorists more generally) to overcome.

As we saw in chapter one, Prinz defends a somatic theory of the emotions according to which emotions are mental states within our somatosensory systems that register changes in our bodies (see Prinz 2004, p. 58). On Prinz' account, emotions register a variety of physiological changes – from changes in the visceral organs and skeletal muscles, to changes in hormone levels and temperature. However, if emotions are merely perceptions of changes in the body, it seems that their intentional content (i.e., what emotions are directed towards) would consist of particular states of the body. For example, anger would represent, or be directed towards, dilated blood vessels, a clenched jaw, and a scowl (among other things). But this, according to cognitive theorists, is not satisfactory.

There are two reasons for this. First, if emotions merely represented particular states of the body, they would qualify as intentional in an uninteresting and trivial sense – they would simply represent the body as being in a certain state. However, emotions are not trivial. Not only do emotions play a significant role in our everyday lives (e.g., we strive for happiness and try to avoid sorrow), there is also evidence that (at least some) emotions derive from our early ancestors to confer survival advantage, and thus, are biologically specified via genetic inheritance (see LeDoux, 1996; Damasio, 1994). Second, reflection on one's own emotional experiences will demonstrate that emotions do not usually have as their content particular states of one's body. Indeed, not only are our

emotions typically directed towards some particular object (or state of affairs) in our external environment (e.g., the President of the United States, an upcoming party, etc.), but they also seem to be subject to rational assessment.[46] However, if emotions were merely perceptions of changes in the body, then they would not have the appropriate type of intentional content required to be considered either rational or irrational. So, how do somatic theorists overcome these problems, and thus, overcome the intentionality objection?

To date, Prinz (2004) is the only somatic theorist (indeed, the only non-cognitive theorist in general) to offer a plausible response to the intentionality objection. Even Damasio, whose somatic theory is highly praised in both the philosophical and psychological literature, is unable to explain the intentional nature of secondary emotions.[47] According to Damasio (1999), the intentional nature of primary emotions can be explained in terms of an automatic basic mechanism: particular stimuli in one's environment (e.g., a spider, a great height, a loud sound, etc.) excite, via the limbic system circuitry, a specific pattern of changes in the body proper, which subsequently

[46] As Prinz (2004b) notes, "emotions are somatic, but they are also fundamentally semantic: meaningful commodities in our mental economies" (p. 45). That is, emotions can be "right or wrong, appropriate or inappropriate, warranted or unwarranted, rational or irrational" (Prinz, 2004b, p. 54). It is the intentional nature of our emotions that allows us to assess whether they are rational (and thus warranted), or irrational (and thus unwarranted). Cognitive theorists (both pure and impure) take a rational emotion to be one in which the propositional attitudes or appraisals involved in the emotion are reasonable (e.g., Pitcher, 1965; Solomon, 1976). For example, imagine that you see a grizzly bear walking through your camp site. Appraising this situation as dangerous (and thus experiencing fear) is reasonable. On the other hand, appraising this situation as a progression towards the realization of a goal (and thus experiencing happiness), is not reasonable (unless, of course, one either has suicidal tendencies or is passionate about grizzly bears and encounters them regularly).

[47] Recall that Damasio distinguishes between *primary emotions* (i.e., universal emotions that an individual experiences early in life – e.g., happiness, sadness, fear, surprise, and disgust) and *secondary emotions* (i.e., social emotions that an individual experiences only as an adult – e.g., jealousy, embarrassment, guilt, and pride) (see Damasio, 1999, p. 51).

results in the experience of a particular emotion (see pgs. 50-51). In this way, primary emotions have as their intentional content particular objects in the environment. The intentional content of secondary emotions, on the other hand, cannot be explained by the limbic system circuitry. This is because, as we saw in chapter one, secondary emotions are more complex than primary emotions in that they necessarily involve the *conscious evaluation* of emotion elicitors (i.e., triggering emotional objects or events). Damasio, therefore, contends that unlike primary emotions, secondary emotions necessarily involve the perception of bodily changes *and* cognitive evaluations (see p. 139). Thus, in accounting for the intentional nature of secondary emotions, Damasio turns his back on somatic theories of emotion (and on non-cognitive theories more generally).

According to Prinz (2004b), Damasio concedes too much by accepting that cognition is necessary for explaining the intentional content of *any* emotion (see p. 55). This is because, for Prinz, *all* emotions *only essentially* involve the perception of bodily changes. On this account, emotions are fundamentally embodied. He says: "… we should not feel compelled to supplement embodied states with meaningful thoughts; we should instead put meaning into our bodies and let perceptions of the heart reveal our situation in the world" (Prinz, 2004b, p. 58). Again, however, if emotions do not essentially involve cognition, how are we to explain their intentional nature?

In exploring Prinz' response to the intentionality objection, it is crucial to first understand his distinction between the *particular objects* and the *formal objects* of emotions, and, second, to establish how he thinks emotions can be mental representations (i.e., mental states that carry information about that which reliably causes them). Let us

explore each of these topics in turn. In accounting for intentionality, Prinz appeals to Anthony Kenny's (1963) distinction between an emotion's particular object and its formal object. On this account, "a *formal object* is the property in virtue of which an event elicits an emotion, and a *particular object* is the event itself" (Prinz, 2004, p. 62). Thus, for example, a snake may be the particular object of one's fear, but the snake causes fear in virtue of being perceived as an imminent threat – an imminent threat being the formal object of fear.[48] Likewise, the death of a loved one can be the particular object of one's sadness, but the death causes sadness in virtue of being perceived as a loss – loss being the formal object of sadness, etc.

According to Prinz, an emotion (e.g., fear) may be directed toward any number of particular objects, but its formal object will *always* be the same. Although certain situations (or objects) have the potential to elicit different emotions in different people (e.g., I may be afraid of cats, but you may find them to be a source of great joy), Prinz nevertheless maintains that there are some similarities in the external causes of all emotions. These similarities, he argues, are core relational themes: "all people are

[48] According to Prinz, the formal object of fear is not, as I maintain, an imminent threat. Instead, he argues that the formal object of fear is danger. But danger does not seem to capture the extension of fear in all situations. There are a number of objects and states of affairs (e.g., snakes, spiders, tigers, radiation, global warming, etc.) that have the property of being dangerous, but that nevertheless may not cause someone to experience fear. For example, one might recognize (pre-reflectively, and thus, in the absence of cognition) that the tiger is dangerous, but if one is accustomed to being around tigers, or if one does not feel threatened by the tiger then it is likely that one will not experience fear. However, if the tiger poses an imminent threat (e.g., if it escaped from its enclosure at a zoo and began to chase after a person), it is highly probable that the person will enter into a state of fear even if he/she is accustomed to being around tigers. Indeed, it seems to me that in those situations where one perceives that one is facing an imminent threat, one will subsequently experience fear. For this reason, I think the formal object of fear is best understood as an imminent threat. I do not think Prinz would deny that it is possible for one to experience fear in the absence of danger. After all, this is a typical occurrence in the case of irrational fears – e.g., one may experience fear when they hear the buzzing of a bee nearby. When this occurs, however, it is not because one has represented the bee as dangerous, as Prinz' account seems to suggest; rather, it is because one's fear has misfired (i.e., one has misrepresented the bee as an imminent threat).

frightened by scary things, angered by offensive things, disgusted by disgusting things, and elated by pleasing things" (Prinz, 2004, p. 60).[49] As we will see later, on Prinz' account, core relational themes *just are* the formal objects of emotions. More specifically, they are relational properties of states of affairs that pertain to an individual's well-being and are responsible for eliciting emotions. This means that for Prinz emotions are elicited by situations as they relate to *us*.[50] And so, although you and I may disagree about what counts as an imminent threat, we are nevertheless afraid of those situations that appear to be an imminent threat *to us*.[51] In this way, regardless of what one is fearful of (e.g., dogs, large bodies of water, heights, failure, death, etc.) the formal object of one's fear will always be an imminent threat. Likewise, the formal object of one's sadness will always be the loss of something valuable, the formal object of guilt will always be the transgression of a moral imperative, the formal object of anxiety will always be uncertainty, etc. The following table contains examples of how the formal

[49] Notice that Prinz seems to change the specification of the formal object of fear from danger to scariness. He does this again later in the same discussion. He states: "…fear represents the property of being scary to me…" (2004, p. 60). It seems to me that this is a mistake on Prinz' part, and that he would have done better to maintain that the formal object of fear is *always* an imminent threat.

[50] See table 3.1 for an illustrative example of what I take the formal objects of different emotions to be. My list generally reflects Lazarus' (1991) original formulation of core relational themes. However, I make some important modifications. For example, unlike Lazarus, I do not divide fear into "fright (i.e., physical threats)" and "anxiety (i.e., existential threats)". This is because there seems to be a number of things one can be fearful of that do not fall into either one of these two categories – e.g., public speaking. Instead, I maintain that the formal object of fear is an imminent threat.

[51] At first glance, this makes Prinz' view sound trivial: I am angered by things that I find offensive, disgusted by things that I find disgusting, etc. Prinz recognizes this worry, and admits that on the face of it, this aspect of his position appears "uncomfortably circular" (2004, p. 60). He states: "saying that fear represents the property of being scary to me is like saying fear represents whatever scares me. But that is like saying that fear represents whatever causes fear. A vacuous insight" (2004, p. 60). To escape this circularity, Prinz claims that emotions do *not* represent response-dependent properties (i.e., properties that depend on being represented a certain way). As such, something can be dangerous, for example, even if one does not represent it as being dangerous – e.g., global warming would be dangerous even if we did not know or believe that it is. Prinz states: "fear represents the property of being dangerous even though that property is possessed by some things that we do not in fact fear" (2004, p. 64).

objects of some common emotions might be specified. This is not quite the way that Prinz specifies them, although the categories and basic strategy are the same. The reason for this will emerge as the chapter proceeds.

**Table 3.1: emotions and corresponding core relational themes**[52]

| *Emotion* | *Core Relational Theme* |
|---|---|
| Anger | A demeaning offense committed against me or someone (or something) I care about. |
| Disgust | My taking in or being too near an indigestible object (or an unpleasant idea or action). |
| Fear | An imminent threat faced by me or someone (or something) I care about. |
| Happiness | My progression towards (or success in attaining) a desired goal. |
| Sadness | An irrevocable loss faced by me or someone I care about. |
| Surprise | A violation of my expectations. |

Since Prinz conceives of emotions as a kind of perception, it is important to get clear on what he takes emotions to *represent*. Typically, perceptions are thought to represent particular objects (e.g., I perceive by means of hearing the wail of a siren). According to Prinz (2004), however, it is an error to conceive of emotions as representing particular objects. He states:

[52] This way of specifying the formal objects of different emotions is not quite the same as Prinz'. For example, I maintain that the formal object of fear is an imminent threat faced by me or someone (or something) I care about, whereas Prinz maintains that the formal object of fear is an immediate and overwhelming physical danger (2004, p. 16). The reasons for the differences between my preferred way of specifying the formal objects of different emotions and Prinz' will emerge as the chapter proceeds.

> Arguments for the claim that emotions have intentionality often appeal to the fact that emotions can be directed at some particular event… This is a flawed form of argument. While there is a sense in which emotions are directed at particular events, that does not mean that they represent those events… The events are represented by mental states that combine with emotions. When I am sad about the death of a child, I have one representation of the child's death and I have sadness attached to that representation. The sadness doesn't represent the death… Sadness represents the loss of something valued (Prinz, 2004, p. 62).

Thus, according to Prinz, when an individual is sad, for example, he/she has two mental representations: one representation which corresponds to the particular object of one's sadness (e.g., the death of a loved one), and the other, which corresponds to the formal object (i.e., loss). Since emotions, on Prinz' account, represent core relational themes, it follows that the latter representation, the emotion (i.e., sadness) represents the formal object (i.e., loss). It is important to keep in mind that for Prinz, emotions are perceptions of changes in the body that represent core relational themes. And so, although he fails to emphasize this point, it follows that the representation one has of a formal object represents that object by registering a certain pattern of change in one's body state.[53] In other words, although Prinz accepts that emotions can be directed towards particular objects (and thus that emotions can have intentionality), he nevertheless denies that emotions *represent* particular objects. Rather, they represent some *feature* the object possesses – a *relational* feature, one the object possesses by standing in some relation to me.

Interestingly, Prinz also denies that emotions represent bodily changes. If emotions are perceptions of changes in the body, then it would seem to follow (at least

[53] As we will see in 3.6, it is unclear, on Prinz' account, precisely how one's sadness (i.e., one's perception of a certain pattern of change in one's bodily state) links up with, or attaches to, the representation of an emotion's particular object.

intuitively) that they represent those changes. However, Prinz maintains that there is a difference between what an emotion registers and what it represents. On his account, emotions *register* changes in the body (e.g., an increase in heart rate, temperature fluctuations, etc.), but they *represent* core relational themes.

Following prominent American psychologist Richard Lazarus (1991), Prinz identifies core relational themes with the formal objects of emotions, and maintains that they are fundamentally concerned with an individual's well-being.[54] As we have seen, for Prinz core relational themes represent the relationship between an individual and their environment – what Prinz sometimes refers to as "organism-environment relations" – and are responsible for eliciting emotions. He states:

> …sadness represents loss (or some disjunction of closely related properties, including loss, privation, defeat, and so on). Loss is not a state of an organism. It is not a bodily condition. Nor is it something purely external…. Loss is a relational property. It is the elimination of something valued by an organism. Sadness represents the elimination of something valued by me (Prinz, 2004, p. 63).

According to Prinz, therefore, the content of an emotion (i.e., what an emotion represents or is directed towards) is not, as one might assume, bodily states; instead, it consists in the representation of a core relational theme.

At this point a quick clarification is in order. Prinz is inconsistent in his description of core relational themes – sometimes he refers to them as "relations" (see for

[54] Lazarus' (1991) dimensional appraisal theory of the emotions is largely dependent on core relational themes. Although both Prinz (2004) and Lazarus agree that emotions represent core relational themes, they nevertheless disagree on how best to characterize them. On Prinz' non-cognitive theory of the emotions, core relational themes are the external conditions that elicit emotions. Lazarus' cognitive theory of the emotions, on the other hand, holds that core relational themes are not just the external conditions that elicit emotions – instead, they also correspond to six judgments (i.e., six molecular appraisals) that he claims we make in arriving at emotions. Thus, Lazarus maintains that core relational themes are involved in the *structure* of our emotions, whereas Prinz maintains that they merely capture the *content* of our emotions.

example, Prinz, 2004, p. 60), sometimes as "properties" (see for example, Prinz, 2004, p. 228), and other times as "relational properties" (see for example, Prinz, 2004, p. 158). This inconsistency in Prinz' claims about what core relational themes are is a potential source of confusion. Nevertheless, as I understand him, Prinz conceives of core relational themes as *relational properties* (i.e., properties that depend on one thing's relationship to other things – e.g., John's property of being taller than Mike). On his account, these properties belong to actual states of affairs (or objects) and pertain to an individual's well-being. Core relational themes are relational properties on Prinz' account because the same state of affairs (or object) that causes an emotion in one person may not cause the same emotion (or any emotion) in another person. For example, the same event can signify the loss of something valuable for one person and thus cause them to become sad, but not for another person (since they may value different things).

To this point, I have explained Prinz' distinction between the particular objects of emotions and the formal objects of emotions, and have shown that on his somatic theory, emotions are mental states that represent their formal objects – i.e., core relational themes. But *how* is it that our emotions represent core relational themes, on Prinz' view? More specifically, how can emotions be perceptions of changes in the body, and at the same time not represent those changes, but instead represent core relational themes? Prinz' response to these questions relies on philosopher Fred Dretske's (1981, 1986) theory of psychosemantics. In the next section, I outline Dretske's theory and consider a prominent objection against it.

### 3.2 Dretske's Theory of Psychosemantics

In the philosophical literature, a theory of psychosemantics is a theory that explains how a mental state (such as an emotion) can represent something as relevant to one's interests or concerns.[55] According to Dretske, a mental representation is a mental state that carries information. In order to carry information, however, a mental representation must satisfy two conditions: 1) it must be reliably caused by something, and 2) it must have been set in place (via learning or evolution) to detect that thing. Thus, a mental state S represents X *if and only if* S is reliably caused by X, *and* S has been set in place (via learning or evolution) for the purpose of detecting and carrying information about X.[56] As Prinz puts it, Dretske's general idea is that "…a mental representation is a mental state that has been *set up* to be *set off* by something" (2004, p. 54).

Thus, to verify that emotions represent something, Prinz must show that they have been set up (either by learning or evolution) to be set off by something particular, and that they therefore carry information about that particular thing. To do this, he must determine what reliably causes our emotions. Since emotions are perceptions of changes in the body, Prinz maintains that *bodily changes* are what *reliably cause* our emotions, and hence that emotions carry information about bodily changes. He denies, however,

[55] Mental representations have semantic or intentional properties – that is, mental representations have a particular content (i.e., they represent, or are about, something particular) and truth conditions (i.e., they can either be true (i.e., veridical) or false (i.e., non-veridical)). Currently, there is no single agreed upon theory of psychosemantics (see Warfield & Stich, 1994, & Fodor 1987). It is beyond the scope of this dissertation to survey all of the various theories of psychosemantics that appear within the philosophical literature. Instead, I here focus only on Dretske's theory of psychosemantics – since Prinz' theory of the emotions, as well as my own, relies on it.

[56] It is important to recognize that reliable causation does not necessarily entail perfect correlation. As Prinz states, "to say that one thing is reliably caused by another thing just means that, were the second thing to occur, then, all things being equal, the first would have a high probability of occurring as a result" (2004, p. 53).

that emotions *represent* bodily changes. This is because evolution (via natural selection) only confers certain characteristics, traits, behaviors, etc., that provide us with a survival advantage. But it is doubtful, according to Prinz, that evolution would have equipped us with emotions solely in order to detect bodily changes. After all, how does the detection of bodily changes provide an advantage to survival? Prinz states:

> If evolution furnished us with emotions in order to detect bodily changes, then detecting bodily changes must confer a survival advantage. This is a strange hypothesis. It is not clear why it is advantageous to know when my blood vessels are constricting. That knowledge is not, in itself, especially useful for survival (2004, p. 59).

On Prinz' account, therefore, bodily changes are not appropriate candidates for the things that emotions *represent*. Instead, he argues that bodily changes are best understood as that which *reliably cause* emotions, which are themselves perceptions of these changes.

And so, another candidate for what emotions *represent* is required. While emotions are reliably caused by bodily changes, Prinz argues that bodily changes are reliably caused by core relational themes.[57] He states: "…emotions rarely

[57] Notice that on Prinz' view the bodily changes associated with emotions are *causal intermediaries*. This means that there is no direct causal link between the perception of a core relational theme and an emotion (e.g., the perception of a core relational theme → an emotion). Instead, the bodily changes associated with an emotion serve as a necessary causal link between the perception of a core relational theme and an emotion (e.g., the perception of a core relational theme → certain pattern of bodily changes → an emotion). The intermediary role played by the bodily changes associated with emotions is perhaps best compared to the role of the primary visual cortex (V1) during colour perception. One cannot perceive the colour of an object without first having certain changes occur in V1 (e.g., perception of an object → certain pattern of changes in V1 → perception of colour). Both the bodily changes involved in emotions and the changes in V1 involved in colour perception are necessary intermediate causal links within their own perceptual domains – they are necessary for the occurrence of certain perceptual experiences. Notice, however, that there is an important disanalogy between emotion and colour and perception – the causal intermediaries involved in colour perception (i.e., changes in V1) are not themselves perceived (i.e., one does not perceive the changes occurring in V1), whereas the causal intermediaries involved in emotion (i.e., certain pattern of bodily changes) *are* perceived (i.e., one perceives the pattern of bodily changes associated with emotions as they occur).

begin from the inside. They are ordinarily elicited by some external situation. In order for emotions to represent external conditions, it would have to be the case that emotions are *reliably* caused by those conditions" (2004, p. 60). And so, while sadness, for example, is reliably caused by a certain pattern of bodily changes, those bodily changes are in turn reliably caused by the loss of something or someone. In this way, emotions appear to carry information about *both* bodily changes *and* core relational themes.

Unlike information about bodily changes, however, information about core relational themes confers a survival advantage (e.g., by enabling one to flee when one is faced with an imminent threat). This is because core relational themes are directly related to an individual's well-being and goals. It is likely, therefore, that emotions have been set in place, either by evolution or learning, to carry information about core relational themes. On this basis, Prinz concludes that emotions represent core relational themes.[58]

To this point, I have shown that on Prinz' account emotions register certain patterns of changes in the body but do *not* represent those changes. To say that anger, for example, represents heart palpitations and/or blood vessel constriction overlooks the important role that anger plays in our lives. Anger is not a heart or a blood pressure monitor; instead, it serves to tell us that a demeaning offense or an unfavorable action has been committed (or has the potential to be committed) against us or someone (or

[58] Prinz states: "To say that fear represents a racing heart fails to explain the important role that fear plays in our lives. It fails to explain why an inability to predict emotional consequences makes us bad decision makers. The claim that emotions represent core relational themes fares much better. Fear seems to be a danger warning system, not a heart monitor. We flee because we are faced with dangers, not because we have palpitations. This is the key" (2004, p. 66).

something) of significant importance to us. I think Prinz is right to postulate that although emotions register certain changes in the body, they nevertheless represent core relational themes. However, I do not think his account of 'core relational themes' is as robust as it could be. In the next section, I show how Prinz' account of core relational themes can be extended to allow for a greater range of emotional content.

### 3.3 Extending Prinz Account: Empathetic Emotions and Imagined States of Affairs

According to Prinz, core relational themes are relational properties that belong to actual states of affairs (or objects) and pertain to an individual's well-being. On his account, therefore, an emotion represents *actual states of affairs* that are directly relevant to *one's own well-being*. But emotions seem to have a greater range of content than this. Indeed, it seems to me that emotions also represent: 1) relational properties of states of affairs that pertain to the well-being of family and friends, human strangers, brute animals, and the natural environment, and 2) imagined, remembered, and envisaged states of affairs. Let us explore these ideas in turn.

First, not *all* relational properties of states of affairs (or objects) that elicit emotions are directly relevant to one's own well-being. Imagine, for example, sitting in a waiting room with your younger brother (or some other close family member) who is about to receive dental surgery. He has been afraid of both needles and the dentist for as long as you can remember. The longer you sit and wait, the more obvious his fear becomes – e.g., his skin is pale and damp with sweat, he is breathing rapidly, and he is more fidgety than usual. Despite the fact that *you* are not about to have dental surgery,

and that you have never been afraid of needles or the dentist, you nevertheless enter into a state of fear. It seems to me that in this situation, you experience fear because you have the capacity for empathy (i.e., the capacity to recognize, understand, and share in the emotion(s) being experienced by another sentient being). You see that your brother is in distress (via his physiological reactions), recognize that he is afraid, and enter into a state of fear.[59]

As this example demonstrates, emotions do not merely detect relational properties in states of affairs that pertain to one's own well-being, but they also seem to detect relational properties in states of affairs that pertain to the well-being of one's family (and close friends or loved ones). Emotions seem to have the biological function of tracking core relational themes – but core relational themes are not, as Prinz sometimes suggests, merely concerned with one's own well-being. Instead, emotions also appear to track core relational themes that pertain to the well-being of the following (in order of distance from concern for one's own well-being): i) family and friends – e.g., one may become happy upon learning that one's grandparent has overcome cancer, ii) human strangers – e.g., one may become sad when reading the news that a child was murdered in a neighboring province, iii) brute animals – e.g., one may become angry if one witnesses an animal being abused, and iv) the natural environment – e.g., one may become sad when a fire

[59] In witnessing this situation one may experience an emotion other than fear – for example, anger. Empathy apparently involves understanding what it is like for another person or sentient being to have some experience (i.e., empathy involves being able to put oneself in another's shoes, so to speak). And so, if one does not have a fear either of needles or the dentist, one *may not* be able to empathize with one's brother. One may instead experience anger because one's brother is unable to control his fear, or one may not experience any emotion at all (which may necessarily be the case for individuals with various psychological conditions (e.g., schizophrenia, depression, post-traumatic stress disorder, psychopathy, etc.)).

destroys a large section of forest.[60] Let us call those emotions that are elicited by the detection of relational properties in states of affairs that pertain to the well-being of other sentient beings/organisms, *empathetic emotions*.[61]

Since it served in our ancestors' survival, we seem to have evolved (via natural selection) the capacity for empathy. Not only have we evolved to respond to the needs of our off-spring (or loved ones) by being empathetic, but we have also evolved to respond to human strangers, brute animals, and the natural environment in the same way. Empathy appears to have evolved not only out of our need to protect our offspring (and loved ones), but also because our species depends on a thriving natural environment and mutual co-operation (i.e., we are more successful when we are surrounded by other's who are healthy and capable) (see de Waal, 2005). Although emotions may well have evolved prior to empathy, they nevertheless appear to be closely related to it.[62]

I do not think Prinz wants to deny that emotions track core relational themes that are concerned with the well-being of others. After all, in his discussion about core relational themes, Prinz claims that anger represents "a demeaning offence against me

[60] Notice that the scope of empathetic emotions extends beyond classical cases of empathy that might ordinarily be considered appropriate.

[61] It seems to me that empathetic emotions are more easily explained by a non-cognitive theory of the emotions than by a cognitive theory of the emotions. While empathetic emotions *can* involve cognition (i.e., can involve a propositional attitude, such as a belief or a judgment), it is not necessary that they do. It seems to me that empathetic emotions generally tend to occur in the absence of cognition. For example, imagine hearing a loved one cry out in pain. One hears their loved one cry out, recognizes (pre-reflectively and thus in the absence of cognition) that he/she is afraid (via their physiological responses), and automatically enters into a state of fear themselves. Indeed, it seems to me to be clear that one does not necessarily have to form a belief or a judgment that their loved one is afraid (or in pain) prior to entering into a state of fear themselves.

[62] It is interesting to note that the cortical structures primarily responsible for the experience of empathy are the same structures that are centrally involved in our experience of emotions – e.g., the anterior insular cortex (see Mutschler, I., et al. 2013).

*and mine"* (2004, p. 16, emphasis mine). This suggests that, on his account, anger can be elicited when an offence has been committed *either* against me *or* against someone (or something) that I care about. Nevertheless, Prinz does not discuss this idea in any detail, nor does he explicitly make mention of whether other emotions (e.g., sadness, fear, joy, etc.) can represent core relational themes that pertain to the well-being of others. Instead, he tends to speak as though emotions *only* track core relational themes that pertain to one's *own* well-being (e.g., see Prinz, 2004, pgs. 60-66). Although Prinz is not wrong to claim that emotions represent core relational themes that pertain to one's own well-being, his view is incomplete – emotions have a greater range in the way that they hook-up to well-being than he suggests.

Second, it is unclear whether Prinz' claim that emotions represent core relational themes means that emotions only represent *actual present states of affairs* as they relate to well-being. Prinz sometimes suggests that this is the case. For example, he states: "…emotions represent relations between external states and ourselves" (2004, p. 60). But this seems somewhat problematic – surely emotions represent more than just actual present states of affairs? It seems to me that emotions also represent imagined, remembered, and envisaged states of affairs. For example, one may become fearful when one imagines being in an enclosed space, become sad when reflecting on a situation in one's childhood where one was bullied, or become elated when one thinks about a future trip to the Caribbean.

Again, my claim is not that Prinz wants to *deny* that emotions can represent imagined, remembered, or envisaged states of affairs – after all, he acknowledges that

emotions can be elicited via imagination. For example, he states: "emotions can be elicited by actual environmental conditions impinging on an organism or by imagined conditions, as when an emotion is caused by contemplating a future event" (2004, p. 62). However, Prinz is unclear about what he takes the relationship between emotions and imagined, remembered, and envisaged states of affairs to be.[63]

Since Prinz accepts that emotions can be *elicited* via imagined states of affairs, it seems to follow that he should also accept that emotions can *represent* relational properties of imagined states of affairs that are concerned with well-being. This is because imagined states of affairs often involve visualizing actual present states of affairs as being a certain way. And so, emotions generated via imagination serve the same purpose that Prinz argues emotions generated via actual present states of affairs serve (i.e., emotions generated via imagination track core relational themes that pertain to our own well-being and the well-being of those that are of significant importance to us). By extending Prinz' account of core relational themes to include empathic emotions and emotions elicited by imagined, remembered, and envisaged states of affairs, I hope to show that a non-cognitive theory of the emotions can account for a broad range of emotional content just as well as prevailing cognitive theories of the emotions.

[63] Prinz' discussion concerning emotions elicited via imagination is somewhat surprising (see Prinz, 2004, pgs. 49-50). Typically, when one is interested in determining whether emotions can be elicited via imagination, one is interested in exploring whether imagining certain states of affairs (or objects) could result in one experiencing an emotion – e.g., if one imagines jumping from a great height does one experience fear? Prinz, on the other hand, appears to be more interested in determining whether emotions themselves can be generated in one's imagination – e.g., is it possible for one to create the feeling of fear within one's imagination? It is unclear why Prinz directs the discussion in this way. After all, when one enters into an emotional state via imagination, it is generally because one imagines a particular state of affairs (or object) – not because one imagines the feelings associated with a certain emotion.

### 3.4 What Emotions Represent: The Disjunction Problem

Let us now return to the question of what emotions represent. Prinz' somatic theory of the emotions relies on Dretske's theory of psychosemantics. This allows Prinz to develop a powerful and plausible account of the intentional content of emotions, one that is consistent with a non-cognitive theory. It is important to recognize, however, that Dretske's theory is not without its problems. Within the philosophical literature, the most frequent objection put forth against Dretske's theory is that it fails to assign precise contents to our mental states (i.e., a mental state could potentially represent a variety of different things rather than just one particular thing). This objection is sometimes referred to as the "disjunction problem" (e.g., Fodor, 1990).

Consider, for example, the mental representations of anteaters. Anteaters typically feed on ants and termites. As such, anteaters must have mental representations that inform them when ants or termites are present in their immediate environments. These mental representations can perhaps be thought of as internal nutrient detection systems. Various types of spiders (e.g., zodariidae spiders), however, are known to mimic the appearance and behavior of ants in order to prey upon them (this mimicry is sometimes referred to as "myrmecomorphy") (see Cushing, 1997). Anteaters are thus often unable to discriminate between an actual ant and a zodariidae spider mimicking an ant. In this way, anteaters' mental representations are reliably caused by ants but are also reliably caused by zodariidae spiders.

As such, it is unclear, when an anteater's internal nutrition detection system goes off, whether its mental representation represents an ant or a zodariidae spider. Perhaps the

mental representations represent insects in general rather than merely representing ants? Or, perhaps they represent the property of being a small, black, crawling object? Of course, they could also represent nutrition more generally. This is the problem with Dretske's account: there is no way to determine whether anteaters' mental representations represent ants or zodariidae spiders or insects or small, black, crawling objects, or nutrition, or a disjunction (i.e., a combination) of all of these things. After all, these ways of specifying the relevant mental content are all capable of satisfying Dretske's conditions on representation (i.e., they are all reliably caused by something and have been set in place by learning or evolution to detect that thing).

According to Prinz, one way to help determine the content of mental representations, and thus overcome the disjunction problem, is by looking at a species' past to see how the thing being represented has been used over time. In the case of anteaters, studies have shown that the mammals have always subsisted almost exclusively on a diet of ants and termites (see Redford, 1985). And so, although they may, on occasion, eat zodariidae spiders, it is highly unlikely that zodariidae spiders are what an anteater's internal nutrient detection system was *designed* to detect.

I think Prinz' basic strategy for responding to the disjunction problem is correct. Recall that on Dretske's account, a mental representation has the function of carrying information about that which reliably causes it to occur. Although an anteater's internal nutrient detection system is reliably caused by (i.e., becomes active as a result of) encounters with ants, it may, as we have seen, also be reliably caused by zodariidae spiders engaging in mimicry. But, this is simply because ants and zodariidae spiders

engaging in mimicry have a similar appearance – and it stands to reason that any state caused by the one insect can be caused (at least on occasion) by the other. Consequently, an anteater's internal nutrient detection system carries information about *both* ants *and* zodariidae spiders. However, an anteater's internal nutrient detection system was set up (via evolution or learning) to reliably detect and/or respond to ants – *not* zodariidae spiders. Anteaters do not actively seek out zodariidae spiders as a source of nutrition; instead, they typically only consume zodariidae spiders when the insects are engaged in mimicry. It is probable, therefore, that an anteater's internal nutrient detection system was acquired because of its numerous encounters with ants as opposed to zodariidae spiders. Although an anteater's internal nutrient detection system carries information about *both* ants *and* zodariidae spiders, and is reliably activated by both insects, it only *represents* ants because it was set up as a result of ant encounters and with the purpose of detecting ants.

Additionally, Dretske maintains that mental states are set up in such a way that they *can* be activated in error. This further demonstrates that there are (at least some) constraints on the content of mental representations. On Dretske's account, when the mechanism responsible for producing mental representations is activated in response to something other than what it has been set up to represent, it has been erroneously activated. Thus, when an anteater's internal nutrient detection system is activated by zodariidae spiders engaging in mimicry, it is activated in error – that is, it misrepresents zodariidae spiders as ants.

According to Prinz, the content of our emotions is restricted in much the same way. On his view, emotions carry information about core relational themes because that is what they have been set up by evolution (or learning) to do, and they can be activated in error. On the face of it, however, it does not seem accurate to claim that the content of an emotion can be restricted in this way. After all, there is a range of situations that can elicit every emotion, and what causes me to experience a particular emotion may not cause you to experience the same emotion (or may not cause you to have an emotional experience at all). Consider, for example, all of the situations that could cause one to become sad: the breakdown of one's marriage, a rejection letter, the death of a loved one, misplacing one's favorite pair of shoes, finding out that one is adopted, being unable to conceive a child, etc. While all of these situations would cause me to experience sadness, only one or two may cause you to experience sadness. Indeed, sadness (and all other emotions) seems to have a broad range of content.

According to Prinz, although there may be a range of situations that one's sadness can be directed toward, they are all similar in one important respect: they all involve losing something that one takes to be of value (e.g., relationships, self-esteem, careers, family, etc.). On this account, loss is a relational property – i.e., a property of a state of affairs that depends on the elimination of something that is valued.[64] This explains why

[64] On Prinz' account, emotions do *not* represent response-dependent properties (i.e., properties that depend solely on being represented a certain way). For example, although sadness represents losing something valuable, something can still be a loss for me without my representing it as such. For example, imagine that I value a certain painting that my grandmother made for me. If this painting were stolen, I would be sad. I may represent the theft of this painting as a loss. Notice, however, that the loss does not depend on being represented as such – it merely depends on whether I consider the painting to be *of value.* Since I value the painting, its being stolen would be a loss for me even if I did not represent it as such – it would be a loss even before I made the discovery that it has been stolen.

different situations cause different people to become sad. The same state of affairs that causes sadness in one person may not cause sadness in another person because they may value different things. For example, Jane might be sad upon learning that she is unable to conceive a child. Mary, however, might be indifferent when presented with the same news. This is because Mary may not value having children. But, should Mary lose something that she values (e.g., her beloved pet), she may become sad.

One of the most interesting features of Prinz' (2004) response to the intentionality objection is his assertion that emotions are *always* accompanied by a second mental state (see pgs. 62-63). This second mental state represents an emotion's particular object. For example, if Bob is sad about the death of his beloved pet, he will have one mental representation which corresponds to the death of the pet (i.e., the particular object of his sadness), and another – the emotion – which corresponds to there having been a loss (i.e., the formal object of his sadness). When these mental states are combined, Bob has a complex mental representation: sadness at the loss of his pet.[65]

In this way, for Prinz, the content of our emotions is actually quite limited. Although every emotion can be directed toward a range of particular objects, Prinz maintains that emotions *only* represent their formal objects (i.e., core relational themes). And so, one can be fearful of a number of different things (e.g., heights, death, dogs, spiders, etc.), but one's fear will *only* ever represent an imminent threat. Of course, it is possible for one to be fearful in the absence of an imminent threat (and sad in the absence of actual loss, etc.) – such as, for example, when one's imagination is overstimulated.

[65] Determining how emotions link up to representations of their particular objects will be the focus of section 3.6.

However, this does not threaten the proposal that fear represents an imminent threat (or that sadness represents loss, etc.). Instead, this merely demonstrates that emotions can occur in error – which, as we have seen, satisfies Dretske's stricture that mental representations must be capable of being erroneously applied.

Prinz' response to the intentionality objection, especially when modified in the ways I have suggested, represents a promising way of addressing the intentionality objection to non-cognitive theories of the emotions. Nevertheless, it faces some further problems. Perhaps the most significant problem with Prinz' account is that it does not adequately explain how to differentiate between emotions and other interoceptive states (i.e., states of the body produced by internal stimuli) that are not emotions (e.g., ear pressure, nasal congestion, etc.). In the next section, I discuss this objection and offer a solution.

### 3.5 Distinguishing Emotions from Interoceptive States that are Not Emotions

On Prinz' view, as on my own view, emotions are states in our somatosensory cortices that register certain patterns of changes in the body. This means that for Prinz, and for me, emotions are *interoceptive states* – i.e., states that respond to stimuli originating inside of the body (see Prinz, 2004, 2006). However, if emotions are interoceptive states, one must be able to distinguish them from those interoceptive states that are not emotions (hereafter *non-emotional interoceptive states* – e.g., ear pressure, fatigue, nasal congestion, etc.). This is a challenge that all somatic theories of the emotions (and all non-cognitive theories more generally) must meet in order to successfully overcome the

intentionality objection.[66] In this section, I argue that Prinz' somatic theory of the emotions fails to adequately distinguish emotions from all non-emotional interoceptive states. I then argue that it is possible to do better.

If emotions are perceptions of changes in the body, as Prinz and I both claim, then how is one to distinguish between an emotion, for example, fear, and a non-emotional interoceptive state, such as, for example, fatigue, which is also a perception of change in the body?[67] According to Prinz, emotions can be distinguished from non-emotional interoceptive states in virtue of their intentionality. While non-emotional interoceptive states *both* register *and* represent changes in the body, emotions register changes in the body but represent core relational themes. Thus, on Prinz' account, fatigue registers certain changes in the body (e.g., increased levels of serotonin, neuromuscular transmission, etc.,) *and* also represents those changes. Fear, on the other hand, registers

[66] To the best of my knowledge, Prinz (2004) is the only non-cognitivist about emotions to have developed a response to this objection. Most non-cognitive theories of the emotions do not appear to address this objection in any capacity. For example, somatic-feeling theorist William James (1884) argues that emotions are feelings of changes in the body, but he does not explain how emotions can be distinguished from non-emotional interoceptive states. In addition, prominent American psychologist Carroll Izard (1984) explicitly states that "emotion has no cognitive component [and that] the emotion process is bounded by the feeling that derives *directly* from the activity of the neurochemical substrates (p.24). Although Izard denies that emotions involve cognition, he nevertheless fails to explain how emotions can be distinguished from non-emotional interoceptive states (the feeling of which is also directly derived from the activity of certain neurochemical substrates). When explaining how emotions can be distinguished from non-emotional interoceptive states, most non-cognitive theorists about emotions end up turning their backs on non-cognitive theories of emotions. Instead, they tend to argue in favor of an impure cognitive theory of the emotions on which emotions, unlike non-emotional interoceptive states, necessarily involve cognition (see for example Damasio (1994, 1999); Kenny (1963)). Although Jenefer Robinson (2005) does not directly discuss this objection, it seems to me that her non-cognitive theory of the emotions may be able to overcome it. According to Robinson, emotions are not themselves cognitive, but they *can* be caused by cognitive activity that precedes the emotion episode. Thus, on her account, emotions may be distinguished from non-emotional interoceptive states in virtue of their being elicited via a cognitive appraisal. It remains unclear, however, how Robinson would distinguish emotions that are *not* elicited via a cognitive appraisal from non-emotional interoceptive states.

[67] Prinz acknowledges that this objection poses a major difficulty for his account of the emotions. He states: "my definition of emotions has difficulty distinguishing [non-emotional interoceptive states] from emotions. It seems to cast the net too wide" (2004, p. 190).

certain changes in the body (e.g., increased heart rate, blood vessel constriction, etc.) but represents a core relational theme (i.e., an imminent threat).[68]

Luca Barlassina and Albert Newen (2013), contemporary philosophers of the mind, have criticized Prinz' response to this objection. They maintain that although Prinz' approach may distinguish emotions from *some* non-emotional interoceptive states, it does not successfully distinguish emotions from *all* non-emotional interoceptive states. According to Barlassina and Newen, Prinz' somatic theory of the emotions adheres to the following two claims: i) emotions, unlike non-emotional interoceptive states, are those interoceptive states that represent core relational themes; and ii) a mental state represents a core relational theme if it is reliably caused by it and has been set in place to carry information about it (see 2013, p. 18). However, they argue, by adhering to both i) and ii) Prinz cannot successfully distinguish between emotions and all non-emotional interoceptive states. This is because, on their account, some non-emotional interoceptive states *do* represent core relational themes.

Consider, for example, an earache. An earache is a non-emotional interoceptive state that registers certain changes in the body. However, some earaches also appear to represent a core relational theme (i.e., some earaches appear to represent a relational property of a state of affairs that pertains to one's wellbeing) – e.g., low levels of atmospheric pressure. Indeed, an earache is reliably caused by low levels of atmospheric pressure, and since it is useful for one to know whether one is in an unfavorable

[68] Prinz distinguishes between what he refers to as a mental representation's *nominal content* (i.e., the bodily changes it registers) and *real content* (i.e., the relational properties of a state of affairs that it represents) (see 2004, p. 142). Thus, on Prinz' account, while non-emotional interoceptive states have bodily changes both as nominal content and real content, emotions have bodily changes as nominal content but core relational themes as real content.

environment, it is probable that one evolved the feelings associated with an earache to carry such information. Thus, if one accepts Prinz' approach to distinguishing emotions and non-emotional interoceptive states, Barlassina and Newton argue that one is forced to conclude that an earache is an emotion. But, it is absurd to claim that an earache is an emotion.[69] For this reason, Barlassina and Newen maintain that one should reject Prinz' approach to distinguishing emotions and non-emotional interoceptive states – and thus reject his somatic theory of the emotions more generally.[70]

To the best of my knowledge, Prinz has not addressed Barlassina and Newen's concerns. However, I do not think Prinz would accept that an earache is an emotion. Instead, it seems to me that Prinz would probably respond to Barlassina and Newen by disputing their claim that an earache represents a core relational theme. In his discussion about the differences between emotions and non-emotional interoceptive states, Prinz claims that his proposal, on which emotions can be distinguished from non-emotional interoceptive states in virtue of their intentionality, finds support in the fact that the stimuli typically responsible for eliciting emotions, unlike the stimuli typically responsible for eliciting non-emotional interoceptive states, can be perceived through the senses (e.g., sight, hearing, smell, etc.). He states:

[69] I do not want to deny that an earache can be *accompanied* by an emotion (e.g., fear), only that an earache itself is an emotion.

[70] According to Barlassina and Newen (2013), for Prinz to distinguish between emotions and *all* non-emotional interoceptive states, he must abandon i) – i.e., the claim that emotions, unlike non-emotional interoceptive states, are those interoceptive states that represent core relational themes. However, Barlassina and Newen note that by giving up i) Prinz must abandon the idea that emotions are interoceptive states – thus abandoning the idea that emotions are non-cognitive mental states.

> "…emotions can typically be caused by a broad range of stimuli, perceived through a broad range of senses. We become afraid when we see, hear, or smell things. If fear represented a bodily state, it would not regularly require some external cause, perceived through [an] externally directed sense. In contrast, fatigue is generally elicited by a condition that originates within the body. The same can be said about [other non- emotional interoceptive states]" (2004, p. 190).

In this way, Prinz suggests that emotions, unlike non-emotional interoceptive states, are generally elicited by *actual present states of affairs in the environment* – and that these actual present states of affairs can be perceived through the "externally directed" senses.[71]

Thus, on Prinz' account, one can perceive by means of the senses that one is experiencing a particular emotion, such as fear (e.g., one perceives by way of seeing that one is sweating, perceives by means of hearing that one is breathing rapidly, etc.). However, one does not perceive an earache (or other non-emotional interoceptive states) in the same way. For example, one does not perceive by way of hearing the sharp (or sometimes dull) throbbing in one's ear that is often associated with an earache. Instead, one seems to perceive an earache by registering the pathology of one's body. Moreover, Prinz claims that emotions are typically elicited by present actual states of affairs in the environment (e.g., one may become angry when one sees that one's vehicle has been vandalized). However, despite the fact that earaches *can* be elicited by actual present states of affairs in the environment – e.g., low levels of atmospheric pressure – they are more commonly elicited by conditions that originate inside the body (e.g., an infected ear

[71] As I discussed in section 3.3, I do not think that Prinz wants to claim that emotions are elicited *only* by actual present states of affairs – after all, he seems to acknowledge that emotions can represent imagined, remembered, and envisaged states of affairs (see Prinz, 2004, p. 62). Instead, Prinz seems to suggest only that emotions are *typically* elicited by actual present states of affairs.

follicle in the ear canal, excess earwax, trigeminal neuralgia, etc.). Since an earache is not typically elicited by actual present states of affairs in the environment that can be perceived by means of the senses, Prinz should argue that it does not represent a core relational theme, and thus is not an emotion.

However, even if Prinz were to respond in the way I have outlined above, he may still not be able to successfully overcome Barlassina and Newen's objection. This is because an earache can be elicited by actual present states of affairs that *can* be perceived by way of the senses – for example, one's earache may be the result of one sticking a coin in one's ear canal. And so, if one accepts Prinz' approach of distinguishing emotions from non-emotional interoceptive states, one would have to conclude that an earache is an emotion. But, as noted before, this would be absurd.

I think there is a better way for non-cognitive theorists of the emotions to respond to Barlassina and Newan's concerns. It seems to me that Prinz' basic approach is correct – emotions can be distinguished from non-emotional interoceptive states in virtue of their intentionality. Indeed, non-emotional interoceptive states both register and represent certain changes in the body, whereas emotions register certain changes in the body but represent core relational themes. However, it is important to recognize that the very existence of core relational themes depends on what an individual values. I do not take 'value' here to involve a cognitive appraisal; instead, I take it to be akin to caring about something. And so, for example, the death of one's pet is only going to be perceived as a loss if one values the life of one's pet; the graffiti drawing on one's vehicle is only going to be perceived as a demeaning offence if one values one's vehicle; one's upcoming

speech to a group of peers is only going to be perceived as an imminent threat if one values being accepted and/or respected by members of that group, etc.

Core relational themes are relational properties of states of affairs. But these relational properties are not features of states of affairs in and of themselves – instead, they depend on what the individual perceiving them values.[72] This explains why one and the same state of affairs (e.g., the retirement of a famous actor), may cause one person to become sad and not another. Differences in emotional responses can be explained by the fact that two individuals value different things (e.g., one person may value the actor's work while the other may find it to be distasteful). Indeed, a state of affairs has the potential to elicit different emotions in different people depending on what they value.

Although emotions depend on an individual valuing something, non-emotional interoceptive states do not. Imagine, for example, that you are sitting on a commercial airplane with approximately fifty other passengers. If the atmospheric pressure on the airplane were to suddenly drop, it is highly probable that everyone on the airplane (yourself included) would experience an earache (assuming that everyone's bodies are functioning normally). Indeed, if one is in a situation in which the atmospheric pressure suddenly decreases, it does not matter what one values, one will automatically experience an earache. This is because what one values is simply irrelevant to whether or not one's ears inform one of a drop in air pressure.

[72] Core relational themes are *not* response-dependent properties (i.e., properties that depend on being represented a certain way). Imagine, for example, that I value my dog, Socrates, and that he passes away unexpectedly. Socrates' death may represent the property of being a loss to me. However, his death would still be a loss to me regardless of whether it is represented as such – e.g., his death would be a loss to me even if I am unaware that he has passed away. This is because being a loss does *not* depend on being represented as a loss; instead, being a loss depends on an individual representing something as valuable. In this way, loss is the elimination of something that is valued by an individual.

What makes emotions different from things like earaches, fatigue, and nasal congestion is not, as Prinz suggests, that only the former are elicited by stimuli that can be perceived through the external senses, or that only the former are represented by certain kinds of interoceptive states, or that only the former bear on an individual's well-being. It is rather that only the former depend on an individual valuing something. What one values is simply irrelevant to whether or not one's body signals that one has a deficiency of protein or oxygen in one's blood (thus indicating that one is fatigued). But what one values is *not* irrelevant to whether or not a state of affairs is a loss to oneself. Therefore, emotions and non-emotional interoceptive states *can* be distinguished in virtue of their intentionality.

Recognizing that the existence of core relational themes depends on the attitudes of the individual perceiving them is crucial, not only because it allows us to successfully distinguish between emotions and non-emotional interoceptive states, but also because it provides us with a way of accounting for the emotional experiences of human infants and non-human animals. This is significant because prevailing cognitive theories of the emotions struggle to adequately explain the emotional experiences of these sentient beings.

If emotions necessarily involve propositional attitudes (i.e., a belief or a judgment), as cognitive theories of the emotions maintain, then it seems that human infants and non-human animals would be incapable of experiencing emotions (since they do not have the developmental capacity necessary to form beliefs or judgments) (see Deigh, 1994). But this does not seem right. After all, when a human infant giggles (or

when a dog fervently wags his/her tail) it certainly seems as if he/she is experiencing some kind of an emotion (e.g., happiness). Indeed, cognitive theories of the emotions do not fit with our intuitions about the emotional lives of human infants and non-human animals.

My somatic-perceptual theory of the emotions is better equipped to explain how human infants and non-human animals are capable of having emotions. On my account, emotions represent core relational themes by registering certain changes in the body. And, as we have seen, the representation of core relational themes involves an individual valuing something (where 'value' does not involve a cognitive appraisal but is instead akin to caring about something). Human infants and non-human animals are capable of valuing things in this way. For example, both human infants and non-human animals appear to value (at least some) playthings (e.g., toys, dolls, bones, balls, etc.) – this is demonstrated by the fact that both show signs of contentedness when allowed to interact with the playthings and cry or become irritated when the playthings are withheld. Understanding core relational themes in the way I have suggested not only allows us to successfully distinguish between emotions and non-emotional interoceptive states, but also enables us to account for the emotional experiences of human infants and non-human animals.

Prinz is correct to assert that emotions can be distinguished from non-emotional interoceptive states in virtue of their intentionality. While emotions register changes in the body and represent core relational themes, non-emotional interoceptive states both register and represent bodily changes. However, as I have shown, Prinz' account can be

strengthened by highlighting the fact that the very existence of core relational themes depends on the attitudes of the individual perceiving them. Indeed, what makes emotions different than non-emotional interoceptive states is that they necessarily depend on an individual valuing something.

Even if one modifies Prinz' account in the way I suggest above, it nevertheless faces a further problem. In particular, his account does not adequately explain how one's mental representations of particular objects link up to emotions. In the next section (the final section of this chapter), I argue that the relationship between emotions and mental representations of particular objects is best understood in terms of causality.

### 3.6 How Mental Representations of Particular Objects Attach to Emotions

According to Prinz (2004), an emotion necessarily represents its formal object. However, on his account, the representation of an emotion's formal object is always (at least initially) accompanied by a second mental representation – the representation of a particular object. When the representation of an emotion's formal object (e.g., loss) attaches to the representation of a particular object (e.g., the death of one's beloved pet), they combine to make one complex mental representation (e.g., sadness at the death of one's pet). Moreover, Prinz argues, in certain situations it is possible for this complex mental representation to decompose into its constituent parts, and for an emotion to exist as a self-contained mental state.

The problem, however, is that Prinz does not adequately explain how the representation of an emotion's formal object attaches to the representation of a particular

object. He sometimes conceives of this link in terms of co-occurrence (i.e., an emotion co-occurs with the mental representation of a particular object) and sometimes in terms of causation (i.e., a mental representation of a particular object causes an emotion). In what follows, I argue that the relationship between an emotion and the mental representation of a particular object is best understood in terms of causality. I then argue, contrary to Prinz, that an emotion can *never* be disassociated from the representation of a particular object.

In general, I think Prinz' understanding of the basic structure of emotions is accurate. Indeed, it seems to me that what we ordinarily think of as emotions are actually complex mental states made up of two distinct mental representations – the representation of a formal object and the representation of a particular object. And so, for example, if I am angry that my friend insulted me, I have one mental representation that corresponds to the insult (i.e., the particular object of my anger) and another – the emotion – which corresponds to there having been a demeaning offence (i.e., the formal object of my anger). When these mental states are combined, I have one complex mental representation (i.e., anger at the insult) that means the insult is a demeaning offence against me. Saying that my anger is *about* the insult does not mean that my anger represents the insult; instead, it means that the insult is what has caused me to become angry in virtue of it being a demeaning offense against me.

But, how does the representation of an emotion's particular object link up or attach to an emotion? According to Prinz, the answer depends on the type of emotion in question. Prinz maintains that within the class of emotions a distinction can be made between what he refers to as "attitudinal emotions" and "state emotions." This

distinction, he argues, is marked by ordinary language (i.e., the way we speak about emotions in everyday or conventional conversation).

Let us explore both types of emotions. On Prinz' account, attitudinal emotions have the complex structure discussed above – i.e., attitudinal emotions require the combination of both a mental representation of a formal object and a mental representation of a particular object (see Prinz, 2004, pgs. 179-180). An attitudinal emotion and its particular object thus comprise a unified whole. Locutions that signify attitudinal emotions are, for example, the following: Bob was sad that he came in last place in the race, Tom was angry that his car was vandalized, etc. In these cases, Prinz argues that the emotion *cannot* be separated from its particular object. If Bob is sad that he came in last place, his sadness will only disappear after his thoughts about coming in last place subside. Likewise, Tom's anger over the vandalization of his car will only subside once he stops thinking about the vandalization.

According to Prinz, therefore, an attitudinal emotion's particular object is a part of the emotion itself. On his account, "when one is angry about an insult, one's thought about that insult is not merely a cause of one's anger, *it is a part of* one's anger" (Prinz, 2004, p. 181, emphasis mine). And so, the representation of a particular object must somehow bind to the emotion to create a complex mental state. Prinz admits, however, that the nature of this binding is currently unclear. Nevertheless, he postulates that in these cases the emotion co-occurs with the mental representation of a particular object. He states: "In the case of attitudinal emotions… the neural representation of an emotional bodily state fires *at the same time* as the neural realization of the representation of its

particular object" (Prinz, 2004, p. 181, emphasis mine). Thus, for example, if Jane is afraid that the dog will attack, her perception of the bodily changes indicative of fear (e.g., increase in heart rate, pupil dilation, blood vessel constriction, etc.) will occur *at the same time* as her perception of the particular object or state of affairs responsible for causing those bodily changes (i.e., the dog).

However, the idea that an emotion co-occurs with the representation of a particular object seems problematic. First, co-occurrence need not imply parthood. Second, the bodily changes associated with an emotion occur because they *prepare* us to behave in a certain way. For example, the bodily changes associated with fear prepare us to fight or to flee; the bodily changes associated with sadness prepare us to grieve, etc. These bodily changes are predisposed to occur (via evolution) only *after* one perceives that an object or a state of affairs has the property of being an imminent threat, a loss, etc. Thus, if Jane is afraid that the dog will attack, her perception of the bodily changes constitutive of fear occurs only *after* she perceives that the dog poses an imminent threat. In this way, an emotion appears to be *caused by* the representation of a particular object as having a certain property.

On Prinz' account, not all emotions have a complex structure (i.e., not all emotions require both the representation of a formal object and the representation of a particular object). Unlike attitudinal emotions, Prinz maintains that "state emotions" can persist as self-contained mental states – i.e., state emotions do not require the representation of a particular object. Locutions that signify state emotions are, for example, the following: Bob entered into a state of fear after hearing a loud noise behind

him, the insult sent Jane into an angry rage, Fred became sad after learning about the death of a distant family member, etc. In these cases, the mental representation of a particular object or state of affairs *causes* one's perception of a certain pattern of bodily changes (and thus causes an emotion). However, Prinz maintains that the particular object or state of affairs that causes a state emotion (e.g., a loud noise, an insult, a death, etc.) is not a constituent part of the emotion. He states: "…the object that caused the emotion may play little role after the emotion is underway. The emotion has a life of its own" (2004, p. 179). And so, although state emotions are caused by the representation of a particular object or state of affairs, they only essentially involve the mental representation of a formal object.

Prinz suggests that in the case of state emotions the representation of a particular object can be dissociated (i.e., separated) from the emotion it elicits if: 1) one continues to have thoughts about the particular object or state of affairs after the emotion subsides, or 2) an emotional state persists after thoughts of the particular object or state of affairs subside. However, both 1) and 2) appear to be problematic. Imagine, for example, that the recent death of one's beloved pet has caused one to experience sadness. On Prinz' account, having thoughts about the death after one's sadness subsides shows that the representation of the death (i.e., the particular object of one's emotion) has been dissociated from the emotion (i.e., sadness) itself. But perhaps the death simply no longer causes one to be sad? If so, then thinking about the death after one's sadness subsides is simply cognizing about a situation that previously elicited sadness. Although the representation of the particular object *is* disassociated from the emotion, it is only

dissociated *after* one experiences the emotion. The death may no longer elicit sadness, but when one *was* sad about the death (i.e., when one was in the emotional state of sadness), the representation of the death was attached to the sadness (i.e., the representation of the particular object attached to the emotion to make one complex mental representation).

The representation of a particular object or state of affairs seems to be necessary for the occurrence of an emotion (i.e., an emotion *must* be attached to the representation of a particular object or state of affairs). Thus, I do not agree with Prinz' distinction between attitudinal emotions and state emotions – instead, emotions only ever represent core relational themes, but they are always caused by other mental states with which they are (at least initially) closely conjoined, such as, for example, perceptions or imaginings. Although it is currently unclear how the representation of a particular object binds to an emotion, it seems to me that it must hold a particular causal relation – the mental representation of a particular object or state of affairs *causes* an emotion. In this way, emotions are dependent on the representation of particular objects or states of affairs. Once the representation of a particular object or state of affairs causes an emotion, it must somehow bind to it so that the two persist as a single complex mental representation. Without the representation of a particular object, the resulting emotional experience appears to be a mood rather than an emotion. This is because moods, unlike emotions, do

not appear to be directed at particular objects (and thus are non-intentional mental states).[73]

Prinz' claim that an emotion (e.g., sadness) which persists in the absence of a particular object or state of affairs (e.g., the death of a beloved pet) shows that the representation of a particular object or state of affairs can be dissociated from the emotion it elicits is therefore problematic. An emotional state that persists in the absence of a particular object or state of affairs seems to be a mood rather than an emotion. After all, in the absence of a particular object, an emotion lacks intentionality – and as we have seen, emotions represent core relational themes, and are therefore directed toward particular objects.

In this chapter, I have shown how emotions can represent core relational themes by registering certain changes in the body. By doing this, I have shown that non-cognitive theories of the emotions can account for the intentional nature of the emotions just as well as prevailing cognitive theories. Moreover, I have offered a way for non-cognitive theories of the emotions to distinguish emotions and non-emotional interoceptive states without invoking cognition. In this way, I have shown how a somatic-perceptual theory of the emotions can meet the intentionality objection. However, although the intentionality objection is traditionally the most persistent objection against all non-cognitive theories of the emotions, they nevertheless face another serious objection. This objection, which I will refer to as "the physiology objection", is based on the claim that

[73] I think moods are a special class of emotions – differing primarily in terms of their intentionality. While emotions are directed towards particular objects or states of affairs, moods appear to lack particular objects. Instead, moods seem to represent one's overall condition. For example, sadness represents a loss, whereas depression seems to represent something akin to losing oneself. However, a full investigation into the nature of moods is beyond the scope of this dissertation.

different emotions are sometimes associated with the same bodily states. Determining how non-cognitive theories of the emotions can overcome this objection will be the focus of the next chapter.

# CHAPTER FOUR:
# THE PHYSIOLOGY OBJECTION

In the previous chapter, I offered a sketch of how my somatic-perceptual theory of the emotions can account for the intentional nature of emotions without appealing to cognition. By doing this, I provided a way for non-cognitive theories of the emotions to overcome the intentionality objection. However, since it claims that emotions are states that register certain changes in the body, my somatic-perceptual theory is vulnerable to what I will refer to as the *physiology objection* – i.e., the objection that non-cognitive theories of the emotions are unable to distinguish between different emotions that are typically associated with the same bodily states.[74]

Consider, for example, guilt and sadness. Both emotions typically cause one's head to hang low, one's body to become tense, a lump to form in one's throat, and one's stomach to become agitated. Guilt and sadness have many of the same effects on the body, and they *feel* alike (i.e., guilt and sadness are phenomenologically indistinguishable). But, if two emotions register the same bodily states (and thus feel alike), how are we to distinguish between them?

In this chapter, I aim to show that the physiology objection does not pose a problem for my somatic-perceptual theory of the emotions. To begin with, the research

[74] This objection was first made by American physiologist Walter Cannon (1927).

purporting to show that different emotions register identical sets of changes in the body proper is far from conclusive. Moreover, even if some emotions register the same physiological changes in the body proper (i.e., register changes in the autonomic nervous system), they may nevertheless register different changes in the cortex – and thus be neurologically distinct. Recent studies in neuroscience suggest that this is a reasonable hypothesis (see Kassam, et al., 2013). Finally, even if future studies demonstrate that this is not the case – and instead establish that some emotions do in fact register the same changes in *both* the body proper *and* the cortex – my somatic-perceptual theory of the emotions still allows that they can be differentiated via their causal histories.

This chapter divides into three sections. In 4.1, I argue that the famous *adrenaline experiment*, which gave rise to the physiology objection, fails to establish that any two emotions register identical physiological responses in the body proper. In 4.2, I argue that even if one grants that different emotions register identical physiological responses in the body proper, the physiology objection still does not pose a problem for my somatic-perceptual theory of the emotions, since, for all we know, emotions may register different changes in the cortex. In 4.3, I argue that even if future experiments were to establish that some emotions register the same physiological changes in *both* the body proper *and* the cortex, these emotions could nevertheless still be distinguished in virtue of their eliciting stimuli.

### 4.1 The Adrenaline Experiment

Recall that impure cognitive theories of the emotions identify emotions with cognitions and (at least) one other component (e.g., feelings, action tendencies, physiological changes, etc.). One of the most well-known impure cognitive theories of the emotions is the *cognitive labelling theory of the emotions*, developed by American psychologists Stanley Schachter and Jerome E. Singer. The theory is also typically referred to in the psychological literature as the *two-factor theory of emotion* or *the Schachter-Singer theory of emotion* (see Passer et al., 2005, p. 432). On this account, emotions necessarily involve *both* a physiological change *and* a cognitive interpretation (i.e., a judgment) of that change. One perceives an emotionally significant object or state of affairs, which triggers certain physiological changes. These changes are then registered in the cortex and interpreted as resulting from that particular object or state of affairs – the physiological changes are thus labelled (or judged) as having some emotional significance.

Imagine, for example, that upon returning home from work one catches one's spouse being unfaithful. One would likely experience an increase in heart rate, an increase in blood pressure, bruxism, etc., and label those physiological changes as anger resulting from one's unfaithful spouse. According to the cognitive labelling theory, however, if one were to label those physiological changes differently – say, for example, as surprise – one would experience surprise at the actions of one's spouse, as opposed to anger. Moreover, if one failed to assign a label to the physiological changes, one would not experience any emotion, since the physiological changes would not be labeled as

resulting from any emotionally significant object (or state of affairs). And so, on this account, the label that one attaches to certain physiological changes is fundamental in determining what emotion one experiences.

An interesting feature of the cognitive labelling theory is that it implies that labels for emotions cannot be erroneously applied. Indeed, on this view, the physiological changes involved in an emotion may be labeled as resulting from one emotionally significant object (or state of affairs) when in actuality they are the result of some other emotionally significant object (or state of affairs). For example, one may label one's increase in heart rate as jealousy resulting from learning that one's co-worker has been promoted at work. In actuality, however, one's increase in heart rate may be elation resulting from one's recent decision to adopt a puppy. On the cognitive labelling theory, the labelling of one's emotional state is unobjectionable even if one misidentifies the cause of one's arousal.

To show that emotions can arise in just this way, Schachter and Singer conducted what is now considered to be a famed experiment – the adrenaline experiment.[75] In the adrenaline experiment, one hundred and eighty four male participants were injected with a substance that they were told was a vitamin that was being tested for its ability to improve eyesight. In reality, however, half of the participants received an injection of

[75] The adrenaline experiment was inspired by the work of Gregorio Marañon (1924). In order to determine whether subjects' reported emotions varied with physiological changes or with the interpretation of those changes, Marañon injected them with adrenaline without warning of any potential side effects (e.g., an increase in heart rate and respiratory rate). Subjects were then asked to reflect on an emotionally significant life event (e.g., the death of a loved one, an upcoming anniversary or celebration, etc.). Marañon noted that all subjects reported feeling as though they were experiencing the relevant emotion (e.g., sadness, happiness, etc.). Since all subjects were injected with adrenaline (and thus presumably all experienced the same physiological changes), Marañon concluded that the subjects' reported emotions depended on their interpretation of the physiological changes, not on the physiological changes themselves.

adrenaline (the experimental group), and the other half received an injection of saline solution (the control group). Some of the participants (from both the experimental group and the control group) were told to expect heightened autonomic responses such as an increase in heart rate and respiratory rate (typical side effects of adrenaline), while others were not. All participants were then placed into one of two waiting rooms – waiting room 'A' and waiting room 'B'. In waiting room 'A', participants were asked to fill out an insulting questionnaire. Participants watched as an actor became increasingly outraged over the content of the questionnaire and stormed out of the waiting room. In waiting room 'B', participants were not given a questionnaire, but instead watched as an actor engaged in silly antics (e.g., stood on tables, made paper airplanes, played with hula hoops, etc.).

Schachter and Singer noted that the participants who had been informed about the potential side effects of the injection exhibited little emotional response in either condition (i.e., participants who had been informed about the potential side effects of the injection remained emotionally neutral or content throughout the experiment). In comparison, those participants who had not been informed about the potential side effects of the injection exhibited emotional responses similar to those portrayed by the actor who appeared in the waiting room with them. For example, uninformed participants in waiting room 'A' exhibited signs of anger, while uninformed participants in waiting room 'B' exhibited signs of happiness.

Schachter and Singer observed that although the participants who had received the injection of adrenaline all underwent the same physiological changes (e.g., an

increase in heart rate and respiratory rate), they nevertheless labelled those changes as different emotions – i.e., as anger, happiness, or contentment. Additionally, some participants who were told that they were being injected with adrenaline did not label the resulting physiological changes as an emotion *at all*. Schachter and Singer determined that this was the result of the participants being exposed to different background knowledge and stimuli in the waiting rooms. On the basis of this experiment, they concluded that emotions depend not only on physiological changes, but also on one's *interpretation* of those changes.

Schachter and Singer's findings in the adrenaline experiment have led to what I refer to as the *physiology objection* against non-cognitive theories of the emotions. According to this objection, different emotions are often associated with the same bodily states, and so feel alike (i.e., they are phenomenologically indistinguishable). But, if two distinct emotions (e.g., anger and happiness) register the same physiological changes, then bodily states alone cannot be used to distinguish between them. This appears to be problematic for my somatic-perceptual theory of the emotions – since, on my account, emotions are perceptions of changes in the body.

However, there are a number of problems with the adrenaline experiment, some of which have been widely observed. As a result, it fails to establish that different emotions register the same physiological changes. To begin, there are a number of procedural and methodological problems with the experiment. First, Schachter and Singer did not assess participants' mood prior to the injection. This may have had an effect on their results – since, presumably, a participant that is in a good mood prior to the injection

might respond more positively to the actor engaged in silly antics than would a participant that is in a bad mood prior to the injection. Second, the sample used in the experiment was not representative – it included one hundred and eighty four male college students taking psychology classes at the University of Minnesota. Third, Schachter and Singer do not precisely define their key concepts; specifically, 'emotion', 'cognition', and 'arousal'. This is problematic because within the psychological literature, there is no single agreed upon definition of 'cognition' or 'emotion' (see Reisenzein, 1983; Leeper, 1965; Leventhal, 1974; Mandler, 1975; Strongman, 1973). It is also important to note that there have been several unsuccessful attempts to replicate Schachter and Singer's results in this experiment (see Marshall and Zimbardo, 1979; Maslach, 1979).

More significantly, for present purposes, the adrenaline experiment fails to establish that different emotions register the same physiological changes. In analyzing the results of the experiment, Schachter and Singer appear to make the following two assumptions: 1) participants that display distinct emotional behavior are experiencing different emotions, and 2) participants that display different emotional behavior are in the same physiological state. Both assumptions can be challenged.

First, consider the assumption that participants who display distinct emotional behavior are experiencing different emotions. Schachter and Singer measured participants' emotional states by observing their behavior through a two-way mirror (i.e., the behavioral measure) and then asking them to report on their emotional state (i.e., the self-report measure). The behavioral measure seems to support Schachter and Singer's hypothesis that different contexts can cause the same physiological state to be interpreted

in different ways. For example, participants in the experimental condition (i.e., participants that were given an injection of adrenaline) that were placed in waiting room 'A' (i.e., the waiting room with the insulting questionnaire) behaved in ways indicative of anger, whereas participants in the experimental condition that were placed in waiting room 'B' (i.e., the waiting room with the actor engaged in silly antics) behaved in ways indicative of happiness.

However, there are problems with the self-report measure. First, there is no way to verify whether a participant's verbal report about their experiences accurately reflects the contents of their experience. Indeed, when asked to report on their emotions, (at least some) participants in the experimental group that were placed in *both* waiting room 'A' *and* waiting room 'B' reported that they were happy. This is at odds not only with the behavior exhibited by participants in the experimental group that were placed in waiting room 'A' (i.e., the waiting room with the insulting questionnaire), but also with Schachter and Singer's claim that emotions do not depend on physiology alone. Second, introspective data cannot be replicated, since each subjective experience is arguably a singular, and thus non-repeatable, event. This makes introspection problematic as a method for producing generalizations across individuals, or even within individuals across time.

Now, consider the assumption that participants that display different emotional behavior are in the same physiological state. This assumption seems to be grounded in the fact that all participants in the experimental group were given the same drug; namely, adrenaline. However, it is not safe to assume that adrenaline will produce the same

physiological responses throughout the entire experiment. The adrenaline may serve as a catalyst to increase the chances of an emotional response; but once a participant exhibits an emotional response, there is no reason to assume that their physiology will stay the same. After all, participants are placed in very different environments, and it is likely that one's environment has some sort of an effect on one's physiological responses – for example, the insulting questionnaire may cause one's respiratory rate to increase, whereas the actor engaged in silly antics may cause one's respiratory rate to decrease.

In their experiment, Schachter and Singer fail to establish that different emotions register the same physiological changes. However, even if one were to grant that the experiment does, in fact, demonstrate that different emotions register the same physiological changes, the physiology objection would *still not* pose a problem for my account. This is because in the experiment, Schachter and Singer were solely concerned with physiological changes occurring in the body proper (i.e., changes occurring solely within the autonomic nervous system). But, for all we know, emotions may register distinct changes in the cortex. The idea that emotions are neurologically distinct will be the focus of the next section.

### 4.2 Emotions as Neurologically Distinct

In carrying out the adrenaline experiment, Schachter and Singer focused exclusively on the physiological changes occurring in participants' autonomic nervous systems (e.g., changes in heart rate, respiration rate, perspiration, etc.). However, they failed to examine whether there were any physiological changes occurring in the participants' central

nervous systems (i.e., in the participants' brains). This is problematic because, for all we know, emotions may register distinct physiological changes in the brain. If this is the case, as recent studies in neuroscience suggest, then the physiology objection does not pose a problem for my somatic-perceptual theory of the emotions (or for any non-cognitive theory of the emotions).

Since the 1930s, emotion researchers have generally agreed that structures in the limbic system (specifically, the amygdala and cingulate cortex), are critically involved in emotion (see Phelps and LaBar, 2006, p. 568). Until recently, however, very little was known about the processes that give rise to emotions, or how emotions are represented in the cortex. Developments in brain imaging techniques (e.g., positron emission tomography (PET), magnetic resonance imaging (MRI), functional magnetic resonance imaging (fMRI), etc.), have enabled emotion researchers to gain new insight into emotion. For example, researchers are now able to explore the underlying mechanisms involved in emotion, study whether emotions are biologically basic or socially constructed, and determine whether emotions have identifiable neurological signatures.

A recent study conducted by emotion researchers at Carnegie Mellon University is, for present purposes, particularly important. In an unprecedented study, Karim Kassam (2013) and his team of scientists used fMRI scans to explore whether individual emotions can be identified solely by distinct patterns of neural activity. Kassam recruited ten actors (eight adult females and two adult males) from the Carnegie Mellon University Drama Community to participate in the study. Prior to the fMRI scan, participants were asked to write one scenario (derived from a past, a present, or an imagined experience)

for each of eighteen different emotion words (e.g., anger, disgust, envy, fear, happiness, lust, pride, sadness, shame, etc.). Participants' task during the fMRI scan was to actively experience various emotions on cue using the scenarios they had previously constructed (see Kassam et. al., 2013).

Kassam and his team used actors in this study because, unlike non-actors, actors are typically skilled at cycling through various emotional states on cue (see Kassam, 2013). However, in recruiting actors as participants, the research team was concerned that the results of the study would reveal something about the actors' emotions specifically, rather than about emotion more generally. In order to mitigate this concern, Kassam and his team conducted a second fMRI scan whereby participants were presented with photos that they had never seen before. These photos contained neutral content (e.g., scenes of bridges, lakes, apartment buildings, etc.) and content intended to invoke disgust (e.g., rotting food, dirty toilets, etc.). The results of this scan showed substantial similarities in the network of cortical activity in both the self-induced (i.e., acted) disgust trial and the picture-viewing disgust trial. This suggests that the cortical activity observed in the first scan was not dependent on participants' acting (i.e., on participants trying to induce an emotion) (see Kassam et al., 2013).

The results of this study indicate that individual emotions register distinct changes in the cortex, and so can be identified by brain activity alone. Moreover, results suggest that individual emotions have a consistent network of brain activity across participants. This means that at a neuronal level, participants' emotion states (e.g., anger) look very similar. Indeed, despite variations in participants' psychology (i.e., in participants'

personality, sexuality, etc.), they nevertheless seem to encode emotions in a remarkably similar way. Based on these findings, Kassam and his team concluded that emotions can be identified on the basis of their unique neurological signatures, and that "these signatures are reliably activated across episodes and across individuals" (Kassam et al., 2013).

The results of this study are compatible with my somatic-perceptual theory of the emotions. On my account, emotions are states in our somatosensory cortices that register changes in our bodies – but, the changes that emotions register need not be limited to the body proper. Instead, it seems likely that emotions register physiological changes in *both* the body proper *and* the cortex. As results from the study indicate, an emotion is not limited to a specific brain region (e.g., the amygdala) as was once thought; instead, an emotion involves characteristic patterns of activity throughout the cortex (e.g., activity in the frontal, parietal, temporal, and occipital lobes) (see Kassam et al., 2013). For all we know, this cortical activity could be unique to every emotion – i.e., characteristic patterns of activity occur in various regions of the cortex, and these changes are registered by an emotion in the somatosensory cortex.

On my somatic-perceptual theory of the emotions, one imagines, or perceives by way of the senses (e.g., vision, audition, olfaction, etc.), an emotionally significant object or event, which subsequently causes a certain pattern of change in the body (i.e., a certain pattern of change in *both* the body proper *and* the cortex). Emotions register these physiological changes in the somatosensory cortex (located in the parietal lobe) and an emotion just is the perception of these changes. Hence, on my account, there are two

kinds of perception. The first kind involves perception in the more ordinary and familiar sense of the word – i.e., perception via the senses (e.g., vision, audition, olfaction, etc.). For example, one perceives by way of smell that the sandwich is rotten. The second kind of perception does not involve the senses, but is instead, a type of inner perception.

Although this inner perception does not involve the senses, it nevertheless seems reasonable to classify it as a form of perception. This is because, like ordinary perception, this inner perception is: 1) domain-specific (i.e., emotions are interoceptive states that register certain patterns of changes in the body), 2) representational (i.e., emotions represent core relational themes), and 3) at least typically consciously experienced (i.e., emotions have a phenomenal quality – although, like ordinary perception, this inner perception can also occur unconsciously, and thus, be unfelt).[76]

One possible worry about this view is that by allowing that emotions perceive *both* changes in the body proper *and* the cortex, this type of inner perception diverges too much from ordinary perception – since, for example, it seems that one is not consciously aware of changes in one's own cortex in the same way that one is (or least can be) aware of changes in one's heart rate. This divergence, therefore, removes some of the initial appeal of James' somatic-feeling theory of the emotions (from which my somatic-perceptual theory and all other somatic theories of the emotions are derived). However, this seems to be a relatively small price to pay in the pursuit of a complete understanding of the emotions – one that is both scientifically respectable and philosophically persuasive.

[76] I will explain and defend this claim in the following chapter.

To this point, I have argued that on my somatic-perceptual theory of the emotions, one can distinguish between emotions in virtue of the physiological changes they register. Even if emotions register the same changes in the body proper, it is possible that they register distinct changes in the cortex. Indeed, as the study conducted by Kassam (2013) and his team suggests, this may very well be the case. If this is right, then the physiology objection does not pose a problem for my somatic-perceptual theory.

Moreover, even if future studies were to establish that some emotions *do* in fact register the exact same changes in *both* the body proper *and* the cortex, the physiology objection would *still* not pose a problem for my account. This is because my somatic-perceptual theory of the emotions allows that emotions can be differentiated, not only via the different bodily changes they register, but also via their causal histories. The idea that emotions can be distinguished in this way will be the focus of the final section of this chapter.

### 4.3 Distinguishing Emotions via their Different Causal Histories

In the preceding sections of this chapter, I argued that the adrenaline experiment, from which the physiology objection against non-cognitive theories of the emotions originates, fails to establish that different emotions register identical changes in the body. I then argued that even if one were to concede that the adrenaline experiment *does* establish that different emotions register identical changes in the body proper, recent evidence from neuroscience suggests that emotions register different changes in the cortex, and are thus neurologically distinct. In what follows, I argue that even if future studies establish that

this is incorrect, and that some emotions do, in fact, register identical physiological changes in both the body proper and the cortex, my somatic-perceptual theory of the emotions still allows for a further way to distinguish them; namely, via their causal histories.

There are a number of emotions that have similar (or identical) phenomenology (i.e., there are a number of emotions that have the same bodily feelings associated with them). For example, both guilt and disappointment feel like sadness, anger feels like indignation, pride feels like joy, fear feels like exhilaration, and jealousy feels like a blend of disgust, anger, and fear. Emotions that are phenomenologically indistinguishable (and thus feel alike) register the same physiological changes in the body proper. If future studies in neuroscience establish that phenomenologically indistinguishable emotions also register the same physiological changes in the cortex, then how are we to distinguish between them?

Even if future research should determine that different emotions sometimes involve identical physiological changes in both the body proper and the cortex – something that has not yet been shown – this would still not undermine my view, since it would still be possible to distinguish emotions via their different causal histories. That is, emotions with identical physiological states could still be differentiated by determining their eliciting conditions. For example, even if it were to be established that sadness and guilt register identical physiological changes, they nevertheless will always have distinct eliciting stimuli. This is because sadness, at least on my theory, is an emotional state caused by a perceived loss (e.g., the death of one's beloved friend), whereas guilt is an

emotional state caused by one's perceived transgression of a norm (e.g., stealing something from a department store). And so, even if sadness and guilt shared identical physiological states, they would nevertheless be caused by different eliciting stimuli. This allows that by examining the causal history of an emotion (i.e., by determining the cause of an emotion), one could distinguish between emotions that registered the same bodily state.

In this connection, it is perhaps worth nothing that distinguishing between two mental states that share similar (or identical) physiology via their causal histories is not without precedent in philosophy. For example, in the philosophy of perception, perceptual experiences are commonly characterized as being either veridical perceptions (i.e., genuine perceptions), or hallucinations.[77] Those who hold a disjunctive view of perception maintain that the conscious perceptual experiences involved in veridical perceptions, illusions, and hallucinations are all of a different nature. However, there are many different varieties of disjunctivism – not all disjunctivists are in agreement over how best to understand the nature and relationship between the various types of perceptual experiences.[78] Some disjunctivists allow for the possibility that while hallucinations and veridical perceptions can have both the same physiology and phenomenal character (i.e., hallucinations and veridical perceptions can feel alike), they

[77] In the philosophy of perception, perceptual experiences are commonly divided into the following three categories: i) veridical perceptions, ii) hallucinations, and iii) illusions. However, for present purposes, I omit illusion from this discussion because it is the disjunctivist's way of distinguishing between veridical perception and hallucination that is of primary interest here.

[78] In philosophy, the disjunctivist theory of perception holds that the objects of veridical perception are mind-independent. The theory also accepts that hallucination and illusion are possible – but maintains that veridical perception is of a distinct nature than both hallucination and illusion. For further discussion see Soteriou, 2014 <http://plato.stanford.edu/entries/perception-disjunctive/>.

nevertheless have distinct mental properties (see Soteriou, 2014). In particular, while veridical perceptions are constitutively dependent on mind-independent objects, hallucinations are not. This means that unlike hallucinations, veridical perceptions have intentional content (i.e., veridical perceptions are about or are directed toward something) that actually exists in the world.

Since hallucinations and veridical perceptions are often both phenomenologically and physiologically identical, one way to establish whether one is experiencing a hallucination or a veridical perception is by determining the cause of the perceptual experience. That is, it can be established by determining whether one's perceptual experience is caused by a mind-independent object (thus establishing that one is experiencing a veridical perception) or a mind-dependent object (thus establishing that one is experiencing a hallucination). For example, imagine that one sees a snake. In order to determine whether one's perceptual experience is an instance of veridical perception or hallucination, one can reflect upon whether there is an appropriate causal connection between one's surrounding environment and one's seeing a snake (e.g., determining whether one is outside). Even if one cannot determine for oneself whether one is experiencing a veridical perception or a hallucination, an appropriately placed external observer could, and hence there is a real difference between the two states.

In a similar manner, different emotions could be distinguished by their different causal histories in much the way that (on some views) veridical perceptions and hallucinations can be distinguished by their different causal histories. This would allow different emotions to be distinguished without invoking cognition, even if it should turn

out that it is possible for different emotions to register identical physiological changes in both the body proper and the cortex.

### 4.4 Concluding Remarks

In this chapter, I have shown that the physiology objection against non-cognitive theories of the emotions is not as strong as it may initially appear. I have accomplished this by arguing that the physiology objection does not pose a problem for my somatic-perceptual theory of the emotions. First, I have argued that Schachter and Singer's adrenaline experiment fails to establish that emotions register the same physiological changes in the body. Second, I have argued that even if emotions register the same changes in the body proper, recent evidence from neuroscience suggests that emotions register different changes in the cortex – and thus, that emotions are neurologically distinct. Third, I have argued that even if future studies demonstrate that some emotions do, in fact, register the same changes in *both* the body proper *and* the cortex, my somatic-perceptual theory of the emotions still allows that they can be distinguished in virtue of their causal histories. In this way, I have shown that a non-cognitive theory of the emotions can successfully overcome the physiology objection.

However, my somatic-perceptual theory of the emotions still faces a further potential problem. On my account, emotions are a form of perception. Since not all perceptions are felt, my view seems to imply that not all emotions are felt. In fact, I accept and embrace the idea that emotions can be unconscious. However, it might seem simply obvious that any emotion *must* be consciously felt. For this reason, the notion of

unconscious emotions is heavily debated in the philosophical literature. Defending and exploring the idea that emotions can be unconscious will therefore be the focus of the next, and final, chapter.

# CHAPTER FIVE:
# EMOTIONAL CONSCIOUSNESS

In ordinary discourse, emotions are paradigm cases of feelings (i.e., phenomenally conscious mental states). Indeed, the terms "emotion" and "feeling" are often used interchangeably (e.g., "I feel sad"). It sounds strange, and perhaps even contradictory, to say: "I am sad, but I do not feel sad," or "I am elated, but fortunately it is unconscious." If one is currently experiencing an emotion, it seems, at least intuitively, that one must be *aware* that one is in an emotional state. After all, unconscious mental states are those mental states that are not currently being accessed (i.e., they are below the subjective threshold of awareness). Do unconscious emotions exist? If so, how are we to make sense of them?

On the somatic-perceptual theory of the emotions that I have been defending, emotions are perceptions of changes in the body that represent core relational themes. Someone might object to my account on the basis that, due to my classification of emotions as a form of perception, I seem to be committed to the counter-intuitive view that emotions can be unconscious. This is because, as we saw in chapter one, 'perception' is not equivalent to 'feeling' – feeling requires conscious awareness whereas

perception does not.[79] This suggests, therefore, that on my account it should be possible to have unconscious (i.e., unfelt) emotions.

However, I take it to be an advantage of my account that it suggests that emotions can be unconscious. My account of the emotions is consistent with anecdotal and experimental evidence that demonstrates the existence of unconscious emotions – this is significant because cognitive theories of the emotions (and some non-cognitive theories of the emotions) have thus far been unable to account for unconscious emotions (e.g., Greenspan, 1988; LeDoux, 1994). In this chapter, I defend the notion of unconscious emotions and show how my somatic-perceptual theory of the emotions provides a framework for thinking about poorly understood psychological conditions (e.g., alexithymia).

This chapter divides into four sections. In 5.1, I argue, contrary to the majority of emotion researchers, that emotions can be unconscious. In 5.2, I introduce a relatively unfamiliar psychological condition known as 'alexithymia' – whereby an individual experiences the physiological symptoms of an emotion (e.g., increased heart rate and blood pressure) but is nevertheless unable to verbalize the associated emotional state (e.g., anger, fear, etc.). I argue that there are various degrees of alexithymia and that the nature of the condition is best understood on the assumption that emotions can be unconscious. In 5.3, I highlight the emotional processing deficit involved in mild alexithymia and argue that it is best understood through an analogy with blindsight. In

[79] I take it as uncontroversial that perceptions can be unconscious (see Cheesman and Merikle, 1986).

5.4, I highlight the emotional processing deficit involved in severe alexithymia and argue that it is best understood through an analogy with apperceptive visual agnosia.

## 5.1 Unconscious Emotions

The primary role of perception is to provide information about one's external environment via the sense organs. However, perception can occur both consciously and unconsciously. Indeed, perceptual systems (e.g., visual system, auditory system, olfactory system, etc.) regularly allow for unconscious perception. As we saw in chapter one, subliminal perception (whereby a stimuli is presented below the subjective threshold of awareness but nevertheless influences one's thoughts, feelings, and/or behavior) is perhaps the most common example of unconscious perception. And so, if emotions are perceptions of changes in the body, as I argue they are, it seems to follow that they can occur both above and below the subjective threshold of awareness (i.e., emotions can occur *both* consciously *and* unconsciously).

However, the notion of unconscious emotions (and unconscious mental states more generally) have never been well-received among emotion theorists in philosophy and psychology. Even Freud – who is often referred to as "the master of the unconscious" (see for example, Prinz, 2004, p. 199) – did not think that emotions could be unconscious. He writes:

> It is surely the essence of emotion that we should be aware of it, i.e., that it should be known to consciousness… [But] it may happen that an affective or emotional impulse is perceived, but misconstrued. Owing to the repression of its proper representative it has been forced to become connected with another idea (Freud, 1915/1984, p.179).

On Freud's account, although one is often unconscious of the *cause* of an emotion (i.e., unconscious of the emotion eliciting stimuli), the emotion *itself* must nevertheless be conscious (1915/1984, p. 179). Imagine, for example, that Jane begins to weep when she sees man and a child sitting at a picnic table in the park being merry and playing a game of cards. According to Freud, while Jane *must* be aware that she is sad, she may nevertheless be unaware that her sadness is, for example, associated with the long-standing estrangement of her father. In this way, Freud maintains that emotions themselves can *never* be unconscious.

Skepticism about unconscious emotions has continued to this day – a large number of contemporary emotion researchers insist that emotions cannot be *completely* unconscious. For example, American psychologist Gerald Clore (1994) argues that emotions must be felt, and that feelings are, by definition, conscious. According to Clore, the primary function of an emotion is to alert one to (i.e., bring to the center of one's attention) emotion eliciting stimuli (i.e., particular objects or states of affairs in one's environment). Emotions must be conscious, on this account, because if they were unconscious they would be unable to serve their primary function (i.e., emotions would be unable to alert one to emotion eliciting stimuli).

Clore maintains that there are two necessary components of an emotion episode: a cognitive component (i.e., an appraisal or a judgment) and a feeling component (i.e., a phenomenally conscious mental state). He argues that while it is possible for the former (i.e., the cognitive component) to be unconscious, it is not possible for the latter (i.e., the feeling component) to be unconscious (see 1994, p. 285). By taking the position that

emotions must be conscious, Clore does not mean to imply that one always knows the cause of one's emotions; rather, like Freud, he seems to suggest that one can be unaware of the cause of one's emotions, but not of the emotions themselves (see 1994, p. 287).

However, Clore's claim that the primary function of emotions is to alert one to emotion eliciting stimuli appears to be somewhat problematic. As we have seen, he makes the following two claims: i) emotions must be conscious because they cannot serve their primary function if they are unconscious, and ii) it is possible for one to repress knowledge about the cause of one's emotions (i.e., it is possible for one to repress knowledge about emotion eliciting stimuli). But, if the function of emotions is to alert one to emotion eliciting stimuli, then emotions fail to serve their primary function when emotion eliciting stimuli are repressed – this is because emotions cannot alert one to emotion eliciting stimuli if the stimulus has been repressed (and is thus unconscious). For example, anger cannot alert one to one's memory of one's father being physically abusive during one's childhood if one has repressed that particular memory. And so, it seems that Clore is committed to saying that it is possible for emotions to occur without serving their primary function (or at least without serving it fully, since it is likely that one is aware that one is angry at *something*).

American neuroscientist Joseph LeDoux (1994) also argues against the possibility of unconscious emotions – though he allows for unconscious emotional processing. On his view, the processes that give rise to an emotional response (i.e., the processes that occur in primitive brain structures such as the limbic system) are unconscious, but the emotion itself must be conscious. Even those who defend unconscious emotions

sometimes imply that *something* must be conscious during an emotion episode. For example, Patricia Greenspan (1988) argues that emotions are evaluative judgments coupled with feelings of pleasure or pain. On her view, evaluative judgments can either be conscious or unconscious, but one always feels the associated pleasure or pain (i.e., the associated pleasure or pain must always be conscious). Although Greenspan often defends unconscious emotions, it nevertheless seems that she is only willing to say that emotions can only be *partially* unconscious (i.e., the cognitive component of an emotion can be unconscious, but not the feeling component).

However, despite the hesitation among many emotion theorists, there *is* evidence which suggests that unconscious emotions exist. First, there is anecdotal evidence for unconscious emotions. Imagine, for example, that Bob is sitting in a lecture hall preparing to write an exam when he hears gunfire. Bob immediately takes cover under his desk and listens intensely to the sound of footsteps coming from behind him. At the same time that Bob is attending to the footsteps, his body enters into a state of fear – e.g., his heart rate quickens, his breathing becomes strained, he begins to perspire, etc. However, Bob may not consciously experience the fear because his attention is consumed elsewhere (e.g., on the gunman). It is only after Bob hears the gunman exit the room that he becomes aware of the physiological changes that have taken hold – e.g., his heart is racing, his breathing is strained, and his entire body is trembling. Bob was afraid from the moment he heard gunfire, but he did not realize it (i.e., Bob's fear was unconscious).

Second, there is experimental evidence for unconscious emotions. For example, in an experiment by Fischman and Foltin (1992), cocaine addicts were attached to two

intravenous lines – one intravenous line contained saline solution and the other contained varying amounts of cocaine. Subjects were *not* told which intravenous line contained the cocaine. The subjects were asked to press a button that released infusions from either line (e.g., subjects were asked to press the button either for line "A" or line "B"). Subjects reported not having any change in emotion when given very low doses of cocaine as opposed to saline (i.e., subjects maintained that their emotional state remained the same both before and after they were given very low doses of cocaine as opposed to saline). Interestingly, however, subjects pressed the button for the cocaine line twice as often as they pressed the button for saline. This suggests that introspective reports (i.e., subjective reports) of one's emotional states do not always coincide with the typical behaviors associated with particular emotions. One would assume that if the subject's emotional state was not affected by the low doses of cocaine that he/she would have pressed the button for saline just as much as he/she pressed the button for cocaine. This suggests, therefore, that the subjects were not aware of their true emotional states.

In another experiment, Winkielman et al., (2005) subliminally presented subjects with photographs of facial expressions (e.g., neutral, happy, and angry). Although these photographs were presented too quickly to be consciously experienced, they nevertheless had an effect on subjects' behavior. After being presented with the photographs, subjects were asked to pour themselves a drink of a fruity beverage, and to report on their feelings. Subjects who had seen the angry face and the happy face all reported being in the same emotional state – however, subjects who saw the angry face poured less of the beverage, and drank less of it, than did those who saw the happy face. Winkielman et al.

note that emotions are known to impact behavior in this way and maintain that the results of this experiment suggest that emotions had been unconsciously induced.

To this point, I have shown that despite the controversy among emotion theorists, there is good evidence in support of the existence of unconscious emotions. On my somatic-perceptual theory, emotions are perceptions of changes in the body – but perceptions can be either conscious or unconscious. This suggests that on my account it should be possible for emotions to be unconscious. Therefore, unlike cognitive theories of the emotions which maintain that emotions cannot be *completely* unconscious, my somatic-perceptual theory is able to account for unconscious emotions – and thus also provides a framework for thinking about poorly understood psychological conditions, such as alexithymia, which is best understood only on the assumption that emotions can be unconscious.

Despite experiencing the physiological symptoms of an emotion (e.g., increased blood pressure, heart palpitations, sweaty palms, etc.), individuals characterized as alexithymic are nevertheless unable to verbalize the associated sentiment (i.e., individuals characterized as alexithymic are unable to verbalize what emotional state they are experiencing – e.g., anger). In the next section, I consider why the condition is so poorly understood and argue that it is best to recognize it as having a dimensional (i.e., graded) nature.

## 5.2 Alexithymia[80]

When asked to report on our emotions, we typically report that we *feel* a certain way (e.g., angry, elated, anxious, etc.). Not only are we able to put into words what emotion we are experiencing (e.g., "I am angry"), but we are also able to describe what it feels like for us to be in that emotional state, and to recognize these states in others. However, not everyone has the ability to identify and describe their emotions. In particular, individuals who are often characterized as alexithymic are seemingly unaware of their emotional feelings (i.e., their phenomenally conscious mental states). When asked to report on their emotional state, alexithymics will either claim that they do not know how they feel, or they will describe their bodily sensations (e.g., "my heart is palpitating and my face feels flushed"). Despite experiencing prototypical autonomous responses to emotion-evoking stimuli (e.g., perspiration, heart palpitations, pupil dilation, etc.), alexithymics are unable to identify and describe precisely what emotion they are experiencing (e.g., sadness, happiness, surprise, etc.). How are we to understand alexithymia? Are alexithymics incapable of experiencing emotion? Do they simply not recognize their emotional feelings? Perhaps they experience emotions unconsciously (i.e., in the absence of awareness)?

Despite decades of research, there is currently no comprehensive account of alexithymia available. There are a number of reasons for this, including: (1) disagreement

[80] The term "alexithymia" was coined by psychiatrist Peter Sifneos in 1972, and means "no words for emotions" (from the Greek *alexis* (no words), *thymos* (emotion)).

over the fundamental deficit of the condition[81]; (2) disagreement over the proper physiological evidence for the condition; (3) disagreement over how best to conceptualize the nature of the condition[82]; (4) a lack of longitudinal studies (i.e., studies that monitor alexithymic traits across an individual's life span); and (5) a poor understanding of the effect(s) (if any) that modified psychotherapy has on alexithymic characteristics. Here, I focus on (3) – namely, on how best to conceptualize the nature of alexithymia. One promising way of seeking to understand the condition, in the absence of a sound understanding of its neurobiology (see Taylor, et al., 1997, pgs. 93-113), is by analogy with another condition – a condition with a similar deficit to the one involved in alexithymia, but that is nevertheless more adequately understood.

[81] While some argue that the fundamental deficit in alexithymia is – as the name of the condition suggests – the capacity for symbolic representation of emotion – specifically in verbal behavior (i.e., the ability to put feelings into words), fantasy, and dreams (see Nemiah, et al., 1976), others argue that the fundamental deficit is in the conscious experience of emotion (see Lane, et al., 1997, Taylor, et al., 1997). While I agree that an inability to symbolically represent emotion and an inability to consciously experience emotion are deficits associated with alexithymia, I am nevertheless inclined to believe that the *fundamental* deficit associated with the condition is one of emotional information processing. This is because a deficit in emotional information processing can result in both an inability to symbolize and consciously experience emotion. Indeed, as Lane, et al. (1996) demonstrate, a deficit in emotional information processing can include (but is not limited to) a deficit in the symbolic representation of emotion. Moreover, Lane, et al. (1990) show that a deficit in emotion information processing can result in an inability to consciously experience emotion (see also Lane and Schwartz, 1987; Lane, et al., 1997, p. 836-837). My claim that alexithymia involves a fundamental deficit in the processing of emotional information will be more fully developed as the chapter proceeds.

[82] According to Lane, et al. (1997), alexithymia is best understood as "a disruption of transmission of interoceptive emotional information to the ACC [anterior cingulate cortex]" (p. 840). On the other hand, Taylor, et al., (1997) argue that alexithymia is associated with a variation in brain organization – a variation which results in a dysfunction of affect (i.e., emotion) regulation (see p. 91-113). They note, however, that it is "unclear…whether this variation represents a dysfunction of the right hemisphere, an interhemispheric transfer deficit, an inhibition of the right hemisphere by a highly activated left hemisphere, or merely a preferred hemispheric mode" (p. 113). Nevertheless, given the increasing body of evidence suggesting that emotional processing necessarily involves both the left and the right hemisphere, Taylor et al. (1997) predict that "future research will demonstrate that alexithymia reflects a deficiency in how the two hemispheres coordinate their respective operations" (p. 113). As these two accounts demonstrate, further research on the neurobiology of alexithymia is necessary.

As this chapter develops, I explore the merits of two analogies others have used to try to understand the nature of alexithymia. In 5.3, I consider an analogy that characterizes alexithymia as an analogue of associative visual object agnosia (i.e., a visual condition whereby an individual is unable to recognize visually presented objects) (see Prinz, 2004, p. 217). In 5.4, I consider an analogy that characterizes alexithymia as the emotional equivalent of blindsight (i.e., a visual condition whereby an individual who is blind in a certain area of their visual field is nevertheless able to accurately (or, at least above chance levels) respond to stimuli presented in their blind spot) (see Lane et al., 1997). Although the analogy with blindsight more adequately portrays the processing deficit involved in what I will refer to as *mild alexithymia* (i.e., lower degrees of alexithymia), neither the analogy with blindsight nor the analogy with associative visual object agnosia is able to account for the processing deficit involved in what I will refer to as *severe alexithymia* (i.e., higher degrees of alexithymia).

Let us first explore some of the symptoms associated with alexithymia. Alexithymia was first observed in individuals with classic psychosomatic disorders (i.e., psychologically caused physical disorders – e.g., high blood pressure, peptic ulcers, asthma, migraine and tension headaches, etc.) (Nemiah & Sifneos, 1970), and has since been observed in individuals with a variety of other disorders – including, for example, post-traumatic stress disorder, substance abuse and dependence (Krystal, 1988), eating disorders (Taylor et al., 1997), panic disorder (Parker, 1993), and Asperger's syndrome (Fitzgerald, 2004).

More recently, the 20-item Toronto Alexithymia Scale (TAS-20) (one of the most valid methods of measuring alexithymia) has demonstrated a replicable four-factor structure of the condition: difficulty identifying feelings and/or discriminating between feelings and the bodily sensations of emotional arousal (Factor 1); difficulty describing feelings to other people (Factor 2); impaired imagination (Factor 3); and an externally bound thinking style (i.e., a hyper-focus on external events and objective facts) (Factor 4).[83]

At first glance, *some* alexithymics do not appear to exhibit behavior consistent with Factor 1 – i.e., they do not appear to have difficulty identifying their feelings, since they commonly complain of depression and/or anxiety, and exhibit outbursts of sadness, anger, or rage (see Sifneos, 1967). However, as Taylor et al. (1997) note: "intensive questioning…reveals that [these individuals] know very little about their own feelings, and, in most instances, are unable to link them with memories, fantasies, higher level affects, or specific situations" (p. 29). As we will see, although all alexithymics exhibit behavior associated with these four factors, it seems likely that they are more evident in individuals with severe alexithymia than in individuals with mild alexithymia.

A number of additional features are often associated with alexithymia, including: stiff posture (Frawley & Smith, 2001), a lack of emotional facial expressions (McDonald

[83] See Taylor et al. (1997), p. 29 and Larsen et al. (2003), p. 534.

Bermond (1997) and Vorst & Bermond (2001) maintain that alexithymia is associated with a fifth feature; namely, a reduced ability to experience emotional feelings. I do not include this as a distinct feature in my list here because it seems redundant to do so. This is because two of the most valid methods for measuring alexithymia – the Toronto Structured Interview for Alexithymia (TSIA) and the TAS-20 already assess awareness of emotional feelings by factoring into their analysis the first two features in my list above: 1) difficulty identifying feelings, and 2) difficulty describing feelings to others (for further discussion see Bagby, et al. (2009)).

& Prkachin, 1990), a tendency to be interpersonally distant (Kauhanen et al., 1996), an impaired capacity for empathy and self-insight (Haviland et el., 2004), an impaired ability to identify the emotions of others – including their facial expressions (Lane et al., 1996, 2000; Jessimer & Markham, 1997; Parker et al., 1993 (b)), and somatic preoccupation (Krystal, 1988). Despite being associated with the condition, however, Taylor et al. (1997) note that these features "are not part of the theoretical core of the [alexithymia] construct" (p. 29). These features, therefore, can be thought of as "the secondary features of alexithymia".[84]

Currently, there is an ongoing debate in the literature as to whether there are subtypes of alexithymia. Drawing on recent knowledge from neurobiology, Bermond (1997) distinguishes between two forms of alexithymia: Type I alexithymia and Type II alexithymia. Type I alexithymia is characterized by the absence of conscious awareness of emotional arousal, and consequently, by the absence of the cognitive judgment accompanying the emotion. On the other hand, Type II alexithymia is characterized by a selective deficit in both conscious awareness of emotional arousal and the cognitive judgment accompanying the emotion. This distinction, however, is controversial. For example, Bagby et al. (2009) argue that "alexithymia cannot be reliably decomposed into two variants," and that "the existence of Type I and Type II alexithymia is not a veridical representation of the alexithymia construct" (p. 418). Moreover, Mattila et al. (2010)

[84] It is important to remember that alexithymia is incompletely understood – and so, the way that the condition is currently characterized (i.e., the four features that are currently thought to make up the core of the alexithymia construct) might change with findings from future research.

argue that alexithymia is best understood as a dimensional (i.e., graded) construct, as opposed to a categorical construct.

From a philosophical standpoint, it seems unlikely that alexithymia can be systematically organized into subtypes; instead, it is more likely that there exist various *degrees* of alexithymia. This is because emotional awareness is not categorical – that is, one is not either completely conscious (i.e., aware) of one's emotional state or completely unconscious (i.e., unaware) of it. Instead, as Lane and Schwartz (1987) demonstrate, there are various degrees of emotional awareness, ranging from mere awareness of bodily sensations to full blown awareness, whereby one has the capacity for empathy and self-insight.[85] Bermond's characterization of Type I alexithymia assumes that Type I alexithymics have absolutely *no* emotional awareness. However, it is likely that these individuals are, at the very least, aware of (at least some of) their bodily sensations. Other individuals labeled as having Type I alexithymia might have a slightly higher degree of emotional awareness, and be aware not only of their bodily sensations, but also of their hedonic states (i.e., general states of pleasure or dis-pleasure).[86] Indeed, distinguishing subtypes of alexithymia seems futile, since it is likely that there will always be a variation

[85] According to Lane and Schwartz (1987) the five levels of emotional awareness are: 1) sensorimotor reflexive (i.e., awareness of bodily sensations) – at this level, the individual will report that they feel nothing, or bodily sensation only; 2) sensorimotor enactive (i.e., awareness of the body in action) – at this level, the individual will describe action tendencies or global hedonic states; 3) preoperational (i.e., awareness of individual feelings) – at this level, representation of emotion is possible, and so the quality of emotion changes such that it becomes a psychological as well as a somatic experience; 4) concrete operational (i.e., awareness of blends of feelings) – at this level, an individual will describe complex and differentiated emotional states that capture their subjective experience; and 5) formal operational (i.e., awareness of blends of blends of feelings) – at this level, an individual has the capacity for empathy and self-insight (see p. 138-139).

[86] Lane and Schwartz (1987) note that at this level of emotional awareness, an individual's ability to experience emotion as a conscious feeling state has not yet developed – and so, "the words used to convey the hedonic tone of the experience would not refer to emotion alone, e.g., "I feel bad"" (p. 138).

of emotional awareness even between members of the same subgroup. For these reasons, I am here going to assume that alexithymia is of a dimensional (i.e., graded) rather than categorical nature.

As we have seen, alexithymics generally appear to be unaware of their emotional feelings; or, as Frawley and Smith (2001) claim: they "have lost the feel of their feelings" (p. 190). According to Prinz (2004), this lack of emotional awareness is best understood as the result of a *high-level* deficit in emotional processing. On his account, alexithymics have difficulty saying how they feel because they are unable to *recognize* how they are feeling. [87] In the following section, I argue that this is incorrect, since alexithymia is best understood as a *low-level* deficit in emotional processing.

### 5.3 Understanding Mild Alexithymia through an Analogy with Blindsight

According to Prinz (2004), alexithymia is best understood as an analogue of *associative visual object agnosia* (hereafter, AVOA) – a visual condition whereby individuals are unable to recognize visually presented objects (i.e., individuals are unable to interpret, understand, or assign meaning to visually presented objects).[88] Although they are unable to recognize these objects, associative visual object agnosics are nevertheless able to

[87] Although Prinz (2004) does discuss alexithymia in his work, his discussion is largely superficial, and is primarily focused on his general theory of emotional consciousness (see p. 217). Other than Prinz' brief mention of the condition, alexithymia is not, at least to the best of my knowledge, discussed in the philosophical literature. This is both surprising and disappointing, since alexithymia presents us with an example of a severe emotional deficit. Understanding this deficit will undoubtedly help us to better understand emotional consciousness and the nature of emotions more generally.

[88] There are a number of subtypes of associative agnosia, including: category specific agnosia, prosopagnosia, simultanagnosia, pure alexia, color agnosia, etc. (for descriptions see Ghadiali, 2004, p. 19-20). As I understand the condition, AVOA involves a difficulty in recognizing a *variety* of visually presented objects (as opposed to, for example, prosopagnosia, which involves a mere difficulty in recognizing faces).

successfully copy drawings of them. This suggests that they are not blind to shape. Indeed, Prinz notes that associative visual object agnosics "are blind to the meanings of shapes, not to the shapes themselves" (p. 217). To put it differently, associative visual object agnosics develop accurate percepts of visually presented objects, but are unable to assign meaning to them.

Similarly, Prinz argues that although alexithymics are unable to recognize their emotion states, they nevertheless still experience emotions. For Prinz, emotions are perceptions of physiological changes (where perceptions can be either conscious or unconscious). On his account, understanding alexithymia through an analogy with AVOA is appropriate because much as individuals with AVOA perceive visual objects on a low-level (indicated by their ability to successfully copy drawings of them) but fail to recognize (or assign meaning to) those objects, individuals with alexithymia perceive the physiological changes occurring in their bodies at a low-level but fail to recognize (or assign meaning to) those changes. Thus, for Prinz, "like associative [visual] agnosics, who experience shape without comprehension, alexithymics may experience emotions without comprehension" (p. 217).

Although Prinz' analogy with AVOA might initially seem compelling, there are problems with it. AVOA involves a deficit in *high-level* visual processing (Ghadiali, 2004). This means that although low-level (or early) visual processing spaces are intact (i.e., non-defective), high-level processing spaces are defective.[89] Since low-level processing spaces are intact, associative visual object agnosics are unproblematically able

[89] According to Ghadiali (2004), "associative visual agnosia is usually the result of bilateral damage to the inferior temporo-occipital junction and subjacent white matter" (p. 19).

to develop percepts of visual objects; however, the defect in high-level processing spaces prevents them from recognizing (or assigning meaning to) the objects.

This high-level processing deficit is not analogous to the deficit associated with alexithymia. This is because AVOA is all-or-nothing, whereas alexithymia occurs in degrees. Unlike most individuals with AVOA who are *never* able to recognize visually presented objects, *some* alexithymics – in particular, some alexithymics with mild degrees of alexithymia (in my terminology, *high functioning alexithymics*) – are (at least sometimes) able to recognize their emotion states.[90] For example, in their clinical illustration of a patient with alexithymia, Taylor et al. (1997) describe a doctor asking a patient if she ever becomes angry. The patient claims that she does get angry, but implies that she only "becomes aware of her anger…when it is intense" (p. 34). Although this indicates a restricted emotional experience (since the patient is only able to experience the anger when it is intense), she is nevertheless able to recognize that she *has* experienced anger.[91]

Further problems with Prinz' analogy arise when we consider that on his account, emotions register physiological changes, but they represent core relational themes. However, some alexithymics – in particular, some alexithymics with severe degrees of alexithymia (in my terminology, *low functioning alexithymics*)[92] – fail to recognize *both*

[90] On my use of the term, high functioning alexithymics have mid-to-high levels of emotional awareness.
[91] It is important to note that although some alexithymics *can* (at least sometimes) recognize their emotion states, it appears to take a great deal of effort for them to do so. Moreover, they are seemingly unable to discuss the emotion state in any detail. In this way, their level of emotional awareness is still much lower than a non-alexithymic's.
[92] On my use of the term, low functioning alexithymics have low-to-extremely low levels of emotional awareness.

the physiological changes occurring in their bodies *and* particular features in their external environment that reliably cause those changes (in specific, they fail to recognize posed facial expressions of emotion). In developing his analogy with AVOA, Prinz focuses exclusively on interoceptive stimuli (i.e., stimuli originating from inside of the body). However, in order to account for various degrees of alexithymia, it seems to me that Prinz needs to say something about emotion elicitors. Indeed, Prinz' failure to discuss emotion elicitors is surprising, since AVOA is essentially a deficit in the ability to recognize exteroceptive stimuli (i.e., stimuli originating outside of the body).

The basic problem with Prinz' analogy is that both severe and mild alexithymia seem to involve a *low-level* emotional processing deficit. For now, let us focus on *mild alexithymia*. Low-level emotional processing deficits are characterized by intact high-level emotional processing spaces (e.g., working memory), and defective low-level (or early) emotional processing spaces (e.g., the amygdala). This low-level deficit accounts for the fact that high functioning alexithymics are (at least sometimes) able to recognize emotion states. If mild alexithymia was an analogue of AVOA (and thus involved a high-level emotion processing deficit), high functioning alexithymics would *never* be able to recognize emotion states.[93]

[93] As I understand it, most individuals with AVOA are *never* able to recognize visually presented objects. I take it that this is because AVOA is associated with damage to the cortex (Ghadiali, 2004, p. 19). Some associative visual object agnosics are (at least sometimes) able to use shape cues to accurately identify presented objects, but this is not *true* recognition – it is mere inference (Baugh et al., 2010, p. 30). Certainly, associative visual object agnosics are unable to identify objects in the same way that non-associative visual agnosics do – via access to memories or knowledge associated with the object.

Instead, it seems to me that the low-level emotional processing deficit associated with mild alexithymia is more similar to the low-level visual deficit seen in blindsight.[94] Blindsight is a fascinating phenomenon whereby individuals who are cortically blind are nevertheless able to accurately (or, at least above chance levels) respond to stimuli that are presented in their blind spot. This suggests that they perceive the visual stimuli at some level, but are not consciously aware of doing so. In other words, individuals with blindsight seem to process information about visual stimuli on an unconscious level (i.e., below the subjective threshold of awareness). Since high functioning alexithymics deny the experience of emotions, but nevertheless manifest appropriate autonomic responses (e.g., increased heart rate, increased blood pressure, muscle tension, etc.,)[95] and behavioral responses (e.g., smile, frown, etc.), they appear to process emotional information (i.e., information about (at least some) physiological changes in their bodies) unconsciously. And so, to use Lane et al.'s (1997) terminology, alexithymia appears to be a kind of "blindfeel".[96]

[94] Lane et al. (1997) conceptualize alexithymia in this way (i.e., as the emotional equivalent to blindsight). Notice, the same deficit characteristic of blindsight can be seen in other blind syndromes, including: blindtouch, blindsmell, etc.

[95] I claim that high functioning alexithymics manifest an *appropriate* autonomous response to an emotion elicitor; however, there is considerable inconsistency in the scientific literature regarding the relationship between alexithymia and physiological responses to emotion elicitors. Some studies have demonstrated that alexithymia is associated with an increased level of physiological arousal (e.g., Infrasca, 1997), while others demonstrate that it is associated with a decreased level of physiological arousal (e.g., Newton & Contrada, 1994) – to see possible explanations for these inconsistencies see Luminet et al. (2004), p. 743. Regardless of their *level* of physiological arousal, it seems to me that high functioning alexithymics nevertheless exhibit appropriate physiological responses – e.g., when viewing disturbing emotional scenes, high functioning alexithymics might exhibit fewer skin conductance responses than non-alexithymics (Roedema & Simons, 1999), but this response is still appropriate because it is the same type of response exhibited by non-alexithymics. An example of an inappropriate response to the disturbing emotional scenes would be if high functioning alexithymics did not exhibit *any* skin conductance responses.

[96] One could perhaps draw a parallel between the dissociation of implicit emotional states from conscious representations of those states (i.e., mild alexithymia as a kind of "blindfeel" condition), and the

Blindsight is typically understood as a low-level deficit in the processing of implicit (i.e., unconscious) and explicit (i.e., conscious) information. As Frawley and Smith (2001) note: "…relevant information is computed at a low level, implicitly, but the results of this computation are defectively reported out in explicit form to some processing space (like working memory) for uptake and use by other cognitive domains…" (p. 190). And so, individuals with blindsight process information about visual stimuli unconsciously, but fail to have that information sent out in explicit form to higher-level processing spaces that are responsible for conscious awareness. As a result, these individuals will deny awareness of any visual experience in their blind spot, but nevertheless still behave in ways that suggest they have (at least some) visual experience.

The same kind of explanation can be used to explain mild alexithymia. Some high functioning alexithymics are only able to perceive physiological changes on an unconscious level.[97] An emotion elicitor (i.e., a triggering emotional object or event) causes implicit low level emotional responses (e.g., autonomic and neuroendocrine changes), but information about these responses fails to be sent out in explicit form to higher-level processing centers responsible for conscious awareness. And so, when an emotion is elicited, high functioning alexithymics will manifest an autonomic response and a behavioral response (due to the implicit processing of physiological changes), but

dissociation of operational memory from episodic memory (e.g., the ability to perform a well-practiced and complex operation without the ability to describe instances of having performed the operation).

[97] Recall that alexithymia occurs in degrees – it is not the case that all high functioning alexithymics are able (at least sometimes) to recognize their emotion states. Indeed, the majority of high functioning alexithymics appear to be unable to access information about their physiological responses (i.e., they are unaware of most of the physiological changes occurring in their bodies), and as such, are unable to recognize their emotion state. It is these individuals (i.e., high-functioning alexithymics who are *unable* to recognize their emotion states) that I am here concerned with.

will typically claim that they do not know how they feel (Frawley & Smith, 2004, p. 190; Lane et al., 1997, p. 838). In this way, high functioning alexithymics seem to experience emotions – or at least the physiological changes typical of emotions – unconsciously (i.e., in the absence of conscious awareness).

Recall, however, that some high functioning alexithymics *can* report restricted emotional experiences. When they make these reports, they claim to *feel something* (typically, a bodily sensation).[98] This indicates that implicit information about physiological changes *has* been sent out in explicit form to higher-level processing spaces. At first glance, this seems to go against the analogy with blindsight, since individuals with blindsight are unable to *actually* see visual stimuli presented in their blind spot.[99] However, according to Lane et al. (1997), "…with increasing salience of the visual stimulus, such as luminous contrast or velocity of movement, [individuals with blindsight] may acknowledge a feeling associated with the visual stimulus, such as a wavelength impression" (p. 839), at least if they have been trained to develop this ability.

Similarly, in the case of mild alexithymia, if the bodily sensations induced by emotion elicitors are intense enough, information about those sensations could be sent out to higher-level processing centers, enabling high functioning alexithymics to become aware of them. Of course, mere awareness of one's bodily sensations is the most basic form of emotional awareness (Lane & Schwartz, 1987). However, with a great deal of training (perhaps with the help of modified psychotherapy) high functioning alexithymics

[98] Studies have shown that some high functioning alexithymics are more aware of the physiological changes occurring in their bodies than are others (Roedema & Simons, 1999).
[99] Indeed, Prinz (2004) claims that the analogy with blindsight predicts that emotions cannot feel like anything to the alexithymic (p. 217, footnote 3).

*could* develop the ability to pair their bodily sensations with the particular emotion state(s) that they signify.[100] Once they develop this ability, it is likely that their level of emotional awareness would increase – e.g., emotion states would come to have a phenomenal quality associated with them.

Now, there is one important difference between blindsight and alexithymia that is worth mentioning. While blindsight is caused by damage to the primary visual cortex, alexithymia (both mild and severe) is not always caused by cortical damage – it can also be brought on by severe psychological trauma (Sifneos, 1988, 1994; Taylor et al., 1997).[101] It seems probable, therefore, that alexithymics without cortical damage would have better uptake of information from low-level processing spaces to high-level processing spaces (since there is no damage to physically prevent the flow of information to higher processing spaces). And so, high functioning alexithymics without cortical damage might be able to report on their emotional experiences more often, and perhaps with greater detail, than high functioning alexithymics with cortical damage. However, in the absence of empirical studies comparing the processing abilities of high functioning alexithymics without cortical damage to high functioning alexithymics with cortical damage, this remains just a hypothesis (I am currently unaware of any such studies).

[100] Studies indicate that non-alexithymics and alexithymics respond better to different types of psychotherapy. Taylor et al. (1997) note that while non-alexithymics respond better to a cognitive-behavioral approach, alexithymics respond better to a clinical management approach (p. 266).

[101] Sifneos (1988, 1994) uses the term "primary alexithymia" to refer to alexithymic characteristics that are caused by neurobiological deficit(s); and uses the term "secondary alexithymia" to refer to alexithymic characteristics that are caused by massive psychological trauma either in childhood or later life, developmental arrests, socio-culture factors, or psychodynamic factors. As Taylor et al. (1997) note, secondary alexithymia is "state-dependent and disappears after the evoking stressful situation has changed" (p. 35).

To this point, I have argued that mild alexithymia is best understood as involving a low-level deficit in emotional processing. This deficit is similar to the low-level visual processing deficit associated with blindsight. If, as Prinz (2004) argues, alexithymia was an analogue of AVOA, and thus involved a high-level deficit in emotional processing, no alexithymic (high functioning or low functioning) would *ever* be able to recognize their emotion states. However, as we have seen, *some* high functioning alexithymics *are* able to recognize their emotion states (Taylor et al., 1997, p. 34). Since the majority of high functioning alexithymics deny the experience of emotion, but nevertheless still manifest appropriate autonomic and behavioral responses, they appear to process emotional information unconsciously. In this way, mild alexithymia is most usefully understood as a condition that is analogous to blindsight.

What about severe alexithymia? In the following section, I argue that the deficit characteristic of severe alexithymia is not analogous to the deficit associated with *either* AVOA *or* blindsight. Instead, severe alexithymia is best understood as a condition that is analogous to *apperceptive visual agnosia* (i.e., a visual condition whereby an individual is unable to access certain properties of visually presented objects).

## 5.4 Understanding Severe Alexithymia through an Analogy with Apperceptive Visual Agnosia

The most important element of emotional processing is connecting the cause of an emotion (i.e., the emotion elicitor) with higher-level appraisal mechanisms (Frawley & Smith, 2001, p. 194). Indeed, for an individual to have emotional awareness (even at the most basic level), information about emotion elicitors must be made available to higher-level processing centers that evaluate them for emotional content (Frawley & Smith, 2001, p. 194). With this in mind, there is one salient feature of severe alexithymia that we need to consider: most low functioning alexithymics fail to respond in the usual way to emotion elicitors – particularly, to posed facial expressions of emotion (hereafter, PFEE).[102] For example, most low functioning alexithymics show a significantly lower level of physiological arousal to PFEE (Roedema & Simons, 1999; Wehmer, et al., 1995; Gilbert, 2008) than most high functioning alexithymics.

This decreased level of physiological arousal is not surprising, since studies have shown that most low functioning alexithymics have extreme difficulty identifying PFEE – indicating that they are extremely poor at recognizing other people's emotion states (Taylor et al., 1997; Parker et al., 1993 (b), Jessimer & Markham, 1997; Prkachin &

[102] Studies have shown that most low functioning alexithymics also show a significantly lower level of physiological arousal to environmental emotion elicitors (e.g., landscapes, children laughing, victims of injury, etc.) (Wehmer, et al., 1995). However, these results are not consistent in the literature (for an alternative result see Infrasca, 1997). These inconsistencies, however, might be due to the different physiological responses being tested in each study (for further discussion see Luminet at al., 2004, p. 743).

Also, notice that "PFEE" refers to *photographs* of posed facial expressions. To the best of my knowledge, all of the studies that have explored the relationship between alexithymia and recognition of facial expressions of emotion have relied on photographs of facial expressions.

Prkachin, 2001). If low functioning alexithymics are unable to recognize PFEE, it follows that they would not exhibit prototypical physiological responses towards them.

Now, some high functioning alexithymics are also poor at identifying PFEE; however, their accuracy levels are much higher than those of low functioning alexithymics (Jessimer & Markham, 1997, p. 253). For this reason, I am here only going to focus on low functioning alexithymics – though much of what I say is likely applicable to *some* high functioning alexithymics. It is important to keep in mind that, unlike some high functioning alexithymics who are (at least sometimes) able to recognize *both* their own emotion states *and* other people's emotion states, most low functioning alexithymics are always unable to recognize their own emotion states, and are *extremely poor* at recognizing other people's emotion states.

Being extremely poor at identifying other people's emotion states, however, does not imply that low functioning alexithymics are unaware of what is occurring in their environment. In fact, as noted, alexithymics in general have an external thought pattern; that is, they are hyper-focused on external events and objective facts (e.g., they tend to recount trivial events of everyday life in chronological order and in monotonous detail). The problem, however, seems to be that most low functioning alexithymics are unable to *recognize* facial expressions of emotion.

More specifically, since low functioning alexithymics exhibit an extremely low level of physiological arousal to PFEE, they appear to be unable to recognize this type of emotion elicitor on a *low-level*. That is, most low functioning alexithymics appear to be unable to develop complete (or accurate) percepts of facial expressions of emotion. If

they were able to develop complete percepts, then one would expect to see arousal levels similar to (at least) those exhibited by high functioning alexithymics. Of course, it is possible for two individuals with normally functioning visual systems to see the same thing – e.g., a photograph of an individual with a scowl and furrowed eyebrows – and yet have different emotional responses to it. For example, the photograph might evoke a high level of arousal in an individual who has experienced physical abuse, and a low (or perhaps absent) level of arousal in an individual who has never experienced abuse. This difference in arousal level is not caused by a deficit at the perceptual level (since presumably both individuals can identify the expression in the photograph as one of anger), but instead by a deficit at a later stage of information processing.

However, the deficit associated with severe alexithymia seems to be specifically perceptual. Most low functioning alexithymics (and indeed some high functioning alexithymics) are unable to detect certain properties of facial expressions – specifically those properties associated with expressions of anger, sadness, and fear (e.g., Prkachin & Prkachin, 2001). For example, when looking at a photograph of an individual with an angry facial expression, a low functioning alexithymic might fail to detect their furrowed eyebrows/or scowl; thus preventing them from recognizing the expression as anger. Since most low functioning alexithymics are unable to detect certain properties of facial expressions, it follows that they would exhibit low levels of arousal towards them.[103]

[103] It seems to me that the more properties of a facial expression one is able to detect, the more accurate they will be at identifying the facial expression, and the higher their level of arousal will be. This explains why some high functioning alexithymics are both more accurate at identifying PFEE, and exhibit higher levels of arousal towards them than most low functioning alexithymics. Similarly, some high functioning alexithymics exhibit a lower level of arousal to PFEE than non-alexithymics – it seems to me that this is

Given how poor most low functioning alexithymics are at identifying PFEE, their deficit seems to be even lower than the emotional processing deficit associated with most high functioning alexithymics. This extremely low-level deficit in emotional information processing is, I shall now argue, similar to the extremely low-level visual processing deficit associated with apperceptive visual agnosia (hereafter AVA).

AVA is a visual condition whereby individuals are unable to access spatial and structural properties of visually presented stimuli. The condition is characterized by an intact visual ability on a basic sensory level, but a defect in early stage visual processing (Ghadiali, 2004). This means that although apperceptive visual agnosics have intact elementary visual functions (e.g., acuity, color vision, and brightness discrimination) the defect in early stage visual processing prevents them from forming accurate percepts of visual stimuli. As such, apperceptive visual agnosics are unable to recognize, copy, or distinguish between objects. For example, despite their different shapes, an apperceptive visual agnosic would be unable to distinguish between a DVD (which is round) and a videotape (which is rectangular). Of course, apperceptive visual agnosics are able to use cues about an object's color, texture, and size to accurately identify it – but this is merely inferring the identity of an object, without ever actually seeing the object as a whole.

In both AVA and AVOA, individuals are unable to recognize objects. However, apperceptive visual agnosics are unable to recognize visual objects because they are unable to develop accurate percepts of them. Associative visual object agnosics, on the other hand, *are* able to develop accurate percepts of visual objects, but are unable to

---

because unlike non-alexithymics, high functioning alexithymics are not processing *all* properties of facial expressions.

recognize them for what they are because they are unable to access memories or knowledge associated with them (Ghadiali, 2004, p. 18). As we have seen, AVA is associated with defects in low-level (or early) visual processing spaces, whereas AVOA is associated with defects in high-level visual processing spaces. This accounts for the fact that while associative visual object agnosics can copy drawings of visually presented objects (since low-level processing spaces are intact), apperceptive visual agnosics cannot.

Severe alexithymia is not analogous to AVOA. If it were, then low functioning alexithymics would be able to recognize PFEE on a low-level (i.e., they would exhibit appropriate autonomic responses to PFEE), much as associative visual object agnosics are able to recognize visually presented stimuli on a low-level (indicated by their ability to accurately copy drawings of objects). However, as we have seen, this is not the case. Not only do most low functioning alexithymics exhibit significantly lower levels of physiological arousal to PFEE than most high functioning alexithymics (Roedema & Simons, 1999; Wehmer, et al., 1995; Gilbert, 2008), but there is also evidence that some low functioning alexithymics fail to exhibit certain physiological responses altogether. For example, in their study, Wehmer et al. (1995) found that 74% of low functioning alexithymics failed to exhibit *any* electrodermal response (i.e., skin conductance response) to all of the slides presented.[104]

[104] In comparison, 46% of high functioning alexithymics failed to exhibit *any* electrodermal response to all of the emotion-inducing slides presented. It is important to note that the slides used in this study did not consist solely of PFEE; they also included environmental emotion elicitors – e.g., four slides depicted victims of injury, accident, or starvation, four were of young children laughing or playing, four were of scenic landscapes, and four showed nude couples in romantic/erotic interactions. Although the authors were

I have argued that the deficit involved in mild alexithymia is analogous to the deficit involved in blindsight. However, as I will now argue, blindsight fails to provide a good analogy for the deficit found in severe alexithymia. Much as individuals with blindsight deny any visual experience in their blind spot, but nevertheless manifest behavioral responses indicative of (at least some) visual experience, high functioning alexithymics deny the experience of emotions, but nevertheless manifest appropriate autonomic and behavioral responses indicative of emotional experience. Information about these responses, however, is not sent out in explicit form to higher-level processing centers responsible for conscious awareness. As such, high functioning alexithymics fail to recognize the physiological changes occurring in their own bodies. For this reason, blindsight provides a useful analogy for understanding the deficit involved in high functioning alexithymia.

Most low functioning alexithymics, on the other hand, fail to recognize *both* the physiological changes occurring in their bodies *and* certain features in their external environment that reliably cause those changes – in particular, they fail to recognize PFEE. This is not to say that most low functioning alexithymics are unaware of exteroceptive stimuli (i.e., stimuli originating outside the body). After all, as noted, one of the defining characteristics of alexithymia is an external thought pattern. Thus, when

---

unclear about the details of the slides, the ones of the victims, children, and romantic couples presumably depicted more than just facial expressions of emotion (some were likely full body shots that included background scenery). In this way, the reactions of the low functioning alexithymics were not merely to facial expressions, but also to background scenery, body language, etc.). In order to gain a more accurate percentage of how many low functioning alexithymics fail to exhibit any electrodermal responses to PFEE, one would have to run the experiment using slides that *only* showed facial expressions of emotion. I am currently aware of only one study like this (see Gilbert, 2008). In this study, Gilbert did not find an *absence* of certain physiological responses in low functioning alexithymics, but she did find a *reduced level* of arousal.

low functioning alexithymics see a photograph of a facial expression, they are not only aware of the photograph, but are aware that the photograph is of a particular individual's face (e.g., Barack Obama). However, most are *typically* unable to recognize what emotion the individual is portraying (recall that *some* low functioning alexithymics *are* able to recognize other people's emotion states, but are *extremely* poor at doing so).

Unlike the deficit associated with blindsight, the inability of most low functioning alexithymics to recognize PFEE cannot be explained as a deficit in the processing of implicit and explicit information. This is because most low functioning alexithymics do not appear to have implicit knowledge of certain properties of facial expressions of emotion. As such, there is no information to be reported out in explicit form to higher-level processing spaces. This is demonstrated by the fact that most low functioning alexithymics are *both* unable to accurately identify PFEE, *and* fail to manifest appropriate autonomic and behavioral responses towards them.

This deficit is similar to the deficit associated with AVA. Of course, low functioning alexithymics *can* sometimes recognize other people's emotion states. However, in these cases, it is unclear whether they exhibit the same low levels of physiological arousal that they exhibit when they are unable to recognize PFEE. To the best of my knowledge, this information is not available in any of the psychological literature. Nevertheless, it seems to me that if low functioning alexithymics exhibit the same low levels of physiological arousal in both cases (i.e., in cases where they are *unable* to accurately identify PFEE, and cases where they *are* able to accurately identify them), then, in cases where they *are* able to accurately identify PFEE, they are merely

guessing what facial expression is being presented. If low functioning alexithymics were truly able to recognize PFEE, then one would expect them to exhibit higher levels of physiological arousal towards them. Indeed, it seems likely that the more attributes (or properties) of PFEE an alexithymic is able to recognize, the higher their level of physiological arousal.[105]

Thus, if low functioning alexithymics exhibit higher levels of physiological arousal when they *are* able to accurately identify PFEE, as opposed to when they are *not* able to do so, it seems probable that they are able to develop accurate percepts of at least *some* properties of posed facial expressions. However, unless low functioning alexithymics exhibit physiological levels of arousal similar to those found in non-alexithymics, it can be assumed that they are not developing accurate percepts of *all* properties of PFEE.

At this point, it might be objected that being poor at identifying other people's emotions is only a *secondary feature* of alexithymia – it is not one of the four defining characteristics of the condition. As such, understanding this deficit will not advance our understanding of the core deficit involved in alexithymia – namely, the inability to recognize one's *own* emotions. However, it seems to me that the inability to recognize other people's emotions might be a useful diagnostic tool – it could help to determine the

[105] One can imagine cases in which an individual is able to recognize PFEE but nevertheless still exhibit low (or non-existent) levels of physiological arousal towards them – as is characteristic of psychopathy. Like some alexithymics, psychopaths fail to exhibit insight into the emotional feelings of others, are unable to introspect on their own feelings, and often create superficial narratives about their feelings in social situations. However, while alexithymics are typically socially conforming, anxious, ethically consistent, and submissive, psychopaths are non-conforming, anxiety-free, ethically inconsistent, and dominant (see Haviland et al., 2004). Although there is a tendency to link alexithymia and psychopathy in the psychological literature, the relationship between the two constructs is currently unclear (Haviland et al., 2004).

severity of an individual's alexithymia (e.g., the more inaccurate/poor an individual is at identifying other people's emotions the more severe is their degree of alexithymia). Moreover, etiological studies have shown that the deficit in emotion recognition associated with alexithymia is attributable to a dysfunction of the right hemisphere (Sifneos, 1988; Kano et al., 2003) – and that most low functioning alexithymics show less levels of right hemisphere activity than most high functioning alexithymics (Jessimer & Markham, 1997). I think this is an important finding that should be explored further, since it will help us to better understand more severe degrees of the condition.

In this chapter, I have shown that there is both anecdotal and experimental evidence which demonstrates that emotions can be unconscious. Unlike cognitive theories of the emotions, which maintain that emotions cannot be *completely* unconscious, my somatic-perceptual theory of emotions is able to account for unconscious emotions. This is because on my account emotions are perceptions of changes in the body – and perceptions can be either conscious or unconscious. Since my somatic-perceptual theory of the emotions allows for unconscious emotions, it provides a way of thinking about poorly understood psychological conditions, such as alexithymia (which is best understood only on the assumption that emotions can be unconscious).

I argued that alexithymia cannot be systematically decomposed into sub-types or categories, and that it is instead best understood as having a dimensional (i.e., graded) nature. On the one hand, I argued that mild alexithymia involves a low-level emotional processing deficit that is most appropriately understood through an analogy with blindsight. This is because high functioning alexithymics appear to implicitly process

information about physiological changes at a low level, but fail to have that information sent out in explicit form to higher-level processing spaces responsible for conscious awareness. Much as individuals with blindsight deny any visual experience in their blind spot, but nevertheless manifest behavioral responses indicative (of at least some) visual experience, high functioning alexithymics deny emotional experience, but nevertheless manifest autonomic and behavioral responses indicative of emotional experience.

On the other hand, I argued that severe alexithymia involves a lower-level emotional processing deficit than mild alexithymia. Unlike high functioning alexithymics who merely fail to recognize the physiological changes occurring in their bodies as indicative of an emotional experience, low functioning alexithymics fail to recognize *both* the physiological changes occurring in their bodies *and* certain features in their external environment that reliably cause those changes. In particular, they fail to recognize PFEE. Since low functioning alexithymics are extremely poor at identifying PFEE, and exhibit extremely low levels of physiological arousal towards them, I have argued that it is likely that low functioning alexithymics are unable to develop accurate percepts of this type of emotion elicitor. This deficit, I have argued, is most appropriately understood through an analogy with AVA.

Of course, blindsight and AVA are both conditions that are not yet fully understood, and so it might be objected that it is problematic to understand other conditions through analogy with them. However, it seems to me that we know a great deal more about both blindsight and AVA than we do about alexithymia. As a result, I think both blindsight and AVA can be useful for understanding mild alexithymia and

severe alexithymia, respectively. Of course, upon further investigation it might turn out that these analogies are no longer appropriate. For the moment, however, I think they are the best analogies available.

Despite decades of research, alexithymia is still incompletely understood. Not only is there currently no comprehensive account of the condition currently available, but the term itself – "alexithymia" – is used inconsistently throughout the literature. This makes navigating the literature difficult, as researchers very often seem to be talking past each other. Nevertheless, it is crucial that we get clear on the deficits involved in alexithymia. This is because alexithymia presents us with an example of a disturbance in emotional experience. Understanding the condition will undoubtedly prove helpful in our quest to understand the nature of emotions more generally.

# CONCLUDING REMARKS

In this dissertation, I developed and defended a somatic-perceptual theory of the emotions. On this account, emotions are perceptions of physiological changes. More specifically, emotions are states in one's somatosensory cortex that register physiological changes. My somatic-perceptual theory of the emotions stands in contrast to cognitive theories of the emotions. The majority of emotion theorists hold some kind of a cognitive theory of the emotions – a theory in which cognition is necessarily involved in any emotion. I argued, in opposition to these theories, that cognition is never necessary to emotion. Instead, I argued that the perception of physiological changes alone is necessary for emotion. Additionally, I showed that my somatic-perceptual theory of the emotions is able to account both for the intentional nature of the emotions and the distinctiveness of different emotions just as well as leading cognitive theories of the emotions. This is significant because these objections have not yet been adequately met by other somatic theories of the emotions (or non-cognitive theories of the emotions more generally).

In the first chapter, I offered a sketch of the somatic-perceptual theory of the emotions that I defend and distinguished it from the somatic theories of William James, Antonio Damasio, and Jesse Prinz. I also highlighted inconsistencies and areas in Prinz' account where he was unclear about what he takes the basic structure of the emotions to be. In the second chapter, I surveyed some prominent cognitive theories of the emotions

and argued, in opposition to these theories, that cognition is never an essential part of an emotion. I then used evidence from psychology and neuroscience to argue that the perception of physiological changes *alone* is necessary for emotion. In the third chapter, I offered a response to a prominent objection against all somatic theories of the emotions – the intentionality objection. I argued that emotions register changes in the body, but represent core relational themes. By doing this, I showed that my somatic-perceptual theory of the emotions is able to meet the intentionality objection just as well as prominent cognitive theories of the emotions. In addition, I showed that my somatic-perceptual theory of the emotions is able to distinguish between emotions and other non-emotional interoceptive states. In the fourth chapter, I offered a response to the physiology objection against somatic theories of the emotions. I argued that, on my somatic-perceptual theory, emotions can be distinguished *both* by the physiological changes that they register *and* by their causal histories. Since, on my account, emotions are perceptions of changes in the body, it should be possible, on my account, for emotions to be unconscious. In the fifth and final chapter I embraced this implication, and argued that emotions *can* be unconscious. In this chapter, I also showed how the somatic-perceptual theory that I defend provides a framework for thinking about poorly understood psychological conditions, such as alexithymia.

Emotions play a fundamental role in the quality our lives. We typically strive to achieve the feelings associated with emotions such as happiness, love, and pride, and avoid the feelings associated with emotions such as sadness, anger, and guilt. Despite their obvious importance, however, there is no single agreed upon theory of the emotions

in either the philosophical or psychological literature. I believe that this is partly due to the fact that research on the emotions has typically lacked true interdisciplinarity. Although there are a number of emotion theories in the literature, most are either purely philosophical or purely psychological. Not only do philosophers and psychologists use emotion terms differently, but they also tend to talk past one another. Indeed, when it comes to current research on the emotions, philosophers and psychologists appear either to be unaware or uninterested (or both) in the work that is being done in one another's discipline. In this dissertation, I defended a theory of the emotions that is interdisciplinary in nature – a theory that is both philosophically persuasive and consistent with the most recent empirical research. As such, this theory is, I believe, the strongest theory of the emotions currently available.

# References

Arnold, Magda. *Emotion and Personality.* New York: Columbia University Press, 1960.

Arntz, Arnoud. "Endorphins Stimulate Approach Behavior, But do not Reduce Subjective Fear: A Pilot Study." *Behavior Research and Therapy* 31.4 (1993): 403-405.

Bagby, Robert M., et al. "Are there subtypes of Alexithymia?" *Personality and Individual Differences* 47.5 (2009): 413-418.

Baillargeon, Renée, et al. "False-Belief Understanding in Infants." *Trends in Cognitive Science* 14.3 (2010): 110-118.

Barlassina, Luca and Albert Newen. "The Role of Bodily Perception in Emotion: In Defense of an Impure Somatic Theory." *Philosophy and Phenomenological Research* 88.1 (2013): 23 Aug. 2014 <http://onlinelibrary.wiley.com/doi/10.1111/phpr.12041/abstract>.

Bartels, Andreas and Semir Zeki. "The Neural Basis of Romantic Love." *NeuroReport* 11.17 (2000): 3829-3834.

Baugh, Lee A., et al. "Agnosia." In *Encyclopedia of Behavioral Neuroscience*. Ed. George Koob, Michel Le Moal, and Richard F. Thompson. London: Academic Press, 2010.

Bear, Mark F., et al. *Neuroscience: Exploring the Brain.* 3rd ed. Baltimore: Lippincott Williams & Wilkins, 2007.

Bedford, Errol. "Emotions." *Proceedings of the Aristotelian Society* 57 (1957): 281-304.

Bechara, Antoine. "The Role of Emotion in Decision Making: Evidence from Neurological Patients with Orbitofrontal Damage." *Brain and Cognition* 55.1 (2004): 30-40.

Ben Ze'ev, Aaron. *The Subtlety of Emotions.* Cambridge: MIT Press, 2000.

_____. "Appraisal Theories of Emotions." *Journal of Philosophical Research* 22 (1997): 129- 143.

Bermond, Bob. "Brain and Alexithymia." In *The (non) Expression of Emotion in Health and Disease.* Ed. Ad Vingerhoets, Frans van Bussen, and Jan Boelhouwers. Tilburg: Tilburg UP, 1997.

Blood, Anne J., and Robert J. Zatorre. "Intensely Pleasurable Responses to Music Correlate with Activity in Brain Regions Implicated with Reward and Emotion." *Proceedings of the National Academy of Sciences* 98.20 (2001): 11818–11823.

Broad, Charlie D (1954). "Emotion and Sentiment." In *Critical Essays in Moral Theory.* London: Allen & Unwin, 1971.

Buss, David M., et al. "Sex Differences in Jealousy: Evolution, Physiology, and Psychology." *Psychological Science* 3.4 (1992): 251-255.

Cannon, Walter B. "The James-Lange Theory of Emotions: A Critical Examination and an Alternative Theory." *The American Journal of Psychology* 39.1 (1927): 106-124.

Clore, Gerald. "Why Emotions are Never Unconscious." In *The Nature of Emotions: Fundamental Questions.* Ed. Paul Ekman and Richard J. Davidson. New York: Oxford UP, 1994. 285-290.

Chalmers, David. *Philosophy of Mind: Classical and Contemporary Readings.* Oxford: Oxford UP, 2002.

Cheesman, Jim and Philip Merikle. "Distinguishing Conscious from Unconscious Perceptual Processes." *Canadian Journal of Psychology* 40.4 (1986): 343-367.

Cushing, Paula E. "Myrmecomorphy and Myrmecophily in Spiders: A Review." *Florida Entomologist* 80.2 (1997): 165-193.

Damasio, Antonio. *The Feeling of What Happens: Body and Emotion in the Making of Consciousness.* London: Heinemann, 1999.

_____. *Descartes' Error.* New York: G. P. Putnam's Sons, 1994.

_____. "Time-locked Multiregional Retroactivation: A Systems Level Proposal for the Neural Substrates of Recall and Recognition." *Cognition* 33.1-2 (1989): 25-62.

Damasio, Antonio and Gary W. Van Hoesen. "Emotional Disturbances Associated with Focal Lesions of the Limbic Frontal lobe." In *Neuropsychology of Human Emotion.* Ed. Kenneth M. Heilman and Paul Satz. New York: Guilford Press, 1983. 85-110.

Deigh, John. "Cognitivism in the Theory of Emotions." *Ethics* 104.4 (1994): 824-854.

D'Esposito, Mark, et al. "Maintenance versus Manipulation of Information Held in Working Memory: An Event-Related fMRI Study." *Brain and Cognition* 41.66-86 (1999): 66-86.

De Waal, Frans. *Our Inner Ape: A Leading Primatologist Explains Why we Are Who we Are.* New York: Penguin Group, 2005.

Dretske, Fred. "Misrepresentation." In *Belief: Form, Content and Function.* Ed. Radu Bogdan. Oxford: Oxford UP, 1986.

_____. *Knowledge and the Flow of Information.* Cambridge: MIT Press, 1981.

Fischman, Marian W. and Richard W. Foltin. "Self-administration of Cocaine by Humans: A Laboratory Perspective." *Cocaine: Scientific and Social Dimensions.* Ed. Gregory R. Bock and Julie Whelan. 166.166-73 (1992): 165-180.

Fish, William. *Perception, Hallucination, and Illusion.* Oxford: Oxford UP, 2009.

Fitzgerald, Michael, and Guy Molyneux. "Overlap Between Alexithymia and Asperger's Syndrome."*American Journal of Psychiatry* 161.11 (2004): 2134-2135.

Fodor, Jerry. "A Theory of Content, I and II." In *A Theory of Content and Other Essay's.* Cambridge: MIT Press, 1990. 51-136.

_____. *Psychosemantics.* Cambridge, MIT Press, 1987.

Frawley, William and Raoul N. Smith. "A Processing Theory of Alexithymia." *Journal of Cognitive Systems Research* 2.3 (2001): 189-206.

Freud, Sigmund (1915). "The Unconscious." In *The Pelican Freud Library: Vol. 11. On Metapsychology: The Theory of Psychoanalysis.* Ed. A. Richards. Harmondsworth, UK: Penguin, 1984. 159-222.

Frith, Uta and Christopher Frith. "Development and Neurophysiology of Mentalizing." *Philosophical Transactions: Biological Sciences* 358.1431 (2003): 459-473.

Geertz, Clifford. *The Interpretation of Cultures: Selected Essays.* New York: Basic Books, 1973.

Ghadiali, Eric. "Agnosia." *Advances in Clinical Neuroscience and Rehabilitation* 4.5 (2004): 18-20.

Gilbert, Alison. "The Physiological Response to Implicit and Explicit Fear Faces in Alexithymia." *Diss. University of Pittsburg,* 2008.

Greenspan, Patricia. *Emotions and Reasons.* New York: Routledge, 1988.

Haviland, Mark G., et al. "Alexithymia and psychopathy: Comparison and application of California Q-set prototypes." *Journal of Personality Assessment* 82.3 (2004): 306-316.

Hinton, John M. "Visual experiences." *Mind* 76.302 (1967): 217-227.

Hohmann, Gerald. "Some Effects of Spinal Cord Lesions on Experienced Emotional Feelings." *Psychophysiology.* 3.2 (1966): 143-156.

Infrasca, Roberto. "Alexithymia, neurovegitative arousal and neuroticism." *Psychotherapy and Psychosomatics* 66.5 (1997): 276-280.

Izard, Carroll E. "Emotion-Cognition Relationships and Human Development." In *Emotions,Cognition, and Behavior.* Ed. Carroll E. Izard, Jerome Kagan, and Robert Zajonc. Cambridge: Cambridge UP, 1984.

James, William. "What is an Emotion?" *Mind 9.34*: 188-205.

Jessimer, Matthew and Roslyn Markham. "Alexithymia: A right hemisphere dysfunction specific to recognition of certain facial expressions?" *Brain and Cognition* 34.2 (1997): 246-258.

Johnson-Laird, Philip and Keith Oatley. "Cognitive and Social Construction in the Emotions." In *Handbook of Emotions.* Ed. Michael Lewis and Jeannette Haviland-Jones. 2nd ed. New York: Guilford Press, 2000. 458-475.

Kano, Michiko, et al. "Specific brain processing of facial expressions in people with alexithymia: An $H_2{}^{15}O$-PET study." *Brain* 126.6 (2003): 1474-1484.

Kassam, Karim, et al. "Identifying Emotions on the Basis of Neural Activation." *PLoS ONE* 8.6 (2013): 52 pars. 11 Mar. 2015. <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0066032>.

Kassam, Karim. "Mapping Emotions in the Brain." *Huffington Post* 21 Sept. 2013 (b). 20 Mar. 2015 <http://www.huffingtonpost.com/karim-kassam/mapping-emotions-in-the-b_b_3635636.html>.

Kauhanen, Jussi, et al. "Alexithymia and risk of death in middle-aged men." *Journal of Psychosomatic Research* 41.6 (1996): 541-549.

Kenny, Anthony. *Action, Emotion, and Will.* London: Routledge & Kegan Paul Limited, 1963.

Krystal, Henry. *Integration and Self-Healing: Affect, Trauma, Alexithymia*. New Jersey: The Analytic Press, 1988.

Lane, Richard D., et al. "Pervasive Emotion Recognition Deficit Common to Alexithymia and the Repressive Coping Style." *Psychosomatic Medicine* 62.4 (2000): 492-501.

Lane, Richard D., et al. "Is Alexithymia the Emotional Equivalent of Blindsight?" *Biological Psychiatry* 42.9 (1997): 834-844.

Lane, Richard D., et al. "Impaired Verbal and Nonverbal Emotion Recognition in Alexithymia." *Psychosomatic Medicine* 58.3 (1996): 203-210.

Lane, Richard D., et al. "The Levels of Emotional Awareness Scale: A Cognitive-developmental Measure of Emotion." *Journal of Personality Assessment* 55.1-2 (1990): 124-134.

Lane, Richard D. and Gary E. Schwartz. "Levels of Emotional Awareness: A Cognitive Developmental Theory and its Application to Psychopathology." *American Journal of Psychiatry* 144.2 (1987): 133-143.

Lange, Carl. (1885). In *The Emotions.* Ed. Carl Lange and William James. Trans. I. A. Haupt. Baltimore: Williams and Wilkins, 1922.

Larsen, Junilla K., et al. "Cognitive and Emotional Characteristics of Alexithymia: A Review of Neurobiological Studies." *Journal of Psychosomatic Research* 54.6 (2003): 533-541.

Lazarus, Richard. "The cognition-emotion debate: a bit of history." In *Handbook of cognition and emotion.* Ed. Tim Dalgleish and Mick Power. Chichester: Wiley, 1999.

_____. *Emotion and Adaptation.* New York: Oxford UP, 1991.

LeDoux, Joseph. *The Emotional Brain*. New York: Simon & Schuster, 1996.

_____. "Emotional processing, but not emotions, can occur unconsciously." In *The Nature of Emotion: Fundamental Questions.* Ed. Paul Ekman and Richard J. Davidson. New York: Oxford UP, 1994. 291-292.

Leeper, Robert W. "Some needed developments in the motivational theory of emotions." In *Nebraska Symposium on Motivation.* Ed. David Levine. Vol. 13. Lincoln: U of Nebraska P, 1965.

Levensen, Robert W., et al. "Voluntary Facial Action Generates Emotion-Specific Autonomous Nervous System Activity." *Psychophysiology* 27.4 (1990): 363-384.

Leventhal, Howard. "Emotions: A basic problem for social psychology." In *Social Psychology.* Ed. Charlan Nemeth. Chicago: Rand McNally, 1974.

Luminet, Olivier, et al. "A multimodal investigation of emotional responding in alexithymia." *Cognition and Emotion* 18.6 (2004): 741-766.

Mandler, George. *Mind and emotion.* New York: Wiley, 1975

Marañon, Gregorio. "Contribution à L'Étude De L'Action Émotive De L'Adrenaline." *Revue Française d'Endocrinologie* 2 (1924): 301–325.

Marshall, Gary D. and Philip G. Zimbardo. "Affective consequences of inadequately explained physiological arousal." *Journal of Personality and Social Psychology* 37.6 (1979): 970-988.

Martin, Michael, G.F "The Transparency of Experience." *Mind and Language* 17.4 (2002): 376-425.

Maslach, Christina. "Negative emotional biasing of unexplained arousal." *Journal of Personality and Social Psychology* 37.6 (1979): 953-969.

Mattila, Aino K., et al. "Health related quality-of-life profiles in nonalexithymic and alexithymic subjects from general population." *Journal of Psychosomatic Research* 68.3 (2010): 279-283.

McDonald, P.W. and Kenneth M. Prkachin. "The expression and perception of facial emotion in alexithymia: A pilot study." *Psychosomatic Medicine* 52.2 (1990):199-210.

Morris, J., et al. "A Subcortical Pathway to the Right Amygdala Mediating "Unseen" Fear." *Proceedings of the National Academy of Science* 96.4 (1999):1680-1685.

Mutschler, Isabella, et al. "Structural basis of empathy and the domain general region in the anterior insular cortex." *Frontiers in Human Neuroscience* 7.177 (2013). 29 July 2014 <http://journal.frontiersin.org/Journal/10.3389/fnhum.2013.00177/full>.

Nash, Ronald. "Cognitive Theories of Emotion." *Noûs* 23 (1989): 481-504.

Nemiah, John C., et al. "Alexithymia: A View of the Psychosomatic Process." In *Modern Trends in Psychosomatic Medicine.* Ed. Oscar W. Hill.Vol. 3. London: Butterworths, 1976.

Nemiah, John C. and Peter E. Sifneos. "Affect and Fantasy in Patients with Psychosomatic Disorder." In *Modern Trends in Psychosomatic Medicine.* Ed. Oscar W. Hill. Vol. 2. New York: Appleton-Century-Crofts, 1970.

Neu, Jerome. *A Tear is an Intellectual Thing: The Meanings of Emotions.* Oxford: Oxford UP, 2000.

Newton, Tamara L. and Richard J. Contrada. "Alexithymia and repression: Contrasting emotion-focused coping styles." *Psychosomatic Medicine* 56.5 (1994): 457-462.

Nussbaum, Martha. *Upheavals of Thought.* Cambridge: Cambridge UP, 2001.

Parker, James D.A., et al. "Alexithymia in Panic Disorder and Simple Phobia: A Comparative Study." *American Journal of Psychiatry* 150.7 (1993):1105-1107.

Parker, James D.A., et al. "Alexithymia and the recognition of facial expressions of Emotion." *Psychotherapy and Psychosomatics* 59.3-4 (1993(b)):197-202.

Passer, Michael, et al. *Psychology: Frontiers and Applications.* 2nd ed. Toronto: McGraw-Hill Ryerson, 2005.

Perner, Josef and Ted Ruffman. "Infants' Insight into the Mind: How Deep?" *Science* 308.5719 (2005): 214-216.

Phelps, Elizabeth and Kevin LaBar. "Functional Neuroimaging of Emotion and Social Cognition." In *Handbook of Functional Neuroimaging of Cognition.* Ed. Roberto Cabeza and Alan Kingstone. 2nd ed. Cambridge: MIT Press, 2006. 421-453.

Philippot, Pierre, et al. "Respiratory Feedback in the Generation of Emotion." *Cognition and Emotion* 16.5 (2002): 605-627.

Pitcher, George. "Emotion." *Mind.* 74.295 (1965): 324-346.

Place, Ullin T. "Is Consciousness a Brain Process?" *British Journal of Psychology* 47.1 (1956): 44-50.

Prinz, Jesse. "Is Emotion a Form of Perception?" *Canadian Journal of Philosophy* 36 (Suppl. 32) (2006): 137-160.

_____. "Are Emotions Feelings?" *Journal of Consciousness Studies* 12.8-10 (2005): 9-25.

_____. *Gut Reactions: A Perceptual Theory of Emotions.* New York: Oxford UP, 2004.

_____. "Embodied Emotions." In *Thinking About Feeling: Contemporary Philosophers on Emotions.* Ed. Robert C. Solomon. New York: Oxford UP, 2004(b). 44-58.

_____. *Furnishing the Mind: Contents and their Perceptual Basis.* Cambridge: MIT Press, 2002.

_____. "The Duality of Content." *Philosophical Studies* 100.1 (2000): 1-34.

Prkachin, Glenda C. and Kenneth M. Prkachin. "Alexithymia and detection of facial expressions of Emotion [Abstract]." *Psychosomatic Medicine* 63.1 (2001):135-136.

Redford, Kent H. "Feeding and Food Preference in Captive and Wild Giant Anteaters (Myrmecophaga Tridactyla)." *Journal of Zoology* 205.4 (1985): 559-572.

Reisenzein, Rainer. "The Schachter Theory of Emotion: Two Decades Later." *Psychological Bulletin* 94.2 (1983): 239-264.

Roberts, Robert C. *Emotions: An Essay in Aid of Moral Psychology.* Cambridge, Cambridge UP, 2003.

_____. "Feeling's One's Emotions and Knowing Oneself." *Philosophical Studies* 77.2-3 (1995): 319-338.

Robinson, Howard. *Perception.* London: Routledge, 1994.

Robinson, Jenefer. *Deeper than Reason: Emotion and its Role in Literature, Music, and Art.* Oxford: Oxford UP, 2005.

Roedema, Thomas M. and Robert F. Simons. "Emotion-Processing Deficit in Alexithymia." *Psychophysiology* 36.3 (1999): 379-387.

Ryle, Gilbert. *The Concept of Mind.* Chicago: U of Chicago P, 1949.

Schachter, Stanley and Jermone E Singer. "Cognitive, Social, and Physiological Determinants of Emotional State." *Psychological Review* 69.5 (1962): 379-399.

Shin, Lisa, et al. "Activation of Anterior Paralimbic Structures during Guilt-Related Script-Driven Imagery." *Biological Psychiatry* 48.1 (2000): 43-50.

Sifneos, Peter. "Affect Deficit and Alexithymia." *New Trends in Experimental and Clinical Psychiatry* 10 (1994): 193-195.

_____. *Short-Term Dynamic Psychotherapy: Evaluation and Technique.* Ed. John C Nemiah. 2nd ed. New York: Plenum Medical Book Company, 1988.

_____. "Clinical Observations on Some Patients Suffering from a Variety of Psychosomatic Diseases." *Acta Medicina Psychosommatica* 7 (1967): 1-10.

Smart, John J.C. "Sensations and Brain Processes." *Philosophical Review* 68.2 (1959): 141-156.

Solomon, Robert. "On Emotions as Judgments." *American Philosophical Quarterly* 25 (1988):183-191.

_____. *The Passions: Emotions and the Meaning of Life.* Indiana: Hackett Publishing Company Inc., 1976.

Soteriou, Matthew. "The Disjunctive Theory of Perception." *Stanford Encyclopedia of Philosophy.* 10 Jul. 2009. 26 Feb. 2013. <http://plato.stanford.edu/entries/perception-disjunctive/#Int>.

Spinoza, Baruch. (1677). "Ethics." In *A Spinoza Reader.* Ed. and Trans. E. Curley. Princeton: Princeton UP, 1994.

Stout, George F. *A Manual of Psychology.* London: University Correspondence College Press, 1899.

Strack, Fritz, et al. "Inhibiting and Facilitating Conditions of the Human Smile: A Nonobtrusive Test of the Facial Feedback Hypothesis." *Journal of Personality and Social Psychology* 54.5 (1988): 768–77.

Strahan, Erin, et al. "Subliminal Priming and Persuasion: Striking While the Iron is Hot." *Journal of Experimental Social Psychology* 38.6 (2002): 556-568.

Strongman, Kenneth T. *The Psychology of Emotion.* New York: Wiley, 1973.

Sturgeon, Scott. "Visual Experience." *Proceedings of the Aristotelian Society* 98 (1998): 179-200.

Taylor, Graeme J, et al. *Disorders of Affect Regulation: Alexithymia in Medical And Psychiatric Illness.* Cambridge: Cambridge UP, 1997.

Vorst, Harrie C.M. and Bob Bermond. "Validity and reliability of the Bermond-Vorst Alexithymia Questionnaire." *Personality and Individual Difference* 30.3 (2001): 413-434.

Warfield, Ted and Stephen P. Stich. *Mental Representation: A Reader.* Oxford: Blackwell, 1994.

Warner, Richard. "Enjoyment." *Philosophical Review* 89 (1980): 507-526.

Wehmer, Francine, et al. "Alexithymia and Psychological Reactivity to Emotion-Provoking Visual Scenes." *The Journal of Nervous and Mental Disease* 183.6 (1995): 351-357.

Weiss, Anthony P. and Heckers Stephan. "Neuroimaging of Hallucinations: A Review of the Literature." In *Psychiatry Research: Neuroimaging* 92.2-3 (1999): 61-74.

Wierzbicka, Anna. *Emotions Across Languages and Cultures: Diversity and Universals.* Cambridge: Cambridge UP, 1999.

Wimmer, Heinz and Josef Perner. "Beliefs about Beliefs – Representation and Constraining Function of Wrong Beliefs in Young Children's Understanding of Deception." *Cognition* 13.1 (1983):103-128.

Winkielman, Piotr, et al. "Emotion, Behavior, and Conscious Experience: Once More without Feeling." *Emotion and Consciousness.* Ed. Lisa Barrett and Paula Niedenthal. New York: The Guilford Press, 2005.

Zajonc, Robert. "On the primacy of affect." *American Psychologist* 39.2 (1984): 117-123.

Zajonc, Robert, et al. "Feeling and Facial Efference: Implications of the Vascular Theory of Emotion." *Psychological Review* 96.3 (1989): 395-416.