Emotion and intentionally in understanding values and personal identity: A philosophical study of emotion from a humean point of view

Sunny Yang

EMOTION AND INTENTIONALITY

IN UNDERSTANDING

VALUES AND PERSONAL IDENTITY

— A Philosophical Study of Emotion from a Humean Point of View–

By
Sunny Yang

A thesis submitted for the degree of
Doctor of Philosophy of
The University of Durham

Department of Philosophy
2006

Declaration

I confirm that the thesis confirms with the prescribed word length for the degree for which I am submitting it for examination.

I confirm that no part of the material offered has previously been submitted by me for a degree in this or in any other University. If material has been generated through joint work, my independent contribution has been clearly indicated. In all other cases material from the work of others has been acknowledged and quotations and paraphrased suitability indicated.

Sunny Yang
2006

ABSTRACT

Emotion theorists in recent discussion generally take a piecemeal approach. For example, some concentrate on envy, others on regret, revenge, pity, guilt or shame. However, there is no serious attention paid to the historical and socio–cultural contexts of their topic. Therefore, I, firstly, address how emotions have been explained in the history of philosophy with particular reference to philosophical issues, for example, 'intentionality' (Chapters 1, 2, 3 and 4). Next, I emphasize the role that socio-cultural context plays in explaining emotions with particular reference to several philosophical issues, for example, 'direction of fit', 'values' and 'personal identity' (Chapters, 4, 5 and 6 respectively).

In explaining emotion, I address three kinds of crude views: the feeling-centred theory in Chapter 1, judgmentalism in Chapter 2, and a Humean functionalism in Chapter 4. Next, I demonstrate the difficulties confronting these crude views. In order to avoid such a reductionism, I suggest two alternative approaches to emotion: a qualified cognitivism and 'an embodied appraisal' theory (Chapter 3). I demonstrate that this latter view faces difficulty in explaining moral emotion. In order to meet this challenge, I provide an alternative which focuses on the socio-cultural nature of the emotions (Chapters, 5 and 6).

I argue that specific emotions are individuated and identified, firstly, with reference to other mental states, notably, perceptions, beliefs, and desires (Chapters 1, 2, 3, 4). Secondly, they are individuated and identified with reference to personal narrative and socio-cultural context (Chapters 4 and 5). Now the reason why personal narrative plays a role in individuating emotion is because emotions are often changeable, unstable, and are sometimes ambivalent (Chapter 2). However, despite their instability, emotions are sometimes long-standing (Chapter 3). They are, on the other hand, related to morality (Chapter 5). They also have perspective (Chapter 6). These characteristics of the emotions, stability, long-standing duration, and possession of perspective, help us in solving one of philosophy's most enduring problems, that is, the problem of personal identity (Chapter 6).

CONTENTS

ACKNOWLEDGMENTS

I am deeply indebted to my Supervisor, Dr. Andy Hamilton, whose many learned and imaginative suggestions I have done my best to incorporate. He provided references and discussed my drafts with sustaining optimism along the way. His intellectual guidance has left its marks on every page of this thesis.

The last two Chapters-Chapter 5and 6- were given as conference papers and I have received a wealth of comments from a variety of audiences, some of them outstandingly helpful. I am very grateful to the audience at the conference, especially, Lilii Alanen, Kathy Behrendt, and Michael Lacewing for their useful comments. I also thank Roger Squires, who suggested, in personal correspondence, some important improvements.

Several people generously made time to read the draft of this thesis, and have benefited from points raised by Peter Goldie, Roger Squires, Louise Richardson, and Francis O'Sullivan. I would especially like to thank to Peter Goldie. He gave me the useful suggestions and kindly sent me his article, which was enormously useful for the discussion in Chapter 3 and 4. Roger Squires offered me fruitful suggestions of Chapter 1, 2 and 3 of this thesis that helped me improving my own view. He brought his meticulousness to bear on the draft of this thesis which was much improved by his comments. I would particularly like to thank Louise Richardson and Francis O'Sullivan, who read all chapters of this thesis with great care. At a later stage, Francis O'Sullivan cast his eagle eye over the proofs. My thanks to them both. Among the people in Durham who were kind enough to discuss aspects of my research, I would like to thank Professor E. J. Lowe, whose zest for philosophy is infectious. I am also grateful to Professor David Cooper for his warm concern for my study and Durham life.

On a more practical note, my life has been enormously eased by my husband, who has been an unfailing support to me in various ways and gave me a constant encouragement. I hope he knows how much I appreciate his solidarity. I also want to thank to my daughter, Soyun Cha, for her cheerfulness, and also for keeping me half-way sensible. Finally, I wish to express my deepest appreciation to my mother-in-law for her unselfish support.

General Introduction

1. The Aim of This Project

What are the emotions? Are they a kind of sensation, perception or a state of consciousness of a specific feeling, a specific bodily response, or conglomeration of these various mental phenomena? In order to answer these questions I shall firstly address the problem of *what the emotions are identified with.* This problem is concerned with whether the various emotions, such as 'fear', 'pity', and 'shame' can really be regarded as a unified set of mental phenomena, i.e., an 'emotion'. There has been huge disagreement among philosophers and psychologists about what emotion cover. For example, according to A. Rorty, "the emotions do not form a natural class,"[1] since as Susan James notes, "our emotional dispositions are far more complicated and varied than systematic theorists have allowed," and therefore "we are unlikely to be able to arrive at any satisfactory equivalent of their highly general claims."[2] Yet, other philosophers argue that emotions have enough in common to form a type of unified mental phenomenon within the general field of mental phenomena. According to many of these philosophers, emotions are reducible to bodily feelings. Others hold that emotions are reducible to belief, desire or evaluative judgment. I shall call these three views reductionist. My aim in this project is to offer a strongly non-reductionist treatment of emotion.

The second problem which I shall address throughout this project is how a given emotion type, for example the mental phenomenon of fear can be distinguished from

[1] Rorty (1980), p. 1.
[2] James, S. (2003), p. 226.

another emotion, for example, pity. In other words, I attempt to illuminate the meaning of emotion and feeling by spelling out the problem of how given emotions can be *individuated.* In order to illustrate these two problems, that is, *identification and individuation*, I first of all examine extreme reductionist views: the *feeling-centred theory* in Chapter 1, *judgmentalism* in Chapter 2, and *a Humean functionalism* in Chapter 4. Second, I present the difficulties of these reductionist views.

In an effort to avoid such reductionist views, I shall then suggest two approaches to emotion. The first alternative which I shall consider in order to avoid the reductionism of the above extreme views will be hybrid theory (Chapter 2). The second alternative which I shall suggest will be a qualified cognitivism (Chapter 3). The third alternative which I shall propose will be 'non-cognitivism,' which takes emotions to be 'embodied appraisal' (Chapter 3). I shall demonstrate that these alternatives face difficulties in explaining the normativity and moral emotion. In order to meet this challenge, I shall provide an alternative which focuses on the socio-cultural aspect of the emotions (Chapters, 5 and 6).

In defence of my alternative which is non-cognitivism, I shall focus on the problem of the intentionality of emotions. Given that emotions are intentional but not reducible to mere feelings or judgments, belief or desire, a question arises: what are they, and how should we explain their intentionality? In order to answer this question, I apply the concept of direction of fit to emotions (Chapter 4). Finally, I attempt to extend the intentionality of emotions so as to include a whole variety of factors and phenomena, that is, the whole range of things and events that affect us emotionally (Chapter 5 and 6). If we use "about" in a very loose sense, we can say, as A. Baier argues, that "our

personal past is always what our emotions are about, whatever else they are, and that some special emotions or moods are directed on the whole emerging contour of that past, at all past emotions with their grounds and their fates." This is because when we are asked "what is it about her that makes you love, hate or admire her, what is it about her that disgusts, amuses, frightens... you", our answer to those question will always be to some extent biographical. If this idea is right, "the graph of her past life can be said to be what all a person's emotions are about"[3] (Chapter 6). The view which I shall defend is opposed to the Stoic and the neo-Stoic account which make emotions depend on rational, voluntary assent (Chapter 2). Rather the view which I shall support is a non-reductionist naturalist one, which takes seriously not only our animal nature but also our social condition, considering these elements as essential constituents of the natural facts that should be taken into account (Chapters 5 and 6).

2. The Scope and Limits of Discussion

An emotion is broadly understood as a mental phenomenon. Emotion as a kind of mental phenomenon has been divided into two categories by a number of philosophers who never see it the same way: an emotion, on the one hand, is a mental state, on the other, a mental disposition. As mental states, emotions are episodic, transient events. The specific emotions, for example, 'pity', 'fear', 'shame' refer to the mental episode of a person in some context, rather than to dispositions or character traits of the person. As mental dispositions, emotions are more or less persisting modifications of the mind which underlie the sequence of mental states. For example, when we say that Jones has a 'fear' of snakes until he dies, we do not say that he has been a state of feeling fear

[3] Baier (1990), p. 25.

from his birth to death. In this case we are referring to his emotional disposition or propensity. In a similar vein, we can feel jealous of a person's particular action at a certain time, but some people could be jealous of a certain person for their whole life. My aim in this project is not to focus on emotion as mental disposition in the latter sense, but on mental states at a certain time, which are nowadays called 'mental events' or states in philosophy of mind.

I shall argue that the dominant view in the philosophical study of emotion during the last thirty years has been cognitivism. There are variations of cognitivism, and I shall discuss this spectrum of views. There has been agreement that the pure cognitive theory, which identifies emotion with evaluative judgment or belief, is inadequate. I shall call the pure cognitive theory 'judgmentalism.' As we shall see, the main objection to the standard pure cognitive theory is that it cannot explain non-cognitive aspects of emotion. It has been argued that judgmentalism is opposed to the feeling theory, according to which emotions are identified with feelings of bodily change. In Chapter 2, I shall also discuss a hybrid theory, according to which emotions include both cognitive and non-cognitive constituents. The hybrid cognitivist attempts to avoid the difficulty of judgmentalism by adopting a non-cognitive element from traditional feeling theory in the analysis of the concept of emotion. Thus, the reason why we call this view the hybrid theory is that it is a combination of the traditional feeling theory and the cognitivist view. On the one hand, it draws on the traditional feeling theory in the sense that it reintroduces the element of feeling, that is, 'bodily feeling', in order to analyze the emotions. On the other hand, it follows judgmentalism in the sense that it individuates the emotions in accordance with types of judgments or beliefs.

In Chapter 2, I also raise several objections to cognitivism, specifically to a modified version of judgmentalism, and hybrid cognitivism. My main objection to the modified version of judgmentalism is the following. Although the judgementalist tries to distinguish emotional from unemotional judgment by adding a *specific additional characteristic* of the cognitive element to the combination of judgment and desire, that is, '*a focused attention*' or '*serious concern*', they cannot distinguish some emotions- for example, pity or sympathy - from unemotional judgment. For these emotions are other-regarding, and not, as the judgmentalist would claim, characterised by serious self-concern or overvaluation. On the other hand, my main objection to hybrid cognitivism concerns the inadequacy of the concepts of physiological or bodily changes, which the hybrid cognitivists utilise in order to solve the judgmentalist's difficulty.

In Chapter 3, I shall argue that both judgmentalism and hybrid cognitivism have difficulties which need to be addressed. When the judgementalist stresses the role of cognition in emotion they presuppose that cognition and bodily feeling are distinct. The hybrid theorist suggests that emotions incorporate both cognitive and non-cognitive constituents. Although the hybrid theorist admits emotions are essentially linked with bodily feelings, they also think that cognition and affect are distinct.

In order to meet these objections, in Chapter 3, I attempt to provide alternative accounts. The first position which I shall consider in order to provide such alternative accounts to both judgmentalism and hybrid cognitivism is a qualified cognitivism. I shall differentiate qualified from hybrid cognitivism in the following way. According to the qualified cognitivist, feeling is a necessary condition of emotion, and the cognitive element is a secondary addition which is not always present. However, according to

hybrid cognitivist, emotions essentially include both cognitive and non-cognitive constituents. Some qualified cognitivists have proposed componential views. Patricia Greenspan, for example, takes emotions to be compounds of affective states of comfort and discomfort and evaluative propositions.[4] Recently, Peter Goldie and Robert Roberts propose new versions of cognitive theories, defending the irreducibility of the affective element.[5] What is at issue for an adequate theory of emotion seems to be, as Goldie claims, some explication of "what holds cognition and bodily feeling together." Goldie attempts to solve this problem by claiming that *an emotion is a state incorporating both bodily feeling and cognitive states.* This leads to the following question, namely, *how do emotions bridge the gap between bodily feelings and thoughts?*

To answer the above question, that is, what holds bodily feeling and thought together in our minds when we have emotional experience, Goldie attempts to show that two kinds of feeling, that is, 'bodily feelings' and 'feelings towards,' are intimately connected in our emotional experiences. Now what are then the difference between bodily feeling and feeling towards? According to Goldie, the difference between bodily feeling and feeling towards is as follows: as far as phenomenology is concerned, bodily feelings are reflectively emotionally engaged with the condition of one's body in a certain condition, whereas feelings towards are unreflectively emotionally involved in the world beyond one's body. Now as far as intentionality is concerned, feelings towards are feelings which are directed towards the object of our emotion as such. On the other hand, Goldie writes that 'bodily feelings' have a restricted form of intentionality in that they are

[4] See Greenspan (1988).
[5] See Goldie (2000, 2002, 2004); Robert (2002).

directed only toward the condition of the body. The bodily feelings 'borrow' their intentionality from the intentional states, such as beliefs, that they occur in conjunction with. These two sorts of feelings, according to Goldie, are "'united in consciousness' in being directed towards its object: united 'body and soul,' 'heart and mind.'"[6] I shall criticize this view on the ground that there need not to be a conceptual connection other than the fact that both are needed for something to be an emotion.

I shall then show how an 'embodied appraisal' account overcomes our problem, that is, a unity requirement for a hybridist or a qualified cognitivist. The embodied appraisal theory suggests a solution to our problem as follows: The theory, on the one hand, agrees with feeling theory in the sense that emotions are embodied, and disagrees with feeling theory in the sense that judgments are needed for emotion elicitation. The theory, on the other hand, agrees with cognitive theories in the sense that emotions represent core relational themes. According to Prinz, this explains why emotions interact with thinking. If emotions represent things like danger and loss, that is, core relational themes, thoughts pertaining to those themes will be rationally tied to emotions. "Thoughts that provide evidence that one is in danger warrant fear, and fear warrants thoughts about strategies for coping with danger." [7]

I shall also discuss how the embodied appraisal theory can explain the problem of intentionality. In explaining this problem, firstly there is the claim that emotions have formal objects; secondly, the account of how emotions attain particular objects. The former can be explained in terms of the core relational themes that emotions have the function of reliably detecting. The latter can be explained in terms of the fact that

[6] Goldie (2000), p. 55.
[7] See Prinz (2003), p. 81.

emotions get linked to representations of particular objects. According to Prinz, formal objects are 'core relational themes' – dangers for fear, losses for sadness, insult for anger, and so on. In this respect, we can say that the formal object is not merely a perceptual property but the norm by which each emotion type defines its own standard of fittingness. Particular objects, according to Prinz, elicit the emotion because they can be appraised as exhibiting the core theme expressed in the emotion. If we adopt these two ideas, we can explain how bodily feelings and thoughts combine to constitute different emotions. For example, when we stand on the edge of a tall cliff, we can say that the perception of that bodily state represents danger, since "it is under the reliable causal control of dangerousness." "Danger is the property in virtue of which these highly desperate eliciting conditions have come to perturb our bodies."[8] In this case, our mental state "contains a representation of the particular object (high place) as well as a representation of the property that makes it fearful (the danger of falling)." Hence Prinz claims that the object represented by an emotion falls under a concept representing a core relational theme. A particular emotion monitors the body and detects salient phenomena such as danger, threat and loss. "Emotions are gut reactions; they use our bodies to tell how we are faring in the world."[9]

In short, according to embodied appraisal theory, emotions are "structurally simple embodied states, but they carry the kind of information that full-blown cognitions can carry." Hence Prinz argues that "cognitive theories have been right about content, and non-cognitive theories have been right about form."[10] However I shall raise an objection to this embodied appraisal theory. The typical objection to the theory is that

[8] Prinz (2004 a), p. 55.
[9] Prinz (2004 b), p. 69.
[10] Prinz (2003), p. 82.

the embodied appraisal theory cannot explain the full range of emotion we experience. It has been pointed out by cognitivists that embodies appraisal theory cannot explain disembodied emotions such as loneliness or aesthetic appreciation, and long-standing emotions such as love. Furthermore, it has been argued that the embodied appraisal theory cannot explain cases of moral emotions such as guilt, regret, and shame. If this is right, we can say that embodied appraisal theory does not generalize, across the entire spectrum of emotional phenomena.

In Chapter 5, I shall present my alternative to meet this difficulty. According to my view, long–standing or moral emotion can be explained in terms of socio-cultural context, since the minds of men are mirrors to one another, not only because they reflect each others emotions, but also because those rays of passions, sentiments and opinions may often reverberate and flow through other people. Thus we can say that the man who is proud of his wealth sustains his pride by the recognition of his worth which he finds in others. His self-esteem is reinforced by other people, when it is mirrored by them. The proud person, hence, can sustain his pride only when others let him have it – namely by their admiration and approval.

Thus understood, I will argue that our emotional reactions and responses play an important role for our moral evaluations. Furthermore, I shall argue that emotions can be changed. Yet I shall argue that it is not modified by the power of our own rational judgment or will alone, as traditional philosophers, Aristotle and Descartes, or judgmentalists like the Stoics and neo-Stoics understood, but changed only through the influence of approving or disapproving responses of those who are affected by our emotion. We pursue other's approval, since their company and judgments are important

for us, and we can change our emotions and improve them only in the course of attuning our emotional responses and behaviour reactions to theirs. Our emotional reactions and behaviour reactions are judged by others, and at the same time we evaluate these other people's judgments. These two elements are, I shall argue, essential to being a virtuous person. I shall demonstrate that this view lends support to a 'response dependent theory of emotional properties', according to which all particular objects that cause an emotion in us, for example fear, are represented as frightening only if they *merit* fear. Then, also in Chapter 5, I shall show the difficulties that this theory encounters. It may be objected that the response dependent theory of emotional properties cannot explain the normative aspect of evaluative judgments, since the theory begs the question concerning whose emotional response is normal or appropriate. As many argue, the normal isn't normative. In order to meet this objection I shall argue that our emotional responses cannot be assessed by a logical or conceptual norm, since they are evaluated in terms of their appropriateness, something which depends on a person's perspective or their historical-cultural location.

Finally, in Chapter 6, I shall show that my view is opposed to the Cartesian view, which ignores the role of emotions in their biological and social contexts, by analyzing the passions as atomistic, individual mental states, and by stressing their effects on the individual will and their expressions in the individual body. Yet I shall show the Humean view which I support places our faculties into the context of the biological and social facts of our situation. We can say that this is a turn away from Descartes' passions conceived as atomistic cognitive states of individual souls, to passions as socially sustained, socially reinforced and corrigible mental and behavioural phenomena. I shall show why we need this kind of transition in Chapter 6. According to Hume's

Treatise, the necessity of transition from Book One's solitary self to Book Two and Three's social self is concerned with his idea that the puzzle raised in Book One cannot be resolved by *thought*. Hence, setting aside the metaphysical problem with the self in Book One, Hume tries to strike a balance between thought and passion. In my view, Hume's reason for making this transition is that Hume finds himself to be located in a world of familial, social, and political relations, and not in Descartes' empty room.

'Of personal identity' in Book One of the *Treatise* suggests that the imagination by itself cannot give much substance to the sense of self. Hume has argued in Book One of the *Treatise* that the philosopher in his chamber, isolated from others, doubts his belief in his own identity. In Book One's 'Conclusion', however, he thinks that this is an unstable position. Thus, it should be noticed that in order to understand his positive conception of self and personal identity, we need to take note of his transition from the philosophical system to the vulgar one. In order to do justice to the security of the ordinary person's belief in himself, we have to turn from Book One's solipsistic world to the peopled world of Books Two and Three. Seen in this light, I shall argue that Book Two's transition from solitary reason to social passions provides a way to dissolve the intellectual puzzles of Book One concerning the self. The relation between emotion and personal identity will be discussed in the last Chapter of this thesis (Chapter 6).

Part I: Emotion and Intentionality

Chapter 1: Feeling Theory of Emotion and the Problem of Intentionality

Preliminary Remarks

In this Chapter I shall show that the traditional Cartesian-Humean view has identified emotions with bodily sensations. According to this view, emotions, on the one hand, can be understood as having similar intrinsic properties to the sensations of pleasure and pain. On the other hand, the view distinguishes emotions from bodily sensations in that emotions arise from the soul, whereas bodily sensations stem from the body. In other words, emotions unlike bodily sensations arise from the soul itself. In this respect, we can say that for the traditional Cartesian-Humean view, emotions/passions are *mental* feelings that are introspectable by analogy with bodily sensations. I shall call this Cartesian-Humean view the *traditional feeling theory* and demonstrate that the view faces difficulty in explaining the intentionality of emotion. We shall also discuss in this Chapter another version of the feeling theory of emotion which we find in William James's work. As we shall see, James also identifies emotions with bodily sensations that have a certain pattern. I shall argue that given James's conception of emotion, he also encounters difficulty in explaining the intentionality of emotion. However, I shall show in this Chapter James's position can in fact explain the intentionality of emotion by considering his later work. This will involve his view that cognition and affect are not distinct.[1] I shall then discuss the present-day Jamesian's treatment of this theme. We shall call this view the 'embodied appraisal' theory. This view will be important in order to solve one of the problems encountered in the philosophical treatment of the

[1] I shall use the term, cognition in a broad sense, which encompasses thoughts, judgments, and beliefs, whereas I mean by affect a 'mere feeling', or bodily feeling, although 'affect' is sometimes used in another sense. In psychology, affect is an emotion or experienced feeling. In abnormal psychology, affect can also refer to emotional expressiveness. Flat affect, constricted affect or restricted affect is a mismatch between experienced emotion and its expression. However, according to present-day mainstream usage in philosophy of emotion, the word 'affect' is used in a more specific way, to refer to a 'mere feeling' or bodily feeling. Throughout this thesis, I shall use the terminology in this restricted sense.

emotions in Chapter 3, namely, how do (bodily) feelings and thoughts combine to constitute different emotions? However, my objection to this account will then be: it cannot generalize across the entire spectrum of emotional phenomena, since the theory cannot satisfactorily explain disembodied emotions such as calm passions. Furthermore I shall demonstrate that the theory cannot account for cases of moral emotions such as guilt, regret, and shame.

In order to do all these jobs, in this Chapter, I focus on the object of emotion. What does 'the object' of emotion mean? The nature of the object of emotion is complex. The discussion of intentionality will lead into a more detailed classification of the objects of emotion. It is generally claimed that the emotions are intentional – they are directed towards an object: if I feel fear, then there is something, some object, which is the object of my fear. Contemporary cognitivism, which puts the emphasis on the necessity of the cognitive element having intentional content, takes the object of an emotion to be its propositional content. According to cognitivists, emotions must have propositions as their objects. That is, if I am angry, I must be angry that *p*. For example, if I am angry that Kate stole my purse, then this assertion presupposes the truth of the proposition that Kate stole my purse. I shall elaborate this view in more detail in Chapters, 2, 3 and 4. Before illustrating this view, in what follows, I shall trace the shifts in philosophy of mind which lead to intentionality becoming understood as propositional thought.

1.1. Emotion and Feeling

1.1.1. Emotion and Desire

In order to illustrate the 'feeling-centred' view of emotion, in this section, I focus on the relation between emotion and desire.[2] The association of emotion with desire has a long history, from Aristotle to contemporary philosophers. [3] Aristotle in the *Nicomachean Ethics* attempts to show the intimate relationship between emotion and desire. He classifies emotions such as fear, anger, envy, and hatred (1105b21-b23) as *pathe*, i.e., short-term states of the *appetitive* soul (1102b30-1103a3), and the most basic example of pathe is *epithumia,* desire. Aristotle's interest in the relation between emotion and desire arises in the context of his discussion of ethics. We can formulate this relation as follows:

1) Excellence of character is determined by actions and pathe (1109b30).

2) Motivation is impossible without desire (De Anima, 433a21-a29).

3) Emotions contribute to the formation of the excellence of character in virtue of their intimate link with desire.

4) Hence emotions are a subject of ethics.

This Aristotelian conception of emotion as involving desire has exerted a considerable influence on modern philosophers. According to Susan James "early modern

[2] The terminology 'feeling-centred' view has been proposed by John Deigh (1994), who attributes this view to the British empiricists and William James.
[3] Concerning the relationship between emotion and desire in contemporary discussion, see: C. C. W. Taylor (1986), 'Emotions and Wants', in Joel Marks (1986); see also Robinson (1983); and also Goldie (2000); Doering (2003).

philosophers interpret desire as an affect or passion, and classify it alongside states that are now generally described as emotions."[4] For example, Hobbes classifies the various emotions as elaborations of the basic impulses to pursue some objects (appetite, desire) and to shun others (aversion)[5] Then he uses the concepts of desire (or appetite) and aversion as the central categories to structure his theory of passions. How then for Hobbes are desires related to passions or emotions? According to Hobbes, we begin with desire and aversion, through which are internal motions, and add pleasure (delight) and pain (trouble), which the motions are overtly manifested. In addition, according to Hobbes, the emotions of 'love' and 'hatred' are basic, since when we *desire* something, we *love* it, while if we are *averse* to something, we *hate* it. Finally if we obtain something that we expected, we then have the passion of *joy*, while if we fail to obtain something we desire, we have the passion of *grief*. Now for Hobbes, on the one hand, desire and aversion are passions from which other basic passions – eight basic emotions – are elicited, and the rest of the basic emotions are understood in terms of them. On the other hand, for Hobbes, other, more detailed passions are variations of the basic passions. For instance, *appetite* with the view of acquiring something is *hope*, *aversion* with the view of hurt from something is *fear*, *desire* of office or superiority is *ambition*, and so on. Thus understood, for Hobbes, desire is "the natural disposition, the motivating force, in terms of which the rest of our passions are to be characterized."[6] Yet, for Descartes, emotions are primarily causes of desire. He writes:

---

[4] James, S. (1998), p. 22. The term emotion has vicissitude in the history of philosophy. Emotion in modern philosophy, for example, used to mean 'violent passion,' whereas present-day philosophers use 'passion' to mean 'violent emotion.' In ordinary usage, there are differences between passion and emotion. The former suggests enthusiasm, the latter mood swings and turbulence (whatever that comes to). Thus, we recognize a difference between passionate and emotional speeches; an emotional person is unstable whereas a passionate one is highly committed. Emotions may be expressed, passions demonstrated.

[5] Hobbes (1991), Chapter 6.

[6] James, Ibid., p. 28.

> The principal effect of all the passions in men is that they incite and dispose their soul to desire those things for which they prepare their body, so that the feeling of fear incites it to desire to fly, that of courage to desire to fight, and so on.[7]

When Descartes calls passions bodily sensations or sentiments, he means that "they are received into the soul in the same way as the objects of the external senses, and that they are not known by the soul any differently."[8] However, Descartes suggests that we should call these kinds of passions emotions due to their power to move us. Indeed, emotions are "more agitating and disturbing than sensory perceptions of bodily sensation" and hence they are called 'passion' in the sense that "we are often passive in the sense of being unable to control them."[9] In this respect, we can say, as Susan James notes, that "the passions are effects, but they are powerful effects which in turn move us to further thoughts and actions."[10] In the following Chapters I shall demonstrate that the contemporary cognitivists ignore this fact. Indeed, as John Deigh notes, "in the philosophical study of emotions cognitivism arose from unhappiness with affective conceptions of the phenomenon that had been a staple of British empiricism for more than two centuries."[11] Now how then, according to Descartes, do emotion and desire relate? According to Descartes, desire is one of six basic passions which is related to the future. For him, "the passion of desire is an agitation of the soul which causes the soul to wish, in the future, for the things it represents to itself as agreeable."[12] This desire can be modified by our judgments about the likelihood of realizing future states. For example, when I desire something and believe I am likely to get it, I feel hope; when I desire something and do not expect to get it, I feel anxiety; and so on. When I love

[7] Descartes, The Passions of the Soul, 40.
[8] Ibid., p. 8.
[9] James, S. (1998), pp. 21-2.
[10] Ibid., p. 21.
[11] Deigh (1994), p. 824.
[12] Descartes, Ibid., p. 86.

something, it "gives rise to the desire to get into the state that will enable me to maintain my love." When I hate something, "hatred gives rise to a desire to get away from the object of the hatred, and so forth."[13] In this picture, we can see that desire is conceived of as one of passions and that it is embedded in the sequence of passions. Thus Descartes argues that "the passions cannot lead us to perform any action except by means of the desire they produce."[14] If this is so, as Susan James points out, "desire is always a necessary condition of action." Then a question arises: is desire also sufficient for action? It is difficult to answer this question. However, Descartes might say that in the case of involuntary action, a desire is sufficient for action. For example, as S. James puts it, "when a particularly strong movement of the animal spirits constitutes a particularly strong desire, it can cause motions of the limbs without further ado." Now in the case of voluntary actions, on the other hand, for Descartes, desire may or may not be sufficient for action; for example, when desire and volition conflict, the cause of action can be said to be the relative strength of the passion and volition concerned, whereas when the desire and volition coincide, both are causes of the action. However, in general, according to Descartes, we act on desires when the will has assented to them.[15] So far we have shown that for Descartes desires are embedded in the sequences of passions, and thus are part of the complex of causes of action. For the sake of simplicity, I take this idea to be that emotion implies desire. In what follows, I shall show that this idea, that is, emotion implies desire, was fully developed in Hume's theory of passion.

What then does desire mean for Hume? Hume claims that a desire does not have "any

---

[13] c.f. James, S. Ibid., pp. 24-5.
[14] Descartes, *The Passions of the Soul*, 144.
[15] As argued by Susan James. Ibid., 25.

representative quality which renders it a copy of any existence or modification",[16] whereas a belief has a representative quality. There are, for Hume, two categories of desires. According to his official theory of desire, desires are known by their phenomenology. Hume contrasts desire with aversion thus; "*Desire* arises from good considered simply, *Aversion* is deriv'd from evil."[17] In other words, in virtue of desire "the mind...tends to unite itself with the good" (i.e., pleasure), whereas in virtue of aversion it tends "to avoid the evil" (i.e., pain)[18] He classifies this narrow notion of desire and aversion with the so-called 'direct passions.' The reason why he calls desire and aversion 'direct passions' is that they "arise immediately from good or evil, from pain or pleasure." Examples of these passions are 'desire and aversion', 'grief and joy, hope and fear.'[19] Here what we should notice is that other direct passions, for example, joy, grief, hope and fear have the same characteristics as desire and aversion in that they "arise from a natural impulse and instinct which is perfectly unaccountable"; "these passions strictly speaking produce good and evil, and proceed not from them like other affections,"[20]

Now the question arises as to the relation between emotions and desires. Hume takes the term 'emotion' to mean a bodily disturbance, like some pleasures and pains. To that extent he seems to regard it as an 'impression of sensation' rather than as an 'impression of reflexion.' In this respect, recent writers who discuss emotion regard Hume's view as 'the most notable exponent' of feeling theory concerning the nature of emotion.[21] Hume, like the seventeenth-century philosophers, regards desires as a passion, and classifies it

[16] Hume, *A Treatise of Human Nature* (hereafter T.), T. 416.
[17] T. 439.
[18] T. 438.
[19] T. 438.
[20] T. 439.
[21] See Lyons (1980), p. 8; Alston (1980), p. 480.

with states that we now call emotions, such as joy, grief, hope and fear. Let's consider some examples in order to understand how such passions arise *directly* from the feelings of pleasure and pain: I enjoy the taste of ice-cream after dinner, and so, I want some ice-cream now. The death of my loved one is painful, and so, I feel deep grief on hearing that he has died. The winning of the team which I support is a pleasant prospect, and so, I feel joy at the news when they in fact do win. The victory of the team which I support was pleasing to me; so, I hope for more of the same with the next match. But the news that the best player in the team has been injured is unpleasant, and so, I fear that my team will lose in their next match.

But one might say that these characteristics of desire cannot explain reasons for action, since Hume's narrow notion of desire suggests that it has what is sometimes called 'qualitative' aspects or 'phenomenal senses' like perceptions and bodily sensations.[22] Like bodily sensation and perception, desires which have 'qualitative' or 'phenomenological' content, can also be defined by how they *seem* to the subject. To put it another way, for Hume, to have desire is similar to what it is like to experience a sensation. Yet Hume holds that desire differs from impressions of sensation or belief in that it has no representational content. For example, visual experience represents the world as being a certain way, but desire which immediately arises from pleasure has no such representational content. Thus it can be said that for Hume the differences between desires and other psychological states can be found in the precise characteristics of their qualitative content. We can say that this characteristic of desire is similar to what John Locke called 'an *uneasiness,*' which he describes as "all the pain of the body, of what so

[22] The 'qualitative' characteristics of any mental state such as belief, desire and emotion are well addressed by McGinn (1982), Ch. 1; John Bricke also addresses this subject well in the context of Hume on the basis of McGinn's distinction. See Bricke (1996), pp 35-48; See also Hacker (2002).

ever, and disquiet of the mind...".[23] According to Locke, being in this state itself is a reason to do what will relieve that discomfiture.[24] 'Uneasiness' appears first in the *Essay* as a synonym for pain. Sensation presents the simple ideas of pleasure or delight and pain or uneasiness. On the basis of this idea, Locke tries to define our passions. For example, "desire is the uneasiness that a person feels upon the absence of a thing whose present enjoyment carries the idea of delight. Hatred is what we feel for things that are apt to produce uneasiness, sorrow is uneasiness of the mind upon the thought of a good lost, and so forth"[25] What is then the relation between desire as uneasiness and the emotion for Locke? In order to understand the relation, we need to look at his view of the relationship between desire and the will. According to Locke, the uneasiness of the state of mind determines the will. Here what Locke wants to suggest is, as E. J. Lowe points out, that "the immediate cause of one's exercising one's will in a particular way is the prevailing balance of one's desire."[26] On the other hand, according to Locke, the will is determined by the various passions that are modifications of uneasiness. In this respect we can say that Locke's view differs from the traditional Aristotelian view, according to which the will determines or controls the passions. He applies his modified view of the will to explain reason for action. He writes: 'the most...urgent *uneasiness*, we at that time feel, is that which ordinarily determines the *will*.'[27] If this is so, then it can be said that on the one hand, the will is not an immediate motive for our action, but

---

[23] Essay II. xxi. 29.

[24]This characterization of desire, as Schiffer (1979) notes, is an acknowledgment of the 'reason-providing' character of those desires. 'Reason-providing' desires have, according to Schiffer, the following two features: 'an r-p desire to φ is a desire to φ to relieve the discomfort of that desire, a desire to φ for the pleasure of its own relief.' If this is so, it can be said that r-p desires are states that have a certain uncomfortable phenomenological character and are themselves desires for the pleasurable relief of that very discomfort. If r-p desires have these phenomenological characteristics, then we can say that their satisfaction will be pleasurable. See Schiffer (1979), p. 199.

[25] Locke, Essay, II. xx. 6. Quoting from this text I was helped by Susan James(1998), p. 29.

[26] Lowe (1995), p. 134.

[27] *Essay,* II. xxi. 40.

is determined by the uneasiness. On the other hand, Locke says that the will is determined by other passions, since the uneasiness sometimes takes the form of passions. Yet for Locke, like Hobbes, not every passion has motivational force, since the passion that plays a central role in determining the will is desire. That is, "an uneasiness in the absence of a thing whose present enjoyment carries the idea of delight."[28] According to Locke, "the will seldom orders any action, nor is there any voluntary action performed, without some desire accompanying it".[29] Now the reason why the will is determined by passions is because the passions imply desire and contain the uneasiness that determines the will. In order to appreciate this point, let us consider the following example: one fears an approaching lion and tries to escape from it. The person seeing a lion on his path, first, undergoes bodily changes such as variations in his heartbeat, bloodflow and muscular tensions, causing his legs to tremble; second, the *uneasiness* of these physiological changes feed back into the emotional reaction at the mental level, causing the passion proper (his fear): third, the state of fear which contains the uneasiness causes one to exercise one's will in a particular way; he tries to run away to avoid the situation. Now if this is true, we can say that we cannot feel emotion and perform any action unless we feel uneasiness. Following Hume's and Locke's lead, some people argue that desires have *per se* authority.[30] D. Stampe argues that desire's possession of qualitative content can move us to act and give a reason for action. He writes:

> A pain is a reason to do what will relive that pain, a reason in itself, not through its object - for pains, and itches, *have* no objects, no propositional or other content.

[28] Ibid., II. xxxi. 33.
[29] Ibid., xxi. 39.
[30] See Stampe (1987), pp. 335-81.
On the other hand, according to Dennis W. Stampe, the belief that p is not in itself a reason to act so as to bring about p. (p. 342).

> Itches make it reasonable to scratch and the rationality of scratching is not derived from the rationality of itching. Such sensory states are not themselves either reasonable or unreasonable. They too are the origin of rationality.[31]

Now it can be said that Hume, on the other hand, has a broader conception of desire which encompass the narrow notion, 'direct passion.' The broader conception of desire involves what Hume calls 'calm' and 'violent' and 'indirect' desires that lack phenomenological content, for they "are more known by their effects than by their immediate sensation."[32] The distinction between 'calm' and 'violent' passions rests on how emotions feel to us. It can be said that calm passions are felt in a 'soft' manner, whereas violent ones come over us with a greater force and intensity.[33] Indirect passions arise from good or evil plus other qualities or ideas. That is, with the indirect passions, 'the conjunction of other qualities' [34]is needed. The other quality is the relation of the cause to the object.[35] With this broader conception of desire, he wants to accommodate emotions such as benevolence, anger, pity and malice as well as desire more narrowly conceived. Benevolence, for example, is "a desire of the happiness of the person belov'd, and an aversion to his misery"[36]; anger is "a desire of the misery of the person hated, and aversion to his happiness".[37] Pity is "a desire of happiness to another, and aversion to his misery; as malice is the contrary appetite."[38] In this spirit, we can say that emotions involve desire in the sense in which a sentence that asserts an emotion

---

[31] Stampe (1987), p. 348.
[32] Hume, T. 417.
33 A further discussion of the distinction between 'calm' and 'violent' passions in Hume's theory of passion, see Chapter 3 of this thesis. Also see a wonderful analysis of Hume's calm and violent passions in Baier (1991), pp. 167-71.
[34] T. 276.
[35] In Book 2 of the *Treatise* Hume takes up the indirect passions before the direct passions on the ground that the distinction between object and cause is easier to discern. An example and model of Hume's analysis can be found in his discussion of the four indirect passions: pride, humility, love and hatred.
[36] T. 382.
[37] T. 382.
[38] T. 382.

entails that the subject is related to desire in some specific way. Many contemporary philosophers who are committed to the belief-desire theory of action explanation seem to agree with this idea that emotion implies desire. According to the belief-desire action explanation, an emotion can motivate our action only if it implies desire.[39] I shall discuss the validity of this account in Chapter 4 ('Do emotions have Direction of Fit?') Before looking at this issue, let us look at the difficulty of Hume's theory of desire.

Although Hume presents the broad conception of desire as encompassing emotions that lack phenomenological content, that is, emotions that have propositional content, he gets into trouble again by attributing the motivational force of desire to its natural instinct – pleasure or pain. If it is assumed that the motivational force of desire is in nature, it might be thought that it cannot explain reasons for action since, as many claim, the idea of reasons for action implies that desire *rationalizes* action and intention. In this respect, Stampe remarks that "to recognize the natural force of desire is to refuse *the rational* authority of desire."[40]

Having established Hume's view of desire in this way, we can say that Hume seems to deny the intentionality of desire. Now in order for desire to have intentionality it must represent a state of affairs as being its object. If so, it might be thought that in order for desire to have adequate authority it must have propositional content. In this sprit, Michael Smith revises Hume's account of desire. He takes desires as having propositional content and a certain functional role, for "the propositional content of a desire may simply be determined by its functional role."[41] Now those who take desires

[39] E.g., see Smith (1994), Ch. 4; see also Goldie (2000), Ch. 5.
[40] Stampe, Ibid., p. 349.
[41] Smith (1994), pp. 113-4.

as having propositional content argue that all ascriptions of desire can be described in the same way: *A desires* that *p* (where A is subject and p is a sentence). According to this view, the sentence *p* that follows "that" is the object of desire. They claim that the object of desire like that of belief can be expressed in propositional form, or by means of a sentence. This is why we sometimes call both beliefs and desires propositional attitudes. It is said that the advantage of this view is that 1) it allows for the intentionality of desire by making the object of that desire explicit; 2) it shows us a way of representing the appropriateness of the propositional content.

We have seen so far that for the traditional feeling-centred view, desires which have phenomenological qualities need not have intentionality. But if the view is committed to the claim that emotion implies desire, it follows that emotion could not have intentionality either. Hence the contemporary attempts to revise Hume's view depart from him by incorporating propositional content into the account of intentionality. I shall return to the issue in Chapter 4('Do emotions have Direction of Fit?') Before looking at this issue, I shall focus in what follows on the contemporary successor of the traditional feeling-centred theory of emotion.

1.1.2. 'Feeling-Centred View' and Intentionality

I have argued so far that the feeling-centred conception encounters difficulties. The main criticism of this view concerns intentionality. Intentionality refers to the aboutness of mental states, for example, my belief about John, my wish about tomorrow. When we are angry or afraid, for example, we are angry at someone or something, afraid of someone or something. This someone, this something, is the intentional object of

emotion, that at or toward which it is directed. Yet it is said that bodily sensations of pleasure and pain, which Descartes and the British empiricists identified with emotion, are not directed at or toward anyone or anything, and hence they are not intentional states. If this is right, then it can be said that the feeling-centred view doesn't capture the intentionality of emotion as it identifies emotion with bodily feeling. However I shall show that there is another kind of feeling theory which incorporates affect into intentionality. This picture is developed in William James's work, since James attempts to reconceptualise intentionality by reuniting cognition and affect.

Following Hume's lead, William James adopts another version of feeling theory. It is commonplace in the literature for James's view to be called a feeling theory, and to be offered as a standard example of a non-cognitive view. In a remarkable passage of his *Principles of Psychology* William James writes:

> My theory... is that bodily changes follow directly the perceptions of the exciting fact, and our feeling of the same changes as they occur is the emotion. Common-sense says, we lose our fortune, are sorry and weep; we meet a bear, are frightened and run; we are insulted by a rival, are angry and strike. The hypothesis here to be defended says that this order of sequence is incorrect, that the bodily manifestations must first be interposed between, and that the most rational statement is that 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.[42]

According to James, a certain feeling which is intrinsic to the various emotions can be identified with bodily sensation. In other words, for him an emotion is simply the complexity of bodily sensation having a certain pattern. Following Darwin, who observes that there are correlations between emotional states and changes in the body,

[42] James W. (1890/1950), Vol. 2, pp. 449-50.

James views fear, for example, as symptoms which include widely opened eyes and mouth, raised eyebrows, dilated nostrils, stiff posture, motionlessness, a racing heart, increased blood flow through the body, pallor of the skin, cold perspiration, shivering and trembling, hurried breathing, dry mouth, faltering voice and so on.[43] In order to appreciate this point, let us look at Darwin's and the Darwinians' view of the basic emotions.

1.1.2.1. The Traditional Feeling Theory and the Darwinian Theory: Basic Emotion Theorists

Basic emotion theorists assume that there are a privileged set of emotions from which all other emotions are derived. The discussion of basic, or primitive emotion can be traced back to Descartes, Hobbes, and Hume, which I have called the traditional feeling theorists. As we have seen in section 1, for Hobbes and Hume, it is our natural disposition that we pursue good or pleasure and avoid pain and trouble. Moreover, for Hume we have a motive to move toward or avoid some object due to direct passions arising from immediate responses toward pleasure and pain. Thus understood, on the one hand, for Hume basic emotions (direct passion) are the motive of our action. On the other hand, for Hobbes, a set of basic emotions or primitive passions- that is pleasure, pain, love, hate, joy, and grief - are elicited from desire and aversion. For instance, *appetite* with the view of acquiring something is *hope*, *aversion* with the view of hurt from something is *fear*, *desire* of office or superiority is *ambition*, and so on. Thus understood, we can say that for Hobbes desire and aversion no longer belong to a set of basic or primitive emotions. Rather, as Susan James observes, for Hobbes, desire "is the

[43] James, W. (1890/1950), 447ff.

natural disposition, the motivating force, in terms of which the rest of our passions are to be characterized."[44] Thus, we can say that although Hobbes and Hume are in the same position when they say that desires are natural dispositions, or the motivating forces, they are differ in the sense that the one holds that desires and aversions do not belong to basic emotions, whereas the other holds that they do.

Yet for Descartes, his posited six primitive passions are not natural dispositions and motivating forces, but are passions from which all other passions are derived. Descartes, for example, argues that six emotions are 'primitive'; joy, sadness, desire, love, hatred, and wonder.[45] In classifying passions, Descartes, first of all, tries to find the simplest and the most primitive ones, and suggests that we should deduce more complex and derivative ones from the primitive. The above six primitive passions are basic, and therefore are not be reducible to other more primitive passions. Furthermore, according to Descartes, each primitive passion has a corresponding subclass of passions. Although there is no hierarchy among these six primitive passions, "an enumeration of the passions requires only an orderly examination of all the various ways having importance for us in which our senses can be stimulated by their objects."[46] According to Descartes, the passion which comes first in enumeration is wonder; secondly, come to love and hatred; desire, joy and grief come next. He presents this enumeration in terms of three criteria: novel, beneficial to us, and time.

Wonder, he claims, is the 'novel' passion. He means by 'novel' that when something hitherto unknown to us appears, this causes our soul to wonder. Considering the fact

[44] James, S. (1998), p. 28.
[45] Descartes, *The Passions of the Soul.*
[46] Ibid., 52.

that for Descartes all passions cause a change in the soul, wonder is the most important passion for him. The reason why the passion of wonder is regarded as the novel one is because the passion of wonder occurs before we know whether the thing is beneficial to us or harmful; second, the passion has no opposite one to itself. Descartes claims that according to the 'novel' criterion, wonder is the first of all the passions.

The second criterion of enumeration of the primitive passions concerns whether or not something is beneficial to us or harmful. The passions arise not due to objective qualities of the objects themselves, but due to their advantageous, or not, to us. The object which is represented to us as advantageous produces love, that which isn't produces hatred. Love and hatred are rather complex, compared to wonder, since they are related to time, for example.

The third criterion for enumeration of the primitive passion is time. In order to put desire, joy and grief in order, Descartes takes time into account. Descartes locates desire in the first order among these three passions, since "they lead us to look much more to the future than to the present or the past."[47] The passions, joy and grief, which are related to the present, follow the passion of desire in the order of the enumeration. Descartes explains other passions on the basis of these six primitive passions. For instance, according to him, *anxiety* is defined as the belief that there is only a small probability that one will have what one desires[48]; *jealousy* is defined as an anxiety relating to the desire to preserve something that one values.[49] In what follows, I shall demonstrate that contemporary feeling theorists, William James and Prinz, for example,

[47] Descartes, *The Passions of the Soul*, 57.
[48] Descartes, Ibid., 58.
[49] Ibid., 167.

do not follow the traditional feeling theorists. Furthermore, I shall argue that they adopt Darwin's view when they take up the basic emotion theory.

Darwin was also interested in primitive or basic emotions. Yet Darwin did not follow the traditional feeling theorist's lead in that his main concern was the involuntary facial expressions which are common to all human beings. In his book *Expression of Emotion in Man and Animals*, Darwin argues that if expressions are innate, then we can say that emotions are also innate.[50] In order to lend support to this idea, he tries to prove that human emotional expressions are innate, and attempts to find evidence for similar facial expression among people who live in isolated cultures. Darwin asked Englishmen who lived in remote regions of the globe to answer questions about the facial expressions they observed. According to their reports, the similarity between the facial expressions of Englishmen who had been isolated from the West and those remaining in England was obvious. Following Darwin's lead, the present-day Darwinians, for example Paul Ekman and Carroll Izard, believe that emotions are partially constituted by facial expressions, and they regard the universality of facial expressions as direct evidence for the universality of emotions.[51] It has been argued that Darwin was influenced by Charles Bell's work[52] which addressed the unique musculature of the human face. Bell claimed that the musculature of the human face is given by God's design for the purpose of expressing the emotions distinctive of human beings.[53] Yet, unlike Bell, Darwin developed this idea to explain the superiority of evolutionary theory.[54] According to Darwin, the musculature of human face is not given by God's design, but

[50] Darwin (1889/1988).
[51] Ekman (1972); Izard (1971).
[52] c.f., *Anatomy and Philosophy of Expression.*
[53] I owe this point to Deigh (2004).
[54] Darwin, Ibid., 7.

is inherited from our ancestors, who had themselves acquired it as a result of its adaptive utility, and so it got passed on to future generations. In this respect, Darwin claims that some expressions have an adaptive function. For example, the relationship between surprise and raised eye-brows was not random, because the movement was merited in the environments that usually motivated surprise in prehistoric circumstances. Hence he holds that the muscular movements of the human face are the products of inherited dispositions.

The trouble with this view lies in explaining other cases which are not the products of inherited dispositions but are involuntary expressions that are merely conventional. Confronted with this difficulty, Darwin claims that not every expressive emotion is the product of inherited dispositions. He tries to distinguish the former from the latter. The latter cases, for example, curling one's lips as an expression of scorn, according to Darwin, tend to be learned in childhood and to be different across cultures. Darwin does not seem to discuss this kind of emotional response further. I shall present an alternative account for this position in Chapters 3 and 5, where I shall discuss the point that our emotional expressions which are conventional and different across cultures involve 'response–dependent emotional properties.'

Following Darwin's lead, contemporary Darwinians focus on basic human emotions, each of which is defined by facial movements that qualify as the true facial expressions of all emotions that belong exclusively to that category. More recent authors have developed this idea with six basic emotions (anger, disgust, fear, joy, sadness and surprise) and claim that each of these corresponds to a unique bodily pattern.[55]

[55] See Levensen, Ekman and Friesen (1991).

According to contemporary Darwinians, the true expression of an emotion is manifested by facial and overt bodily movements. Basic emotions theorists differ from each other in that they include different items on their lists. For example, Izard offers the following list: anger, contempt, disgust, distress, fear, guilt, interest, joy, shame, and surprise.[56] In recent works, Ekman develops his initial list of six affect programs into fifteen basic emotions: amusement, anger, contempt, contentment, disgust, embarrassment, excitement, fear, guilt, and pride in achievement, relief, sadness, satisfaction, sensory pleasure, and shame.[57] Basic emotions theorists also differ in choosing criteria for the inclusion of particular emotions. Some people claim that basic emotions are psychologically primitive, and others holds that they are biologically primitive. Contemporary Darwinians take biological criteria as primary. The main idea that biological criteria theorists defend is that basic emotions are innate or present in all normally developing members of the species. They are evolved patterns of response. Although there is disagreement among Darwinians about how many basic categories there are and what emotions belong to them, they usually include basic categories such as *joy, anger, disgust, surprise, fear, distress, and sadness.* According to them each of these is understood to include a range of cognate emotions. The category of joy covers happiness, delight, gladness, satisfaction, and so forth; that of anger covers annoyance, indignation, rage, resentment, and so forth.

Now a question arises: how can we identify basic emotions? In order to answer this question, Ekman, for example, suggests that we need to look for culturally universal facial expressions. If the same expressions are elicited by the same conditions across

[56] Izard (1971).
[57] Ekman (1992).

cultures, there is reason to believe that the response is mediated, at least, in part, by an innate emotion. But it seems to me, as Prinz points out, this kind of evidence is neither necessary nor sufficient for basicness.[58] It is not necessary, because, as Ekman admits, some innate emotions may lack corresponding facial expressions.[59] It is not sufficient, because some innate emotions may be nonbasic. In this respect, Ortony and Turner, who cast doubt on the inference from facial expressions to basic emotions, claim that if basic emotion theorists present different lists, then it seems that such lists are arbitrary, in that there is no objective criterion or principle to justify the different choices.[60] Now it seems to me, that the lack of agreement among basic emotion theorists concerning basic emotions, for example, anger, shows us that basic emotion lists must be open to revision. I do not discuss basic emotion further. Instead, in what follows, I shall explain how contra Darwin, many of those influenced by him take a reductionist approach to the emotions, in that they attempt to reduce them to physiology.

I have shown so far that the Darwinians develop Darwin's idea of the involuntary facial expression into the basic emotion theory. Yet the Darwinians conceive of an emotion as a neurophysiological event whose manifestations typically include the facial and overt bodily movements that are the emotion's true expressions.[61] Now when the Darwinians divide emotions into basic categories they think these categories specify differences among the emotions. Furthermore, they claim that the distinction between joy and sadness, for example is a distinction between the neurophysiological mechanisms that produce the facial movements defining those categories.

[58] Prinz (2004b).
[59] Ekman (1999a).
[60] Ortony and Turner (1990).
[61] See Izard (1977), 4; Ekman (1980), pp. 79-82.

Now several questions arise: given this conception of emotion, first, how can the Darwinians explain long-standing emotion such as 'love' which does not appear to involve physiological mechanisms? Second, how can the Darwinians explain moral emotions, such as guilt, or regret? Third, how can the Darwinians accommodate the intentionality of emotion? I shall discuss these issues in the following sections of this Chapter. Provisionally, we can note that they may encounter the same difficulty in explaining the intentionality of emotions as the feeling theorists did. I shall present my alternative which solves these kinds of difficulties in Chapter 3 and 5. Before looking at this problem, let us consider other difficulties that James and the Jamesian theorists might face.

1.1.2.2. Some emotions do not involve bodily change.

The first difficulty which James's theory encounters is a conceptual one: it does not generalize, since some emotions are not associated with bodily changes. If this is true, it follows that James's theory implies that someone who cannot feel the relevant bodily changes will have no emotions. In order to appreciate this point, let us consider the following example. Suppose that someone is in hospital after a terrible car accident. Due to the accident, he cannot move his body. Fortunately, his consciousness remains. In this case, presumably, the person may not feel physiological changes, such as a tightening in the stomach and sweating. Thus he cannot feel bodily sensation. If James were right, the patient should not feel any emotion. But it seems to me that this is unlikely to happen. He may feel *joy* because he is alive. He may feel *guilt* for his carelessness. He may *regret* that he didn't drive more carefully. He may feel fear that he will soon die. How then can James explain the fact that these emotions are able to arise

in the absence of bodily change? It has been argued that the retention of emotions in subjects with severe spinal cord lesions, which prevent all neural communication between brain and body shows us James's theory is false.

In reply to this kind of objection, more recent writers, for example Damasio, suggest several points against the evidence from spinal cord lesions. First, information concerning many of the relevant physiological changes could be transmitted to the brain via nerves such as the vagus, which enters the brain stem at a level too high to be impaired by survivable spinal cord damage. Second, hormonal changes will still be communicated to the brain via the bloodstream. Third, some diminution of emotion is indeed found in many patients with spinal cord lesions. Damasio then concludes that emotional responses may be able to bypass the body by means of an 'as-if loop.'[62] If an emotion is a perception of a bodily change, then the very brain state that underlies that perception must be able to arise in the absence of a bodily change, acting as if the body had changed. But some people criticize the idea of an 'as if loop' on the grounds that it is ad hoc to assume that spinal patients simply hallucinate bodily changes when they are in an emotional state. Confronted with this objection, Jesse Prinz, who is the contemporary successor of William James, defends Damasio's ideas by claiming that "if emotions are evolved from reflexive bodily response, the brain may have adapted a way of anticipating bodily movements before they happened."[63] Even if we accept Prinz's suggestion, another question arises: is there a bodily correlate of *guilt* for example? Here Prinz's answer is as follows:

> [...] Perhaps guilt has a face, with downcast eyes and lowered chin. There may

[62] Damasio (2000), p. 289.
[63] Prinz (2004a), p. 48.

> even be a blush of guilt, borrowed from the more primitive emotion of embarrassment. [...] James's cerebral emotions may also have bodily concomitants. Moral passions are widely believed to prod us into action, and aesthetic response can send tingles down our spines. Intellectual emotions can overlap with surprise or delight and almost certainly have a somatic mark.[64]

Now even if we agree with James and the Jamesian theorists, we can raise the following objection: James cannot explain 'long term' emotions. One can be in love for a long time, even though one's body isn't in an enduring state of perturbation. Faced with this kind of objection, James seems to concede to the view that emotions cannot be reducible the basic emotions, and hence cannot form natural kinds. He writes:

> I should say first of all that the only emotions I propose expressly to consider here are those that have a distinct bodily expression. That there are feelings of pleasure and displeasure, of interest and excitement, bound up with mental operations, but having no obvious bodily expression for their consequences, would, I suppose, be held true by most readers.[65]

However James's remark the above does not mean that he retracts his initial claim[66] that emotions are bodily feelings, rather it implies that feeling and cognition cannot be distinguished. James writes:

> In all cases of intellectual or moral rapture we find that, unless there be coupled a bodily reverberation of some kind with the mere thought of the object and cognition of its quality; unless we actually laugh at the neatness of the demonstration or witticism; unless we thrill at the case of justice, or tingle at the act of magnanimity; our state of mind can hardly be called emotional at all.[67]

If this is right, the way in which James can explain the case of long-standing emotion - for example, someone who claimed to be in love but never showed signs of somatic

[64] Prinz, Ibid., 50.
[65] James W. (1884), p. 189.
[66] Ibid.
[67] James W.(1890/1950), pp. 470-1.

response – would be as follows. They *dispose* us to enter into patterned bodily responses, since as Prinz notes, "we do not say that these emotions disappear when they are unfelt, because the disposition is there all the time."[68] In defence of William James, Prinz draws a distinction between dispositional and occurrent states and he takes long-standing emotions to be dispositions. If I say that I love my husband all the time, my love in this case, is an enduring disposition to have occurrent states of love. An occurrent state of love, according to Prinz, "is an embodied reaction of the kind one has when one encounters the object of one's love."[69] The reason why this state is embodied is because being in such an occurrent state shows that "brain activation in limb areas associated with bodily response when subjects viewed pictures of their lovers."[70] In this respect, Prinz suggests that we should take long-standing love to be such an embodied state, since when I say that I love my husband without experiencing cuddly tenderness in relation to him, for example, it seems to be nonsense. Although Prinz's the embodied appraisal theory explains long-standing emotion by distinguishing dispositional and occurrent states, one might still doubt whether the theory can explain calm passions, such as aesthetic appreciation or moral emotions. Confronted with this objection, Prinz appeals to the basic emotion theory, according to which some emotions are basic and others blend these basic emotions. For example, contempt may be anger plus disgust; thrills may be joy plus fear; guilt may be sadness plus thoughts about my transgressions. Obviously some 'basic' emotions – fear, hate, anger, wonder – are fundamental as an equipment of survival in a vulnerable species. However, the history of philosophy shows that the examples of the basic emotions do not remain constant; for example, as

[68] Prinz (2004a), p. 50.
[69] Prinz (2003), p.83.
[70] Ibid., p.83.

A. Rorty puts it, Aristotle considers wonder (*thauma*) to be a basic activity of the mind; for Descartes, admiration is the first of the passions[71]; for Hume, as we have seen in the previous section, desire and aversion are basic passions which arise immediately from good or evil, from pain or pleasure. If this is true, as A. Rorty points out, "beyond those basic attitudes, the focus of phenomenological and analytic attention shifts with changes in primary social, political, and economic activities."[72] Hence I suggest that in order to explain moral emotions, we should take account of them in the socio-cultural, political, and economic activities. In Chapter 5, I shall discuss this issue with regards to socio-cultural context, in terms of the 'adaptive syndrome theory.'

1.2.3. How can emotions be individuated?

Another objection to James's theory of emotion can be cashed out in terms of the following questions: what is the difference between the emotion of anger and irritation, envy and jealousy, shame and embarrassment? Now the traditional feeling theorist, for example Hume, might say that the difference between getting angry and getting irritated is that they are felt differently. But the problem is whether the pure feeling of both inwardly getting angry and inwardly getting irritated are fundamentally different or not. One might, for example, argue with William James that the difference between envy and jealousy is whether particular bodily sensation such as a feeling of weight in the stomach presents itself or not. Another might argue that the pure feeling of both is the same in virtue of the feeling of pain, but it is different in virtue of the object of the feeling. The question now arises: do we have to take a pill to cure the pain caused by extreme jealousy? In a similar way, when we are embarrassed, we can feel that our face

[71] Rorty (2004), p. 276.
[72] Ibid.

becomes red and sweats. But don't we feel this bodily change in the case of flu? If this is right, how can we individuate feelings in terms of the criteria of bodily changes? In this respect one might say the feeling theorists are in a difficult position when it comes to deciding the question of whether embarrassment and shame are different in terms of phenomenology. Suppose I cannot be aware of the difference in the element of pure feeling between getting irritated and getting angry. This objection amounts to the claim that I am only aware of the qualitatively identical element of feeling 'F' in both cases. But according to Hume's view, what I am aware of in both cases ('getting irritated and getting angry) is either 'H' or 'G.' If this is right, it is clear that the pure element of feeling itself cannot be 'immediate, self-evident and incorrigible' to allow us to be aware of it, let alone to solve the problem as to whether the element of pure feeling inherent in different emotions is individuated as the same element (F) or different elements (H or G). If this is right, the Humean view that the difference between getting angry and getting irritated lies in the difference of pure feeling is not available. Since the Humean view attempts to solve the problem of the individuation of emotions by appealing to 'feeling', it should show that the element of pure feeling cannot have distinct epistemological characteristics. In this respect many argue that bodily feelings are not sufficient to individuate the full complexity of emotional experiences and that an element of cognitive evaluation or judgment is also needed.

The Jamesian might reply to this objection as follows: first, while embarrassment and shame are distinct, they are also closely related. Hence we can say that the latter is a species of the former. Shame is something like embarrassment at misconduct. Jealousy is a combination of anger, sadness, fear, and disgust. Envy is a species of jealousy. Now the reason why these emotions share the same bodily states is because these emotions

belong to a common family.[73] Qualifying this reply as insufficient, Prinz presents another response. According to Prinz, emotions can be individuated by their reliable elicitors. He claims that "emotions can be distinguished by their eliciting conditions."[74] Yet Prinz maintains that this is possible in terms of perceptions of bodily states as the vehicles of emotion. Prinz observes that James's theory can make some headway on this account, since James stresses the conditions under which emotions are elicited.[75] If this is right, we can say that embarrassment and shame feel alike, as perceptions of bodily state, but that shame is caused by misconduct, and many cases of embarrassment are not. Although Prinz updates and revises James's account in order to defend it, James himself did not make clear the problem of the individuation of emotions. In Chapter 3, I shall discuss Prinz's updated version of William James's account concerning the problem of individuation in more detail.

1.2.4. How can James accommodate the intentionality of emotions?

Now the most common objection to James's theory in philosophical studies of emotion is that it cannot explain the fact that emotions have intentional content. In what follows I shall argue that James does not in fact overlook the intentionality of emotion. As I have said before, on the one hand, for him an emotion is simply the complexity of bodily sensation having a certain pattern. On the other hand, James sometimes implies that emotions are always consciously felt. For example, James believes that emotion is constituted by certain feelings that are aroused by the thought or, as he says, perception of an exciting object. The object excites in that perceiving it precipitates changes in the

[73] Ekman (1999).
[74] Prinz (2004 a), p. 54.
[75] Ibid., 54-5.

body, and "our feeling of the same changes as they occur is the emotion."[76] Thus when he identifies anger with feelings of certain physiological disturbances-'boiling blood,' 'an agitating stomach,' - he means that feeling hot and agitated are the perception or thought of someone insulting one. Hence we can say that the charge that James distinguishes cognition from affect and puts emphasis on the affect[77] is misleading, for he supposes that one cannot understand emotions independently of thought.[78] It can be said that the reason why James's view is called an extreme case of an affect theory is because he sometimes highlighted the turbulence of emotions rather than their intentionality.[79] Now the difference between James' conception and that of cognitivism consists in the relation to thought that emotion might have. It can be said, as John Deigh points out, that the Jamesian theorist takes the relation as causal, whereas cognitivism regards it as constitutive.[80] I shall discuss the view that the relation to thought that emotion might have is constitutive in Chapter 2, and the view that takes the relation to be causal in Chapter 3.

Now for James, the way in which cognition and affect are related is that when one is angry at P, for example, some thought must be present. That is, one feels hot and agitated in response to some thought about P. If this interpretation is right, we can say as follows: on the one hand, for James, emotions are turbulent states. To experience

[76] James W. (1890/1950), p. 449.
[77] When critics argue that James distinguishes cognition and affect, they regard the latter as 'mere' feeling.
[78] In recent work on the emotions, there is a general consensus among philosophers that James does not distinguish between cognition and bodily feeling. In this respect, some people propose to defend James's view, though not thoroughly. According to these philosophers, on the one hand, James's theory is defective in that he fails to reckon with what can broadly be regarded as the rationality of emotions. On the other hand, they claim that James's theory can be remedied without losing his core idea that emotions are perceptions of patterned changes in the body. See, Deigh (1994), Prinz (2004 a &b), and Ratcliffe (2005a).
[79] In fact James sometimes strays from his explanation. See 1950, pp. 458-59. However, there is a research in much recent work concerning why James strays, but also claiming there is a consistency in his work. See Ratcliffe (2005a), pp. 179-202.
[80] Deigh (1994), p. 829 .

emotion is to be in some state of agitation, commotion, excitation, etc. In this respect, James seems to identify emotions with certain feelings of bodily changes. Yet James does not think bodily feeling or affect is divorced from intentionality, since when we have emotional experience, our attention is directed at our environment, our lives, or ourselves, and our thoughts and sensations oriented toward our emotional response. In his later work[81] James argues that we think or judge due to our needs, capacities, concerns, and bodily capabilities. In doing so affect plays a role. In this respect, according to James, our practical adaptation to our environment is partly constituted by affect. In what follows, I shall argue that James's this idea has been developed by De Sousa and Prinz.

[81] James W. (1897, 1902, 1907).

1.2. Perceptual Theory and the Problem of Intentionality

1.2.1. Emotional Truth

James's practical conception of intentionality has been developed in more recent works. According to de Sousa, an emotion is directed at a 'target.'[82] For example, being afraid of a snake implies that you are thinking of it as dangerous. If this is right, it can be said that an emotion has an intentional content that is evaluative. In order to lend support to this idea that emotions imply evaluations of their targets, De Sousa goes on to argue that each emotion-type has a so-called 'formal object' (corresponding to truth as the formal object of belief). Now in order for the formal object to be the possible target of an emotion, the subject must see the object as having a certain property; otherwise the emotion would not be intelligible. For example, your fear of the snake is only intelligible if there is some feature of it - looking fearful, in this case -which explains why you see it as fearsome. If you were saying that you are afraid of the snake while at the same time denying that there is anything fearsome about it, it would appear to be nonsense. Hence in order for your fear to be intelligible in this case, it is said that the property which has to be ascribed to an emotion's target is the emotion's formal object.

The term 'formal object' is derived from medieval philosophy and has been applied to emotions by Anthony Kenny. According to Kenny, the formal object of a state is the object under that description which must apply to it if it is possible to be in this state with respect to it.[83] According to him, when we describe the formal object of emotion, the description refers to belief, that is, one has to believe that something is dangerous in

[82] De Sousa (1987), p. 116.
[83] Kenny (1963), p. 189.

order to feel fear. Recently, however, it is commonly believed that the formal object of an emotion is a property. Hence, de Sousa holds that "the formal object of fear – the norm defined for its own appropriateness – is the Dangerous."[84]

The formal object of emotion view in much recent times has been defended by those endorsing a perceptual account of emotions.[85] According to this view, an emotion like fear has correctness conditions that are similar to the correctness conditions of perceptual states. Fear is correct or appropriate in so far as its object is really dangerous, in that this is what makes fear correct or appropriate. In his recent discussion, De Sousa claims that emotional truth is concerned with the correctness of the emotional evaluation, and holds that it refers not to semantic satisfaction but to success.[86] In order to lend support to this idea, de Sousa makes a further distinction between emotion with a propositional object and emotion with a direct object. He formulates the former and the latter as follows: "E (p) is *satisfied* iff p is true, [while] E (p) is *successful* if p actually fits E's formal object."[87] For example, the formal object of fear is the property of being dangerous. A rabbit-phobic person's fear of rabbits is satisfied if rabbits exist but it is not successful if rabbits are not dangerous. On the other hand, if someone is afraid of aliens, his or her emotion is not satisfied if there are no such aliens. However, it may be successful, since aliens could be dangerous if they existed. If this is so, emotional truth can be said to be whether the particular emotional object fits the relevant formal object.[88]

---

[84] De Sousa (2002), p. 251.
[85] De Sousa (2002, 2004); Prinz (2004 a &b); Doering (2003) ; Tappolet (2005).
[86] De Sousa (2004), p. 72.
[87] Ibid., p. 72.
88 Recently Mikko Salmela presents an argument for de Sousa's account of emotional truth by arguing that emotions have cognitive content as digitalized evaluative perceptions of the particular object of

Now if De Sousa's argument so far were right, one might raise a question, namely, how do we evaluate the correctness of emotional evaluations? Faced with this question, de Sousa introduces his axiological hypothesis of emotions as perceptions of value. According to him, we can understand what values are only in virtue of our emotiona responses. De Sousa calls his view 'axiological holism' for "it stipulates that we do not apprehend value in discrete units but only in the light of a complex of factors that transcend individual experience."[89] Among these factors are biological facts, social norms, and 'paradigm scenarios' of individual biography. Yet none of them alone constitutes the norm for emotional truth. "[I]nstead it is the totality of all these factors – biological facts, social, personal, and more – that may properly be confronted with one another in the hope of arriving at something like reflective equilibrium." [90]

Now I want raise an objection to de Sousa's perceptual view of emotions. In order to answer the question as to how we evaluate the correctness of emotional evaluations, he appeals to the idea of a wide reflective equilibrium of biological facts, social norms, and individual 'paradigm scenarios.' But this idea remains too unclear. He does not provide us with any example relevant to this definition. Moreover, the notion of a reflective equilibrium does not present any basis for judging between conflicts between biological, social, and personal factors.[91] In this respect, we can say that James[92] goes further than De Sousa as far as practical intentionality is concerned, since James claims that emotions do not simply enforce patterns, appropriately or inappropriately upon

---

emotion, in terms of the relevant formal property. Salmela develops this idea by arguing that an emotion is an actual fit between the particular and the formal objects of emotion, and the emotion's propositional content is semantically satisfied, or the target of the emotion exists. See Salmela, 'True Emotions', (2006), pp. 382-405.

[89] De Sousa (2002), p. 255.

[90] De Sousa (2004), p. 74.

[91] Salmela (2006) also makes this point.

[92] James W. (1897, 1902, and 1909).

situations; rather, according to James, emotions partly constitute the situation. Hence James claims:

> [T]he practically real world for each one of us, the affective world of the individual, is the compound world, the physical facts and emotional values in indistinguishable combination. Withdraw or pervert either factor of this complex resultant, and the kind of experience we call pathological ensues.[93]

James in the above passage seems to deny the notion of correspondence between emotions and facts or states of affairs, since for him our emotional relationship with the world is one of practical 'commerce' rather than theoretical 'correspondence.' If we agree with the idea that there is a correspondence between our emotions and the world, then we are conceding to the idea that emotions are capable of having epistemic warrant. Many contemporary writers regard this kind of warrant as the concept of appropriateness or fittingness that is taken to be analogous to truth in the emotional realm.[94] Yet, if we allow an analogy between appropriateness and truth, it would seem to allow that emotions are capable of being true or false. Many contemporary emotion theorists tend to be rather unclear on the idea that emotions are able to be true or false, when they say such things as the appropriate emotions "enable us to get things right,"[95] since it is said that such idiom as 'getting things right' is usually used in case in which the truth predicate is available. Now the reason why many philosophers have circumvented the concept of truth in the emotional realm is because there are some emotions that cannot be reduced to propositional attitudes which are eligible for being truth-apt, unlike beliefs, thoughts, and judgments etc. I shall discuss this issue in more detail in Chapters 4 and 5.

---

[93] James W. (1902),p. 151.
[94] e.g. Nussbaum (2001), Goldie (2000), D'Arms & Jacobson (2000a); Greenspan (1988).
[95] Goldie (2004), p. 99.

1.2.2. Embodied Appraisal

Now let us look at another version of the practical conception of intentionality which has been developed from James's conception. Prinz, for example, tries to offer an account of the intentionality of emotion without abandoning the idea that emotions are perceptions of bodily states. In order to do this, Prinz, first of all observes that emotions are intentional in two senses: they have formal and particular objects.

According to Prinz, formal objects are what Lazarus calls 'core relational themes'[96] – dangers for fear, losses for sadness, insult for anger, and so on. In this respect, we can say that the formal object is not merely a perceptual property but the norm by which each emotion type defines its own standard of fittingness, as De Sousa points out.[97] Particular objects, according to Prinz, elicit the emotion because they can be appraised as exhibiting the core theme expressed in the emotion. A great height is appraised as dangerous. Thus in this case we can say that our mental state "contains a representation of the particular object (a high place) as well as a representation of the property that makes it fearful (the danger of falling)." Hence Prinz claims that "emotions place representations of particular objects under concepts that represent core themes." He then claims that fear is warranted if heights are in fact dangerous, and not so if they aren't, for example if appropriate precautions have been taken. As such, emotions are representative, and capable of misrepresenting their objects.[98]

Now in order to explain intentionality, Prinz turns to the theory of mental representation. According to this idea, "a mental state gets its intentional content in virtue of being

[96] See Lazarus (1984).
[97] See De Sousa (2002).
[98] Prinz (2004a), p. 54.

reliably caused (or having the function of being reliably caused) by something."[99] If we apply this theory to the account of emotion, we can say as follows: the perceptual contents of emotions are bodily states, in virtue of their being reliably caused by bodily changes. Furthermore, if those bodily changes are brought about by the manifestation of core relational themes, then the emotions can also be said to represent these themes.[100] In order to appreciate this conception, let us consider again the case of fear. When we stand in fear on the edge of a tall cliff, we can say that the perception of that bodily state represents danger, since it is reliably caused by 'dangerousness' or a dangerous situation. "Danger is the property in virtue of which these highly desperate eliciting conditions have come to perturb our bodies."[101] Elsewhere Prinz gives more detail about the way in which emotions represent their formal objects as follows: emotions are simple mechanisms that represent their formal objects or "'core relational themes' (Lazarus) by having the function of being caused by the latter." Each emotion represents the bodily state and the presence of matters of concern such as danger and loss. "Emotions are gut reactions; they use our bodies to tell how we are faring in the world."[102] Prinz claims that we have the body-change elicitors in a mental file. That file can be extended over our life by experiences. Furthermore, Prinz suggests that elicitation files can include evaluative judgments. Hence we can say that "all the representations that trigger the bodily response will do so in virtue of being recognized as dangerous, either explicitly or implicitly by similarity to previously established elicitors."[103]

In short, when bodily perceptions represent core themes by reliable causation, we can

[99] Dretske (1981, 1988); Fodor (1990).
[100] Prinz (2004a), p.55.
[101] Ibid., 55.
[102] Prinz (2004 b), p. 69.
[103] Prinz (2004 a), p. 55.

say that they have formal objects. Prinz calls this amendment to James's theory the "embodied appraisal theory." As I have said before, on the one hand, for James, an emotion is simply the complexity of bodily sensation having a certain pattern. On the other hand, James sometimes implies that emotions are always consciously felt. According to Prinz, emotions are embodied just as James suggested. In Prinz's words, on the one hand, "they are perceptions of changes in our somatic condition: on the other hand, they are also appraisals."[104] Now the difference between the judgmentalist[105] and Prinz is that the former define an appraisal as an evaluative judgment while Prinz regards it as 'any representation of an organism-environment relation that bears on well-being.' [106] Prinz suggests that certain bodily perceptions represent an organism-environment relation that bears on well-being, and hence they are regarded as appraisals. The way in which the bodily perceptions represent an organism-environment relation is in terms of 'the right causal relations', not in terms of using 'concepts or providing descriptions.'[107]

Although Prinz's suggestion as presented so far contributes to capturing the intentionality of emotions without abandoning the Jamesian conception that emotions are bodily changes, I want to raise some doubts about his account. When he say that the formal object of emotion, for example, the formal object of fear, is 'dangerous' it seems to me that it cannot include many cases, such as radiation and the threat of a terror which cannot be perceived. The reason why is that, before there is a formal object, there needs to be a literally or physically presented object, which there is not in the case of

[104] Ibid., 57.
[105] Judgmentalism is the view that emotions are identified with evaluative judgment or belief. I shall discuss judgmentalism in more detail in Chapter 2 – 'Cognitive Theories of Emotions'.
[106] Prinz, Ibid., 57.
[107] Ibid.

radiation. Furthermore his view cannot explain other emotions, such as amusement, happiness, envy, guilt and shame. For these emotions have, as many point out, 'response–dependent' properties. If this suggestion is right, we can say that the response–dependent property of 'fear' is 'frightening' or 'fearful' rather than 'dangerous.' [108]

Prinz admits that we are 'frightened by things, angered by offensive things, disgusted by disgusting things, and elated by pleasing things.'[109] Yet he claims that a response-dependent view is viciously circular if those properties represent anything that causes fear, anger, disgust, and elation in us since the view seems to imply that those emotions can never involve error. I shall argue, however, in Chapter 5 of this thesis that Prinz's interpretation of the response-dependence theory of emotional properties is misleading. Before discussing this matter, I want to point out briefly a crucial point for a response-dependence theory of emotional properties here. The reason why a response-dependent emotional property cannot be involved in error, according to the response-dependence theory, is because all particular objects that cause fear in us are represented as frightening in our emotional experiences of fear but those objects are actually frightening only if they *merit* fear.[110] I shall return to this issue in Chapter 5. (§ 5.1 Sentiment and Value)

Now I want to raise another objection to perceptual account of emotion in general. Recently, some philosophers argue that emotions are very similar to perceptions or even that they are kinds of perceptions.[111] Among these philosophers, some try to explain the

[108] See Goldie (2004); McDowell (1998).
[109] Prinz (2004 b), p. 60.
[110] This point is made by Salmela. See Salmela (2006), p. 386.
[111] See e.g., De Sousa(2002); Prinz (2004).

appropriateness of emotion by making an analogy with the truth. In defence of emotional truth, de Sousa argues that an emotion is true if and only if there is an actual fit between the particular and the 'formal objects' of emotion. Adam Morton agrees with de Sousa in that emotions are capable of being true.[112] Yet he disagrees with de Sousa in that emotional truths are not accurate enough to be considered as truth as such. In order to demonstrate that emotional truth is too cheap, Morton presents the problem, which is implicated by the perceptual account of emotions, by raising the question whether two contrary emotions, for example, fear and delight, can be both present.[113] Morton's argument can be summarised as follows: (1) Fear and delight are emotions that are incompatible, though we can oscillate between them. (2) Both fear and attraction might be equally accurate with respect to the same situation S. If this were right, we could say that (3) the situation S can be both dangerous and attractive. However, according to Morton, if this were true, 'as De Sousa would put it, following Tappolet (2000), the values of danger and of attractiveness are both present.' Yet, Morton suggests that 'we shouldn't count them as truths, as it would be a strange kind of truth, such that having it committed one also to falsity.'[114]

If Morton's argument is right, then the perceptual account of emotion might be rejected. This is because as Tappolet notes, the perceptual account entails that 'if fear is appropriate with respect to a situation, then it is dangerous, and if attraction is appropriate to a situation, then it is attractive'.[115] However, Tappolet holds that if we can prove the way in which both fear and attraction are mutually exclusive, perceptual

[112] Morton (2002).
[113] Morton (2002).
[114] Morton (2002), pp. 274-5
[115] Tappolet (2005), p. 230.

account of emotion can avoid the alleged difficulty. Morton also acknowledges this, and his solution to this problem is as follows: at earlier stages of development, it is not possible to instantiate emotions such as fear and attraction at the same time, even though the situation may possess both the properties of fearfulness and attractiveness. However, as we develop and become more sophisticated, he claims we become capable of manifesting 'subtle emotions', such as 'delighted horror' which reveal both properties of the situation simultaneously.[116] However, if perceptual theorists accept this idea, they should clarify what the formal object of such 'subtle emotions', namely, 'delighted horror' would be. Is it 'the attractiveness of danger' or 'the dangerousness of the attractiveness of some objects'? It would seem that perceptual theorists cannot simply answer this question. Thus, setting aside the problem of what the formal object of such an emotion would be, Tappolet presents the alternative way to meet this challenge without presuming the existence of Morton's 'subtle emotions.'

In order to provide an alternative view to Morton's, Tappolet asks us to imagine that you are walking along a dangerous mountain path. Upon reaching a particularly dangerous section, you develop a certain set of emotions. Firstly, you become terrified as the next few steps might lead to your speedy and painful death. Secondly, you are strongly attracted to making the next few steps as completing the walk would be an exciting thing to accomplish.[117]

If somebody feels these two contrary emotions at the same time, do we have to say that the subject is in a state of conflict? If we assume that the perceptual account is correct, how can the perceptual theorist explain this kind of ambivalent emotion? In order to

[116] Morton (2002), pp. 274-75.
[117] Tappolet (2005), p. 230.

lend support to the idea that ambivalent emotions do not make for contradiction, Tappolet provides an argument as follows:

(1) The properties of danger and attractiveness are both disclosed as present.

(2) The emotions cannot be regarded as truth-disclosing, if it is the case that dangerousness and attractiveness are incompatible properties, which cannot be ascribed to the same situation S.

(3) However, the example shows that such properties may, in fact, co-exist.

(4) Therefore, emotions can unproblematically be regarded as truth-disclosing.

The central issue is whether danger and attractiveness are mutually exclusive values or not. In order to give an answer to this question, Tappolet develops her view as follows. She claims that one's emotions can *attune* one to two aspects of a present or future situation, for example, its danger and its attractiveness. 'Fear and, in this case, accurate awareness of the danger, tends to come with the thought that you should not go further, while the excitement and your attraction pushes you forward.' If this is true, something can unproblematically be both attractive in one respect (as a challenge to one's limits, for example), while unattractiveness in another respect (as dangerous). Thus, she argues that the quite inexorable co-instantiation of fear and attraction do not constitute a contradiction.[118]

Given this, we can say, as Tappolet notes, that 'our ordinary emotions of fear and attraction do not make for a contradiction even on the assumption that the perceptual

[118] Tappolet (2005), pp. 230-31.

account is correct'.[119] If her claim is right, cases of ambivalent emotions do not prove that perceptual theory of emotion is incorrect. Rather the cases show the fact that our emotional dispositions are far more complicated and varied. In order to appreciate the complexity of emotions, we should consider what De Sousa calls 'axiological holism' according to which 'we do not apprehend value in discrete units but only in the light of a complex of factors that transcend individual experience.'[120] These factors include 'biological facts', 'social norms', and 'paradigm scenarios' of individual biography. Hence De Sousa holds that the totality of all these factors constitutes the appropriateness of emotions.[121] If this were right, we could say that the appropriateness of a person' emotions can only be found through considering all these factors. I shall discuss this issue in the following two Chapters. (Chapter 3 – 'Cognitive theories of Emotion' and Chapter 4 – 'Do emotions have Direction of Fit?')

1.2.3. Emotion and Perception in Literal and Non-literal senses

In searching for emotions and their intentionality I have so far discussed the traditional feeling theory and its contemporary and present-day successors. The common point that these theorists share can be said to be that they are perceptual theories of emotion. Yet there are differences between them, depending on which kinds of theories are taken as truly perceptual theories. Hence the discussion which will follow focuses on illuminating in what sense they differ.

I have said that we can take Descartes' and Hume's views to be the traditional feeling theory. Hume says emotions (passions) are impressions of reflections (second-order

[119] Tappolet (2005), p. 231.
[120] De Sousa (2004), p.74.
[121] De Sousa, Ibid.

impressions.) 'Impressions' is often interpreted as Hume's term for perceptual states. However, the term 'Impression' should not be regarded as a synonym for perceptual state. Impressions are a contrasting kind that includes two distinct types of states. These Hume calls impressions of sensation and impressions of reflection. The former may be identified with perceptual states, and the latter, which includes emotions, makes up an entirely distinct class. A second-order impression is one that has been brought on by another impression, or by an idea (ideas are just faint copies of impressions, for Hume). Prinz claims that it is not entirely clear why emotions should qualify as perceptual on Hume's account, since Hume does not show that emotions qualify as states in dedicated input systems.[122] Neither does Hume draw any compelling analogy between emotions and paradigm examples of perceptual states, such as sights, sounds and smells. I have said that for Hume emotions have phenomenological qualities that they share in common with paradigm sensory states, in the sense in which they are similar to what it is like to experience a sensation. But as I have pointed out in section 1 of this Chapter, this is not the evidence for the thesis that emotions are perceptual, since, for Hume, desire or emotion differs from an impression of sensation or a belief in that it has no representational content. For example, as I have said before, visual experience represents the world as being a certain way, but pleasure has no such representational content. If this is true, we can say that Hume's view of emotion is not a perceptual theory.

Prinz finds the seeds of a truly perceptual theory in Descartes. For Descartes, emotions are states of the soul that occur in response to movements of the 'animal sprits' that flow through our bodies. Descartes sometimes says they are 'perceptions' of such

[122] Prinz (2004b), p. 223.

bodily changes. We can see this picture when he describes the way in which the perception of what constitutes the first cause and object of a passion affects one's body causing bodily changes. For example, when one sees an approaching lion on his path, the perception of the lion – the sight of a frightful beast on his path - affects his body, changing his heartbeat, bloodflow and muscular tension, causing his legs to tremble. These physiological changes feed back into the emotional reaction at the mental level, causing and strengthening the passion proper (one's fear). If this is true, we can say that for Descartes, the passions are perceptions of conative states caused by bodily changes or reactions. But according to Prinz, this view does not mean that emotions are states in perceptual systems. Instead, Prinz observes that Descartes regards emotions as pleasant or unpleasant feelings that draw our attention to some feature of the world and impel us to act in response. For example, Descartes defines love as a state of the soul, caused by movement of the animal spirits, that incites one to join those objects that seem agreeable. Love seems to involve a wish (the wish to join another object) and an evaluation (the assessment of an object as agreeable). But Prinz argues that neither of these components can be plausibly identified with a perceptual state. According to Prinz, Descartes, by closely associating the emotion with bodily changes, laid the foundations for perceptual theories. These foundations were developed by James in his identification of emotions with the feeling of bodily changes.[123] Now the reason why Prinz views James as the originator of a perceptual theory is because when James identifies emotions with bodily changes. They are not movements of the animal sprits but perturbations in visceral organs and adjustments in skeletal muscles. Prinz claims that this is clearly a perceptual

[123] Prinz (2004b), p. 224.

theory, because Prinz regards emotions as states in the somatosensory system.[124] In this respect, Prinz argues that his embodied appraisal theory descends from James's theory, because according to Prinz, "emotions are states within systems that are dedicated to detecting bodily changes."[125]

Another perceptual theory of emotion has been developed by De Sousa. De Sousa claims that emotion can suitably be seen as a form of perception, and he holds this claim on the basis of many parallels. For example, both perceptions and emotions can be felt, both are perspectival (one cannot entertain having an emotion without actually having it), and both can be hallucinatory (they can represent objects that are not really there). Although there are parallels between perceptions and emotions, de Sousa's theory of the emotions, as Prinz points out, does not turn out to be a perceptual theory. Let us look at the reason why we cannot attribute the perceptual theory to de Sousa's view. De Sousa first of all characterizes emotions as 'patterns of salience' that imply directions for inference and inquiry. Then he claims that emotions are learned by observing 'paradigm scenarios,' which are situations in which emotions are characteristically elicited. If this so, we can say that emotions arise when we recognize current events as sufficiently similar to paradigm scenarios. Now if we have a representation of a paradigm scenario in our mind, it draws our attention to characteristics of a present situation and disposes us to draw certain inferences. "Anger may arise when we recognize someone's behaviour as offensive, and we may be moved to draw certain conclusions about that person's character."[126]

---

[124] Ibid.
[125] Ibid.
[126] Prinz (2004b), p. 222.

Prinz agrees with de Sousa that paradigm scenarios play an important role in emotion elicitation. Yet Prinz argues that de Sousa's theory is not perceptual in the sense in which Prinz himself uses the phrase. Prinz agrees with de Sousa that there are several parallels between emotion and perception. Both perceptions and emotions are, for example, felt, perspectival, and capable of being hallucinatory. However these and other parallels do not amount to more than a metaphorical affinity between emotion and perception because de Sousa's theory "does not implicate activity in a perceptual input system."[127] Even if perceptions are involved in the recognition of salient similarities between a present situation and 'paradigm scenarios', these are just elicitors of emotion while emotion proper boils down to primitive gut reaction. De Sousa seems to allow this kind of interpretation for he admits that emotions differ from sense perceptions by their opaqueness and internal, bodily locus.

De Sousa explicitly denies that emotions are beliefs. The reason why de Sousa rejects strong cognitivism (emotions are belief), is because we experience groundless emotions in spite of well-founded beliefs.[128] Although he denies strong cognitivism, he is himself committed to cognitivism, since he suggests that perception is a form of cognition that is capable of accommodating groundless emotions. He admits that emotions differ from sense perceptions by being relatively opaque. In this respect de Sousa calls emotions modes of perceiving. But 'mode of perceiving,' according to Prinz, is best interpreted as a metaphor on de Sousa's account. Hence Prinz argues that for de Sousa "emotions are interpretations of things without being explicit judgments." Furthermore Prinz argues that this view can be applied to perceptual states. "If I have yellow lenses in my

[127] Ibid., p. 223.
[128] De Sousa (2004).

sunglass, I may see the world as yellowy without judging that it is yellow." We can also see this character in many nonperceptual states. "A child might 'see' her stuffed animals as living things. A mugger may 'look at' passer-by as victims, when deciding whom to assault. An artist may view a model as Venus or Adonis."[129] Prinz holds that these are forms of perception only in a non-literal sense. If he is right, we can say that emotions for de Sousa are not literally perceptions. In other words, for de Sousa, perceptions are merely analogous to emotions. Prinz describes the relationship between perception and the emotions as follows:

> A dedicated input system is a mental system that has the function of receiving information from the body or the world via some priority class of transducers and internal representations. Dedicated input systems are perceptual modalities or senses. To count as perceptual, a mental state must inhabit a sense. Vision, audition, and olfaction are dedicated input systems. They each have their own neural pathways and proprietary representations.[130]

Hence Prinz argues that "if emotions are literally perceptual, they must reside in such a system."[131] I do not discuss the contemporary perceptual theory of emotion further. Instead, In Chapters 3 and 5, I shall present some difficulties that the perceptual theory encounters. Before looking at this issue, in the next Chapter I shall first of all discuss cognitive theories, which are regarded as the main rival to feeling-centred theories of emotion.

[129] Prinz (2004b), p. 223.
[130] Ibid., p. 202.
[131] Ibid.

Chapter 2: Cognitive Theories of Emotion

Preliminary Remarks

I have argued in the previous Chapter that the discussion of intentionality leads into a more detailed classification of the object of emotion. In doing so I have focused on the traditional feeling-centred conception of emotion that exemplified by Descartes, Hume and William James. Then I have shown some difficulties that these views encounter. The question whether the theory of emotion is cognitive or not is the question whether the cognitive elements, such as belief and judgment, are *conceptually* necessary or not. Hence the discussion which will follow in this Chapter focuses on contemporary cognitivism, which puts the emphasis on the necessity of the cognitive element having intentional content. As we have seen, the traditional feeling-centred theory denies this, and in this respect we can dub the feeling-centred theory 'standard non-cognitivism.' We shall see that contemporary cognitivists argue that the objects of emotion have a special character that justifies the inference from their intentionality to propositional thought. In order to show this, they maintain that what qualifies something as the appropriate object of an emotion is the subject's belief that it has a certain character. Hence, for the cognitivists, belief and so propositional thought, is essential to emotion.

2.1. Cognitive Turn

2.1.1. Belief

I begin by illuminating the reason why cognitivism has dominated the philosophical study of emotions. First of all, cognitivists attempt to solve the problem of 'phenomenological individuation' of the traditional feeling-centred theory of emotion. As we have seen in Chapter 1, on the traditional feeling theory, the distinction between the emotions of irritation and anger, shame and embarrassment cannot be distinguished merely in terms of 'pure feeling.' But if the cognitive element is regarded as essential to the analysis of emotions, the problem may be solved. For example, your *belief* that your action was not ethically right is likely to cause the feeling of shame. On the other hand, your *apprehension* that your action does not fit into your culture, even though it is not ethically wrong, causes you to feel embarrassment. In a similar way, you may get irritated or angry depending on your response to a certain action. Secondly, normative evaluation, which is a problem for the traditional feeling-centred view, lends support to cognitivism. Indeed, the various normative evaluations of emotional responses, i.e., rationality or irrationality, appropriateness or inappropriateness, represent the various cognitive elements involved in the emotions. Thirdly, it is an important characteristic of mental phenomenon that the 'objects' of our emotional responses appear to be enormously various. Let us take the emotion of anger: you can be angry with the bad behaviour of members of parliament, ice over the road, the stiffness of a needle. Sometimes, you may get angry with a certain character in a movie, sometimes with your friend. Indeed, there are a number of 'objects' with which we may get angry: objects, states of affairs, facts, legends, fictions and characters of legends and fictions...etc. The

relationship between the 'objects' of these emotional responses and the states of our emotions cannot simply be a causal one, for the tragedy of a non-existent character of fiction cannot cause my grief. Thus the problem of the various objects of those emotions and the relation between them could be solved by thinking that the concept of emotion implies a cognitive element. Cognitivists also argue that the nature of the cognitive element is intentional. They argue that the objects of intentional attitudes are not simply limited to real objects but whatever is possible for their objects, for example, legend or fiction. In this respect it might be said that the cognitive theory has much greater explanatory power than the traditional feeling-centred theory. Yet we can say so only if we take 'minimal condition of cognitivism', according to which what qualifies something as the appropriate object of an emotion is the subject's belief that it has a certain character. The minimal condition for the cognitive theory is clearly shown in following remarks by K. Walton:

> It seems a principle of common sense, one which ought not to be abandoned if there is any reasonable alternative, that fear must be accompanied by, or must involve, a belief that one is in danger. Charles does not believe that he is in danger; so he is not afraid.[1]

Let us consider the minimal condition for the cognitive theory before examining 'strong cognitivism,' that is, 'judgmentalism,' and 'hybrid theory.' We cannot feel the emotion of 'pity' without the belief that someone encounters difficulties and frustration. We cannot feel the emotion of 'regret' without the belief that I did wrong for someone. A dangerous man would not be feared if he were not known or believed to be dangerous. If you do not *know* of a certain social norm - for example, that it is impolite to point in a certain culture, you may not feel insults, and sometimes your ignorance can cause

1 Walton (1979), pp. 6-7.

pleasure inadvertently. Thus the apprehension of certain circumstances is a necessary condition for the understanding of certain emotions. However, cognitivism, which puts the emphasis on the necessity of the cognitive element runs into difficulty with 'mood'. For example, the emotions of faint anxiety and melancholy are not about certain objects or states.

The typical response to this objection is, following A. Kenny, to regard those emotions as not representative cases for the emotions but as merely peripheral ones.[2] In other words, when we call a certain state the emotion of 'faint anxiety,' it is because of certain similarities to the typical case of anxiety. Thus according to this response, these counter-examples can be understood in terms of the typical cases of the emotions which imply propositional attitudes. In the light of this, the cognitivist might claim that the counter-example is an inappropriate objection to cognitivism. But it seems to me that a better solution to the problem of 'mood' is to notice the following facts: firstly, faint anxiety and melancholy may sometimes be dispositions rather than episodic states of a person. Secondly, even if 'mood' is an episodic mental state, it is wrong to say that it is a non-intentional state. Since the fact that the emotion of 'faint anxiety' is not about a *certain* object does not entail that a 'mood' should be non-intentional. Thus it is more plausible to say, it seems to me, that the state of 'faint anxiety' appears to be non-intentional because their objects are very often rapidly changing and are general.

Another objection concerning Walton's thesis that 'it is impossible to have the emotions without accompanying appropriate belief' is whether the cognitive element involved in the emotions is *the state of belief* or not. For example, a mother is anxious about the

[2] Kenny (1963), p. 63.

security of her daughter when she goes to school. Suppose the mother knows that the distance between her home and the school is just 5 minutes on foot and there is no road crossing on her way, and there is hardly any traffic. Moreover, if she asks: 'is my daughter coming safely back home?' She may feel the emotion of anxiety at the moment in which she is *believing.*[3] Of course these cases are irrational emotional responses. However, the irrational emotions should also be regarded as emotional states. The point of this objection is to doubt whether the cognitive element involved in the emotions is necessarily a state of *belief.* Indeed, cognitive theorists do not always claim that the cognitive element, which is essential to the emotions, should be 'belief.' Belief, among the many cognitive attitudes, is somehow a 'very strong' kind of attitude, for belief itself is necessarily connected with evidential rationality or justification. We do not have the belief-attitude when there is no or insufficient evidence. In other words, if cognitivists claim that the cognitive element, which is essential to the emotions, is belief, there is a risk of ruling out a number of 'irrational emotions.' In this respect, cognitivists substitute 'judgment', 'evaluation', 'appraisal,' 'apprehension', or simply 'construal' for the belief. According to these concepts, for example, if you *judge* that 'X is anxious rather than safe', the judgment does not entail that you believe that 'X is anxious.' Since we may think that the evidence is vague or insufficient to believe it. In this respect, the mother's faint anxiety for her daughter's safety can be cognitively understood as follows: it is possible for the mother to be anxious about her daughter's safety simply because of her *belief* that 'her daughter may not be safe' and her *desire* for the safety of her daughter. Thus it seems to me that cognitivism is plausible, once the cognitive element involved in emotions is broadly understood as I have described. That is, I think,

[3] It is common that this kind of objection is raised. See Bergman (1978); Calhoun (1984); Schaffer (1983).

the minimal condition for cognitivism.

In what follows I shall divide cognitive theories into three categories: first, 'pure' cognitivism or judgmentalism, the claim that the cognitive element itself is not only a necessary but also a sufficient condition; second, 'hybrid theory,' both 'feeling' and a 'cognitive element' are necessary for emotion. Thirdly, 'qualified cognitivism,' cognitive element is a necessary condition but the emotion itself is regarded as non-cognitive element added on the cognitive element. I shall discuss 'qualified cognitivism' in the next Chapter. Before looking at this view, in what follows in this Chapter, I shall discuss 'judgmentalism' and 'hybrid theory' and present the difficulties that they encounter.

2.1.2. Judgmentalism

2.1.2.1. Strong Cognitivism

In contemporary discussion, there has been an increased appropriation of the cognitive dimensions of emotion. Robert Solomon and Martha Nussbaum, for example, have — presented cognitive views that can be called judgment theories. According to them, an emotion can be analyzed solely in terms of beliefs, or evaluative judgments. Solomon holds that "emotions are something far more sophisticated than mere feelings."[4] He goes onto argue that emotions are rational and purposive and like actions, and that we choose an emotion as we choose a course of action.[5] Furthermore, he claims that emotions are normative judgments and that "since normative judgments can be changed

[4] Solomon (1976/1993), p. 102.
[5] Solomon (1980), pp. 251-52.

through influence, argument, and evidence, and since I can go about on my own seeking influence, provoking argument, and looking for evidence, I am responsible for my emotions as I am for the judgments I make."[6] In his book *The Passions,* he maintains that an emotion is an "evaluative judgment, a judgment about my situation and /or about other people"[7] He writes:

> An emotion is a (set of) judgment(s) which constitute our world, our surreality, and its 'intentional objects.' An emotion is a basic judgment about our Selves and our place in our world, the projection of the values and ideals, structures and mythologies, according to which we live and through which we experience our lives.[8]

There have been objections against identifying the emotions with appropriate evaluative judgment. P. Greenspan, for example, argues that our emotions may be in an ambivalent state but the judgment corresponding to the emotions may not be in the ambivalent state.[9] For example, someone feels joy because his friend has won the lottery, and at the same time he may feel jealousy. In this case, his contrary emotions, such as joy and jealousy, are co-existent. But rational judgment can unify the conflict between these opposite emotions as follows: on the one hand, "his winning is good in that it satisfies the desire of someone I identify." On the other hand, "his winning is bad in that it frustrates a desire of my own." But it is evident that the resolution of the conflicting elements in terms of rational judgment will not always be a solution, for my jealousy caused by the judgment may be 'overriding' rather than mitigated even if I judge that "his winning is *on the whole* good." Greenspan argues that "ambivalence seems to be

[6] Solomon (1980), p. 261.
[7] Solomon (1976/1993), p. 187.
[8] Ibid., p. 185.
[9] Greenspan (1980), pp. 223-50.

possible in persons not so irrational as to hold genuinely contrary judgments." [10] In this respect, Greenspan maintains that the 'logic of emotion' differs from the 'logic of judgment', for it is not *irrational* for my emotions to resist qualification. Hence she claims that "emotions may persist even when they are accompanied by a stronger opposing feeling, in a basically rational person."[11]

To counter this criticism, Solomon presents a modified form of his theory. He smuggles in the qualitative notion of 'intense.' According to Solomon, what is distinctive about emotional judgments is that they are "self-involved and relatively intense evaluative judgments....the judgments and objects that constitute our emotions are those which are especially important to us, meaningful to us, concerning matters in which we have invested our Selves" [12] It seems, however, as many have pointed out, that his judgmentalism is easy to criticise. The main objection is that it overlooks the affective aspect of emotion. It fails to do justice to a person's emotional state such as being upset, perturbed, agitated or moved. In this respect, some argue that this problem can be solved by adopting a theory which is in part non-cognitive. I call this view a hybrid cognitive theory. Before looking at this, let us consider the objections against judgmentalism further.

Let us call the problem of identifying the emotion with evaluative judgment or belief 'the problem of insufficiency' of judgmentalism. The reason why we regard judgmentalism as insufficient is that the evaluative judgment or belief does not imply the existence of a corresponding emotional state. Thus what is at issue for

[10] Ibid., p. 223.
[11] Greenspan, Ibid, p. 234.
[12] Solomon (1976/1993), p. 188.

judgmentalism is to find the *specific* kind of cognitive element by which all emotions can be explained. Thus the point of the insufficiency objection to judgmentalism is that the subject who has an appropriate normative evaluation of certain circumstances does not necessarily feel the corresponding emotions. For example, we can say that some expert who may judge that 'London bridge is very dangerous' does not necessarily feel fear. Solomon responds to this objection as follows: the emotion should not be identified with ordinary normative judgment but with a *special kind* of judgment or evaluation. According to him, the judgment which can be identified with a certain emotion is a "self-involved and relatively *intense* evaluative judgment." He argues that sometimes the judgment is important for us, even desperately so.[13] In short, he maintains that the judgments which can be identified with appropriate emotions are not ordinary judgments but *urgent*, even *desperate*, *intense* responses to situations. Now if Solomon's judgmentalism were right, some emotions, for example, pity and sympathy cannot be distinguished from unemotional judgment, since they are not a serious concern for a subject himself or overvaluation. These emotions, for example, pity and sympathy, can be distinguished from unemotional judgment in terms of *altruistic* concepts, which are not imbued with serious self-concern or overvaluation. Hence it seems to me that the judgementalist strategy which attempts to distinguish emotional from unemotional judgment by adding a specific additional characteristic of the cognitive element cannot succeed, for there are many emotions which can be distinguished from unemotional judgment excluding a specific additional characteristic.

Now even if Solomon presents this modified version of judgmentalism in order to meet

[13] Solomon (1980), pp. 251-81.

Greenspan's objection, his view, as J. Robinson notes, still has problems.[14] Firstly, an 'emotional' judgment cannot be a judgment that is particularly *intense*, for *judgments* are not the sorts of things that admit of degrees of intensity. However, Robinson herself accepts that both *emotions* and *desires* allow degrees of intensity, for the intensity of emotion is based upon the intensity of the underlying desire. For example, according to Robinson, if I am intensely afraid of snakes, I have an intense desire to avoid them. Of course, we can talk about the degree of certainty or uncertainty of judgment. But we cannot talk about an intense or moderate judgment. Moreover, 'a desperate, meaningful judgment relevant to the subject' does not always imply the corresponding state of emotion. Thus "I may make the dispassionate judgment that my adored husband is a drunken and a liar." Furthermore, according to Robinson, there are other more obviously urgent judgments which are not emotional. For example, the judgment "I had better go now" may be made urgently. But it may be a purely prudential or practical judgment, and not central to any emotion.[15] Thus Robinson argues that the cognitive element which can be identified with an emotion is not a specific judgment as Solomon claims, but the combination of an appropriate judgment (belief) and *intentional desire.*

Placing special weight on the role of desire in emotion rather than dispassionate emotions, Robinson argues that if we identify the emotions with the combination of judgment and desire, we can solve many problems with Solomon's judgmentalism. The characteristics of many emotional attitudes which are described as 'desperation' of emotional state or 'intense' emotional response can be easily explained in terms of the state of 'desires' which underlie emotions. For desires, unlike judgments, can be clearly

[14] Robinson (1983), pp. 731-41.
[15] Ibid., pp. 733-34.

distinguishable in terms of degree of intensity or desperation. Thus she claims that the concept of desire implied in the emotions can remove the objection related to the 'existence of the ambivalence of emotion' raised by Greenspan. As we have seen, according to Greenspan, ambivalent emotions, unlike judgments, can be co-existent, resisting unification or synthesis, and a specific judgment and emotion cannot be identified. Robinsons' point here is that unlike judgments, the different desires can be co-existent, for the important feature of emotions which links them to desires is their resistance to change. Hence she maintains that "my failure to sum up my conflicting emotions is due to my failure to sum up the corresponding desires."[16] Following Greenspan's lead, Robinson modifies Greenspan's claim as follows: "the 'logic' of emotion *resembles* the 'logic' of desire."[17] If we take this assumption, Robinson argues, we can explain the difficulty presented by Greenspan. The Greenspan's example shows how two inconsistent emotions and their corresponding desires can persist unchanged in a 'basically rational' person. I desire both my own good and that of my friend, and I cannot simply 'drop' one of these desires; hence I cannot simply 'drop' the corresponding emotion. But there is nothing irrational in this behaviour if, as Greenspan argues, there are adequate reasons for the two emotions.[18] Robinson argues that there are adequate reasons for the two emotions in this case, for the two emotions are based on two equally rational desires.[19]

In my view, Robinson's main contribution is an attempted defence of judgmentalism, by adopting intentional desire, in addition to the standard judgmentalist resources,

[16] Ibid., p. 735.
[17] Ibid., p. 734.
[18] Greenspan, Ibid., pp. 234-38.
[19] Robinson, Ibid., p. 736.

namely belief or judgment. She accepts that cognitive assessments are important in emotions, but argues that these assessments are both caused by, and 'coloured' by desire. A desire-determined conception of an object is, she argues, always, necessary and in certain cases sufficient for the phenomenon of emotion. I doubt, however, that the combination of judgment and desire can be identified with emotions. For example, we can imagine the completely rational being who judges the same way as we do in given circumstances and has the same degree of intense desire as we do, but who does not feel the corresponding emotion that we do. He may be 'inhuman' in that he, unlike us, may not feel 'frustration' when his intense desire is frustrated and may not feel the joy of his success. It seems to me that it is not conceptually impossible. In order to understand this point more appropriately, let us examine the following example:

> Two men while walking down a path see a snake. Though they both evaluate it as dangerous, want to avoid harm, and take steps to avoid the danger, one is gripped with fear while the other remains calm.

In this case, we can say that the frightened man is upset or disturbed; he is distracted or perturbed. Thus, it is likely that he trembles or shudders with fear. Let us call the combination of belief and desire of the person who feels fear 'emotional judgment,' and call that of the person who does not feel fear 'unemotional judgment.' The problem is, in this case, how the judgmentalist can explain the difference between emotional and unemotional judgment. N. Nash, who sees himself as offering a new version of Solomon's judgmentalism suggests the solution to this problem as follows:

> ...what makes some evaluation/desire complex emotional is not its association with typical physiological changes, though such changes may accompany all or most emotions; it is rather the presence of a focused attention of a particular sort on the object of emotion, resulting in the agent's overvaluation of that object (relative to

his dispassionate evaluation of it).[20]

According to Nash, emotional judgment can be distinguished from unemotional judgment in that the former has a 'focused attention' and 'overvaluation.' If this modified version of judgmentalism were right, we could say that the man who feels fear of the snake is in a state of focused attention on the snake, resulting in his overvaluation of the danger of the snake, relative to the other man who evaluates it dispassionately. It seems, however, as many have pointed out, that this judgmentalism is easy to reject. The main objection is that it overlooks the affective aspect of emotion. It fails to do justice to a person's emotional state such as their being upset, perturbed, agitated or moved. The judgmentalists also acknowledge that most emotions are accompanied by physiological changes and their feelings. Moreover, they think that these physiological changes are caused by the combination of the appropriate belief and desire. Yet, they believe that physiological changes are just a *symptom* of a given emotion, but are not necessary for the concept of emotion. However, if this kind of cognitivism were correct, then the cognitivist account of emotion would render feeling theory superfluous. In this respect, some argue that this problem can be solved by adopting a theory which is in part non-cognitive. Such a theory is offered by the hybrid theorist, according to which emotions include both cognitive and non-cognitive constituents. I shall criticize this kind of cognitivism later. Before looking at it, let us consider a similar view to Solomon's.

2.1.2.2. Neo-Stoic View

There is another kind of judgmentalist, who identifies emotion with evaluative

[20] Nash (1989), p. 482.

judgment or belief, following the Stoic view which defines emotions in terms of beliefs. We can trace this idea back to Plato's and Aristotle's work. In his work *Protagoras* and the *Laws*, Plato claims that emotion, for example fear, is a cognition: the expectation of impeaching evil.[21] In *Phaedrus,* Plato holds that beliefs and other cognitions play some role or other in emotional struggle. According to him, these beliefs are sometimes, but not always, mistaken ones. The bad horse, which represents the lustful part of the soul, can think it right to given in to temptation.[22] Another way one can give in to temptation is in terms of the deceit of pleasure.[23] Pleasure and distress can, for Plato, make mistakes about their objects because they are accompanied by a false belief.[24] Now the Stoics, for example Chrysippus, developed this idea further that the judgments involved in emotions are always false.[25] According to the Stoics, in order to be free from emotions, one should get one's beliefs concerning the value of things right. Hence the rational person should get rid of emotions, since without any rational basis, their beliefs are false beliefs about the good or bad at hand. There seems to be this kind of strong cognitivism in contemporary discussion of emotion. The strong cognitivist, for example, G. Taylor claims that 'an emotional reaction is unjustified if it rests on irrationally mistaken belief or when it is disproportionate to a given situation.'[26] Given this, she argues that an emotion, for example, malicious envy or jealousy is reducible to false belief.[27] In order to appreciate this point, let us consider an example. Suppose that Kate and Lucy have been friends and they assess themselves by their success in some

[21] Plato, *Protagoras* 358 D; *Laws* 644 C-D.
[22] *Phaedrus* 255 E-256 A.
[23] *Laws* 863B-C.
[24] Philebus 37E.
[25] In searching for the relationship between the Greek philosophers and the Stoics about emotion, I was helped by Richard Sorabji's recent work (2001).
[26] Taylor (1975), p. 393.
[27] Taylor, for example, argues that jealousy is unjustified, since this emotion must include a false or unjustified belief. See Taylor (1975), pp. 401-2; see also Ben-Ze'ev (1990), p. 506.

competition. They apply for a job at the same place. But Lucy is accepted and Kate fails in an interview. So Kate is jealous of Lucy's success. In being jealous of Lucy, Kate thinks of Lucy as an enemy to her self-esteem. So despite their friendship, Kate wishes that Lucy fails. Kate wants Lucy to be absent from the final interview due to illness. If the strong cognitivists were right, in this case they would say that Kate's jealousy involves false belief, and is morally wrong. The reason why, the strong cognitivists might say, Kate's jealousy involves false and morally wrong belief is that her emotion rests on irrational belief which cannot be justified. Now the presumption of this kind of strong cognitivist's inference, it seems to me, is that there is an analogy between our emotional response and the world, and a true belief and the world. According to them, the former relation can be described as 'fittingness', whereas the latter one is maintained by truth. Then the cognitivists argue that these attitudes are both rational. Since they take both to be rational, they argue that emotion can be justified or criticized. Furthermore, they claim that justification or criticism of the emotions is always committed to morality or prudence. But this is, I shall argue, what some call a 'moralistic fallacy.'[28] I shall discuss this issue in more detail later (Chapter 4).

In *On the Soul*, Aristotle also provides a cognitivist view of emotion. According to him, emotions might be calmed either by addressing the physiological process or by addressing the cognitions. For Aristotle, emotion, for example anger, is not only desire for retaliation but also a boiling of blood around the heart.[29] The Stoics, for example Chrysippus, developed the former idea – for example, anger is desire for retaliation – to

[28] See D'Arms and Jacobson (2000a), p. 73. "The moralistic fallacy" is '"imply to infer, from the claim that it would be morally objectionable to feel F toward X, that therefore F is not a fitting response to X." The reason why this inference is fallacious, according to D'Arms and Jacobson, is because "an emotion can be fitting despite being wrong to feel." (pp. 73-74). If this is right, we go wrong when we take evaluative sentences as entailing imperatives.

[29] Aristotle, *On the Soul,* I. I, 403 a25-b9.

cognitive view, since he observes that the idea of retaliation implies *awareness* of distress received and distress to be returned and regards the latter – physiological processes – as mere concomitants. Although Plato and Aristotle addresses the idea that emotions are cognitive, the view attains new depths with the Stoics. On the Stoic view, the mind is monistic: there is only one kind of soul, the rational, and hence we can say that any kind of mental phenomenon is qualified as in some sense rational. According to this view, all emotions are judgments, and hence depend on reason. For example, according to Chrysippus, emotions which he calls passions, *pathe,* are judgments and judgments that lack any rational basis.[30] He holds that right philosophy is all we need to free ourselves from emotions. In order to do this, he suggests that we should change our judgments.[31] The Stoics believe that an essential role of emotion is judgment, which is the only element over which reason is taken to have full control. The control is worked through the power of assent, which following the Stoic theory is 'up to us': reason is free to give or withhold assent from any present impression, no matter how violently it may affect or agitate the mind. The Stoics hold that peace of mind is a necessary and sufficient condition for human happiness. Furthermore they suggest that peace of mind is a recommending state of indifference and that the absence of passions (apatheia) is a prerequisite of virtue. Hence they argue that only the person who is undisturbed by any emotions and the bonds they create is sovereign over their own reason and judgment.

This idea of the Stoic therapy of emotions is recently suggested by Martha Nussbaum in her *The Therapy of Desire.*[32] Although Nussbaum recommends the Stoic theory of emotion, she does not wholly agree with the idea. She agrees with the Stoics that

[30] Annas (1992), pp. 108 and ff. See also; Sorabji, (1993), p. 59.
[31] See Nussbaum (1993).
[32] Nussbaum (1994).

emotions are or involve beliefs, but does not agree that all emotions are false beliefs. Nussbaum takes an emotion to be the full acceptance of, or recognition of a belief. She writes:

> Emotions, it is said, are unreliable, animal, and seductive. They lead away from the cool reflection that alone is capable of delivering a considered judgment...a central purpose of these essays is to call this view of rationality into question and to suggest, with Aristotle, that practical reasoning unaccompanied by emotion is not sufficient for practical wisdom; that emotions are not only not more unreliable, and less deceptively seductive.[33]

According to her, emotions are judgments concerning things important to our own well-being, which acknowledge our reliance on parts of the world beyond our control.[34] The judgments involved in love and grief, for example, are eudaimonistic: "they evaluate the external object or person an important part, not of the world from some detached and impersonal viewpoint, but of the world from the viewpoint of the agent's own goals and projects."[35] She maintains that emotions are judgments of values, which incorporate cognitive aspects. Then she views such judgments as not merely necessary constituent components in the emotions, but also sufficient components, for if the emotion is not there we are entitled to say that the judgments are not fully or really there.[36] Thus she argues that to feel grief or fear is not to have a bodily pain but to form a judgment.[37] She writes:

> Anger is associated with a boiling feeling, fear with a chilled and queasy feeling. But here we should beware of the word "feeling," which is remarkably slippery and likely to mislead. We should distinguish "feelings" of two sorts. On the one hand,

[33] Nussbaum (1990), p. 40.
[34] Nussbaum (2001), p. 19.
[35] Ibid., p. 82.
[36] Ibid., p. 44.
[37] Ibid., pp 57- 8.

> there are feelings with a rich intentional content – feelings of emptiness of one's life without a certain person, feelings of unhappy love for that person and so forth. Feelings like these may enter into the identity conditions for some emotions; but the word "feeling" now does not contrast with our cognitive words "perception" and "judgment," it is merely a terminological variant of them. ... On the other hand, there are feelings without rich intentionality or cognitive content, let us say feelings of fatigue, of extra energy, of boiling, of trembling, and so forth. I think we should say about these exactly what we said about the bodily states: that they may accompany an emotion of a given type and they may not – but that they are not absolutely necessary for it. In my own grief, feelings of crushing fatigue alternated in a bewildering way with periods when I felt preternaturally wide awake and active; but it seemed wrong to say that either of these was a necessary condition of my grief.[38]

Critics have argued against Nussbaum that she ignores the feeling or affective component.[39] In order to meet this kind of challenge, in much of her recent work, Nussbaum distinguishes two kinds of feelings: one with a rich intentional or cognitive content, and the other lacking such content. As an example of the former she mentions "feelings of emptiness of one's life without a certain person." As an example of the latter she mentions "feelings of fatigue, of extra energy, of boiling, of trembling, and so forth." She claims that such feelings, which lack cognitive content, may accompany emotions, such as feelings of fatigue accompanying grief, but they are neither necessary nor sufficient for emotions. Feelings of the former kind, however, with cognitive content, and which she claims are 'terminological variations' of perceptions and judgments, are necessary but not sufficient.[40] Hence she claims that emotions can be defined in terms of the cognitive evaluative element only.[41]

---

[38] Ibid., p. 60.
[39] c.f., Ben-Ze'ev (2004)
[40] Nussbaum (2001), p. 60.
[41] Ibid., pp. 63-4.

Having established Nussbaum's downplaying concerning the role of feeling in emotions, we have arrived at determining the plausibility of her view, that is, whether or not feelings are an essential component of the definition of emotions. It can be said that not all evaluative cognitive judgments are components of emotional states. As I mentioned before, when I say that I'd better go now, this judgment does not cause any emotion in me. How can we then distinguish between emotional evaluative judgments and non-emotional evaluative judgments? Let us look at our previous example again.

> Two men while walking down a path see a snake. Though they both evaluate it as dangerous, want to avoid harm, and take steps to avoid the danger. One is gripped with fear while the other remains calm.

As I have said, Solomon and Nash distinguish emotional judgment from unemotional judgment such that the former has a 'focused attention' and 'overvaluation,' and hence feels fear of the snake, whereas the latter who evaluates it dispassionately does not feel fear of it. Now in order to distinguish those two judgments, Nussbaum suggests various cognitive features, such as a 'personal point of view,' 'cognitive denseness,' and 'cognitive freshness.' These features do not imply noncognitive components since they impart the cognitive content of the judgments. Let us look at such features in more detail.

The first cognitive feature is the 'personal point of view,' which is essential to emotions. When we have emotional experiences we look at the world from our own point of view, attaching to the various events the personal importance they have focus. Thus we can say that in our example, the two men feel differently because they attach their own personal value and preferences to the situation or just have different characters. The man who feels fear may do so because he wants to save his life for attending

tomorrow's job interview, whereas the other man wants to be dead since he thinks his life is miserable.

Now the second cognitive feature is cognitive denseness. She claims that the experience of emotion "contains rich and dense perceptions of the object, which are highly concrete and replete with detail."[42] What she means by 'cognitive denseness' here is that emotions are related to imagination. Nussbaum describes her feelings of grief on hearing of her mother's death as follows:

> [My] feeling, of terrible tumultuousness, of being at the mercy of currents that swept over me without my consent or complete understanding; the feeling of being buffeted between hope and fear, as if between two warring winds; the feeling that very powerful forces were pulling the self apart, or tearing it limb from limb: in short – the terrible power or urgency of the emotions, their problematic relationship to one's sense of self, the sense one has that one is passive or powerless before them.[43]

A close look at Nussbaum's intention in the above passage reveals that non-cognitive or bodily feelings are not necessary or sufficient elements of an emotion. Rather she takes evaluative judgments to be the constituent parts of an emotion, since she identifies emotion with judgment. Nussbaum seems to try to explain how emotions can be violent, stormy, fading, suffocating etc in purely cognitive terms without reference to (bodily) feelings.

Nussbaum holds that grief is not just an abstract judgment plus the ineliminable localizing element: it is very highly particular. Hence, feeling grief involves a storm of memories and concrete perceptions that swarm around that content, but add more than is

[42] Ibid., p 65.
[43] Ibid., p. 26.

present in it. This 'more', she writes, is cognitively laden, or dense, and this denseness is usually, if not always, a necessary feature of the experience of an emotion such as grief. When she grieves for my mother, she *sees* her and the sight is, like a picture, dense and replete.[44] Now, a fading of the imagination takes place as a result of the lost person's absence. When she no longer sees her mother before her, the perceptual image of the object is weakened, leading the content of the emotion to shift. Hence her grief is gradually diminished by this experience. But Nussbaum sees this emotional change which takes place when an emotion fades as a cognitive change. Nussbaum calls this features of emotion 'cognitive freshness.'[45]

Let us look then at 'cognitive freshness,' which is the third cognitive feature of emotions. When "my violent grief for my mother's death has by now grown calmer," this does not mean that "I no longer believe that she is dead," and that "I no longer believe her enormously wonderful and valuable."[46] What then the difference between her calmed state of November 2005 and her violent grief state of August 1995? According to her, the difference is a cognitive one, since when the imagination fades, "it is *because* I no longer see my mother before me that I no longer make her such an important part of my life." This shift can be understood in eudemonistic propositions, such as "my mother is an important element in my flourishing"[47]; I now am more inclined to accept the proposition, "The person who died *was* a central part of my life, "and this judgmental change itself is a large part of what constitutes the diminution of grief."[48] For example, she looks forward to celebrating her mother's birthday, then

[44] Ibid., p. 65.
[45] Ibid., pp. 79-85.
[46] Ibid., p. 80.
[47] Ibid., p. 84.
[48] Ibid., p. 82.

remembers that this is not to be, thus she has to amend her prospect and plans for the future. The recognition of her own future goals or projects shifts the judgments of the importance of her mother into the past tense. Nussbaum calls this feature 'cognitive freshness.'

Now this feature makes the given emotion a background emotion rather than a situational one. When grief progresses, for example, we find it becomes a background emotion rather than a situational emotion. What then is the difference between a background and a situational emotion? According to Nussbaum, the former is one that persists throughout situations of numerous kinds, whereas the latter is one that is elicited in the context of some particular situation.[49] An example of the former may be: many women live in a state of continuous anger at the domestic injustice that is a part of their daily lives; and yet the anger will surface only in certain circumstances.[50] However, Nussbaum claims that background emotions need not be conscious, just as episodic or situational emotions need not be nonconscious; but frequently they will be, since they are persisting conditions that are often unnoticed partly on account of their pervasiveness.[51] Nussbaum does not want to admit nonconscious emotion in her account, in part because they lack the phenomenological and imaginative feature, but she admits them partly because of methodological reasons.[52]

So far we have examined Nussbaum's attempt to distinguish emotional judgment from

---

[49] Ibid., p. 69.
[50] Ibid., p. 70.
[51] Ibid., p. 71.
[52] When Nussbaum mentions nonconscious emotions, this does not mean in the Freudian sense. It means, according to her, in a much more ordinary sense, just the sense in which many of our most common beliefs are nonconscious, although they guide our actions in many ways: beliefs about cause and effect, beliefs about numbering, beliefs about where things are, beliefs about what is healthy and harmful, and so on. In the light of this, she claims that if we weren't like this, if we could use only those beliefs on which we were consciously focusing, we couldn't possibly survive. (2001), pp.71-2.

unemotional judgment by appealing to various cognitive features. Although Nussbaum's account has plausibility in some sense, for example, in explaining the fact that grief fades as time passes by, her view is defective in that she ignores the aspects of *felt* urgency and the affective side of the emotional experience. She seems to admit this point when she says that "it is the emotion itself, and not some further reaction to it, that has urgency and heat."[53] But she views this urgency as a cognitive feature for it "comes not from the unthinking force, but from my thought that my well-being is threatened by that force."[54] For example, if there is urgency in being hit by a gust of wind, it is not after all a noncognitive urgency – the urgency, if it is there, comes not from the unthinking force, but from my thought that my well-being is threaten by that force.[55] My objection that Nussbaum omits feelings might be met by trying to show that the relevant feelings are not bodily but can themselves be unpacked 'cognitively.' She claims that the occurrence of non-cognitive features in emotional experiences is contingent to the definition of emotions.

[53] Ibid., p 78.
[54] Ibid.
[55] Ibid.

2.2. Emotions Are Not Mere Judgment of Value

Peter Goldie, a qualified cognitivist, claims that Nussbaum's cognitivism fails to recognise that "how profoundly and systematically our emotional feelings can mislead us – how the emotions can distort perception and reason."[56] In order to establish the point that emotions are not merely a cognitive-evaluative state, in what follows I demonstrate how affect or bodily feeling also plays an important role in our emotional experiences. I will do this by taking an example from *Remembrance of Things Past* by Proust.[57] When Marcel hears the news of Albertine's death, he suffers great anguish. Up until this point he had believed that he did not love her. However, upon hearing the distressing news, he realizes with certainty that he did indeed love her. According to Proust, Marcel has deceived himself about his love for Albertine. If we hold, like the cognitivist, that recognition of our condition can be best explained by a detached unemotional and non-affective cognition, then we do not have the resources to explain Marcel's shock and suffering. This intellectual judgment that he did not love Albertine is shattered by his anguish. Of Marcel's self- deception, Proust writes:

> I had been mistaken in thinking that I could see clearly into my own heart. But this knowledge, which the shrewdest perceptions of the mind would not have given me, had now been brought to me, hard, glittering, strange, like a crystallised salt, by the abrupt reaction of pain.[58]

Marcel's feeling -anguish- is said to have the power, just in terms of its own *felt* quality, 'like a thunderbolt'[59], and forces his assent to his state of love. His state is revealed by his feelings, not his intellect. Proust tells us that our feelings are systematically buried

[56] Goldie (2004), p. 91.
[57] Nussbaum (1990) also discusses this example.
[58] Proust (1981), Vol. III, p. 426.
[59] Ibid., p. 431.

by the workings of habit, the primary manner in which self-deception operates. The truth of Marcel's condition is brought about by his feelings. His painful anguish demonstrates his prior state of self-deception. If this is indeed a typical or common phenomenon, then it shows that a full account of our condition cannot be given simply by appeal to a cognitive-evaluative state. If this is right, the attempt by Nussbaum to reduce the role of cognition and affect in our recognition of emotional change to a single cognitive element might be implausible. In response to this objection, Nussbaum could say that Marcel's feeling – anguish – wasn't just a physical sensation of pain that articulated his emotion, but the pain of a particular loss. She writes:

> We also have a type of pain that probably is necessary for grief: namely, the pain *that an important element of one's life is gone.* But of course that is not a noncognitive element, and we have already included it in our cognitive/evaluative account, which has stressed, with Chrysippus, that such losses are bad and that it's right to be upset about them.[60]

If this is true, she might say that Marcel's feeling, in this case, is psychological pain, would yield to a purely 'cognitive' analysis. Thus, she maintains that "emotions can be defined in terms of these evaluative recognitions alone, although we must recognize that some feelings of tumult or 'arousal' will often accompany them, and sometimes feelings of a more type-specific kind, and although we must recall that they are at every point embodied."[61] Hence she holds that the existences of noncognitive characteristics in emotional experiences are contingent to the definition of emotions, but part of the characterization of the typical case. But if we accept this, Nussbaum must explain what 'affect' is and whether it can be fairly contrasted with cognition. However, she remains unclear about this matter.

---

[60] Nussbaum (2001), p. 64.
[61] Ibid.

I have argued that Both Solomon's and Nussbaum's cognitivism ignore a feeling component. Although Nussbaum tries to address this feeling component by distinguishing two kinds of feelings, one with a rich intentional or cognitive content, and the other lacking such content, she downplays the feeling component. She writes that the latter variety of feelings is neither necessary nor sufficient for emotions, while feelings of the former kind, which she claims are 'terminological variations' of perceptions and judgments, are necessary but not sufficient. Furthermore, she holds that the existences of noncognitive characteristics in emotional experiences are contingent to the definition of emotions, but part of the characterization of the typical case. Hence she claims that emotions can be defined in terms of the cognitive evaluative element only. But if this is true, as I have said, Nussbaum must explain what 'affect' is and whether it can be fairly contrasted with cognition. If she cannot do this, her cognitivism remains distinctly unsatisfying. *Pace* Nussbaum, if affect is not contrasted with cognition, then the real problem of explaining emotion is how we can develop a thorough account of how emotions dissolve the distinction between thoughts and feelings, or cognition and affect. In what follow, I will examine a view which indeed grants a role to cognition and affect in our emotional experiences.

2.2.1. A Hybrid Theory

Before discussing a hybrid theory, let me remind us of what the strategy of the judgmentalist is. The judgmentalist attempts to distinguish emotional from unemotional judgment by adding a specific additional characteristic of the cognitive element to the combination of judgment and desire: a focused attention, overvaluation, and serious concern. It seems to me, however, that this strategy is not available, for judgmentlaism

cannot explain kinds of depression. When someone feels miserable, sad, or hopeless, due to his belief that he has lost something important, it seems to me, his depression cannot be understood as a 'focused attention', 'overvaluation', or 'serious concern', as R. Solomon and N. Nash argued. Depression seems to be defined in terms of another element different from that which the judgmentalist suggests. There seems to be no 'overvaluation' in the person's depression deriving from losing something important, though those events are terrible. Moreover, anything in the world, even the painfulness of his body and the miserableness of his life cannot be a focused attention for him. Hence we should conclude that it is impossible for the judgmentalist to define emotion in general, comprehending emotions such as fear and depression. Furthermore, this objection is also available for the other emotions. If judgmentalism were right, some emotions, for example, jealousy, and vanity cannot be distinguished from unemotional judgment without a serious concern for a subject himself or overvaluation. However, some emotions, for example, pity or sympathy can be distinguished from unemotional judgment in terms of *altruistic* concepts, which are not serious self-concern or overvaluation. Hence it seems to me that the judgmentalist strategy which attempts to distinguish emotional from unemotional judgment by adding a specific additional characteristic of the cognitive element cannot succeed, for there are many emotions which can be distinguished from unemotional judgment excluding a specific additional characteristic.

The feeling centred theorist might argue that there is a way to demarcate emotional from unemotional judgment. According to the feeling centred view, the difference between the man who fears the snake and the man who does not fear it consists in a 'feeling'. That is, the former feels fear while the latter does not. In this respect, hybrid theorist

suggests that the difficulty with judgmentalism can be resolved by reintroducing a noncognitive element, that is, a feeling.

A hybrid theory is one that attempts to defend cognitivism by adopting the noncognitive element, which is downplayed by judgmentalism. William Alston puts this point as follows:

> [...] we cannot identify emotions with evaluations alone, without completely losing contact with such phrases as "emotional reaction," "getting emotional over it," and "controlling one's emotions." An evaluation can be either emotional or unemotional. Two people can see a snake as equally dangerous, and yet one is gripped with fear while the other is calm.[62]

According to Alston, only a hybrid theory can explain the affective aspect of emotion. He suggests that the difference between emotional and unemotional evaluations is that an emotional evaluation causes in the evaluator bodily disturbances and sensations.[63] Hence it can be said that to be emotionally perturbed is, for example, to tremble or shudder with fear.

The hybrid theory, on the one hand, resists *judgmentalism* in that for judgmentalism, judgment or belief identified with emotion includes 'coldness,' in that if we had nothing but cognitions we would be unemotional. If so, the hybrid theorist argues that the emotion identified with judgment cannot be the appropriate emotion. The hybrid theory, on the other hand, resists the *feeling theory* in that the Jamesian identifies emotion with a certain pattern of physiological changes and bodily sensations, while the hybrid theorist takes them to be only a component of an emotion. Thus, according to the hybrid theorist, if we combine cold judgment with the element of feeling, we get the concept of

[62] Alston (1980), vol. 2, p. 485.
[63] Ibid., p. 486.

emotion in general. Hence the hybrid theorists, most notably, Alston and Lyons[64] maintain that we can have the concept of emotion in general when the appropriate cognitive elements cause bodily feeling or sensations. The way in which the appropriate cognitive elements cause certain feelings is as follows: a) when I judge that someone has been treated unjustly; b) I want people not to be treated unjustly. These states (belief and desire) then cause: c) some kinds of bodily disturbance. Hence getting emotionally worked up, according to Jerome Shaffer, must involve more than just belief and desires, namely, the physiological/ sensational effects of those beliefs and desires.[65] In other words, in explaining emotion, the hybrid theorist takes an emotion to be a complex of physiological process and sensations caused by certain beliefs and desires. For example, in seeing a cat while I am driving round a bend, I *believe* that there may be bodily harm and I *desire* not to be harmed. But in this case it is not sufficient to explain my feeling fear by the belief and desire, since in order to feel fear, the belief and desire must affect me. The way in which the belief and desire affect me is that my face may turn pale, my heart rate may rise, and my stomach may tighten. These physiological changes feed back into the emotional reaction at the mental level, causing the passion proper (my fear). Then it can be said that in this case to experience feeling fear is to experience a particular kind of physiological or sensational state which was caused by a particular belief/ desire complex. If this is right, according to the hybrid theory, an emotion is determined by the kind of beliefs and desires and the kind of physiological and sensational effects. It seems to me, however, that this view also encounters difficulties even though it appears to be 'the ideal combination' of the feeling centred theory and the cognitivist view.

---

[64] Ibid., pp. 479-86; See also Lyons (1980).
[65] Shaffer (1983), p. 161.

2.2.2. Objections to the Hybrid Theory

Although the hybrid theorist adopts feeling to solve the problem which the judgementalist faces, the hybrid theory is also essentially a cognitive theory. In other words, the hybrid theory resolves the problem of individuation of the traditional feeling theory in terms of cognitive elements. Thus it can be said that while for James, each of the different emotion-types, such as anger, fear, grief, jealousy, is distinguished by a unique pattern of bodily disturbances and feelings, the hybrid theory individuates the emotions in accordance with types of judgments or beliefs following the judgementalist. In this respect, we regard the hybrid theory as cognitivism. Then one might wonder whether physiological changes or the combination of bodily changes, which the hybrid theorist takes to be a necessary condition for emotion in addition to the cognitive element, play a role for illuminating emotion in general or not.

The first difficulty which the hybrid theory encounters is a conceptual one. The judgmentalists also acknowledge that most emotions are accompanied by physiological changes and their feelings. Moreover, they think that these physiological changes are caused by the combination of the appropriate belief and desire. However, they believe that physiological changes are just a *symptom* of a given emotion, but are not necessary for the concept of emotion. However if the physiological/ sensational component is taken to be a necessary component as the hybrid theorists do, one might argue, they cannot explain the case of spinal cord lesions. However, as I have shown in Chapter 1, this objection can be met in the following way: emotional responses may be able to bypass the body by means of an 'as-if-loop.'

Now even if the first objection can be met by the hybrid theorist, it seems to me that our understanding of intentional action can be misled by the hybrid theory. In order to appreciate this point, let us take the example of fearing snakes again. According to the hybrid theory, the difference between the man who makes an emotional judgment and the man who does not, lies in the fact that the former feels specific physiological changes, whereas the latter does not, and at the same time both entertain the same belief /desire. But how are the men's reactions caused by those cognitive emotional states? First, the man who fears snakes basically manifests the same reaction as the man who does not feel fear, since we supposed that both have the same belief and desire in the given circumstances. Both may try to escape from the harmful situation. The scared man, as the hybrid theorist argues, may shudder. But although the scared man feels fear, if he reacts in the same way as the man who does not feel fear, since he has been a trainer in the zoo for a long time, then the sensation of pleasure or displeasure which the hybrid theorist adopts, must be unnecessary for understanding intentional human action. Perhaps physiological changes or displeasure, which are the scared man's shuddering hands and sweating, can be understood as involuntary behaviour. But this is not an explanation of human action.

Secondly, the scared man may express a completely different reaction from the man who does not fear. For example, the man might freeze or may run away leaving his friend behind. According to the hybrid theory, in order to appreciate the scared man's behaviour, we should understand the element of sensation, that is, displeasure. In other words, according to the hybrid theorist, the scared man's particular action can either be understood as intending to avoid displeasurable bodily sensation, or be understood as the displeasure hindering us from making a cool judgment. If my objection is right,

hybrid theory, according to which our emotional response is possible since bodily changes cause pleasure /displeasure of bodily sensation, is problematic. A more important point in the case of the scared man. It seems to me that the object of his feeling is not the displeasure caused by his bodily changes but the poisonous snake. At the same time, the reason why he cannot make a cool judgment is not simply because of a sensation of displeasure, but because of the snake's ferocious eyes, which he cannot forget. In this respect, some people suggest that sensation, or the element of feeling, which the hybrid theorist adopts, is unnecessary to explain intentional action. I shall return this issue later and present an alternative to hybrid theory in the next Chapter.

In sum, the judgmental theory of emotion attempts to illuminate the reason why we think of emotions as unified mental states, as the traditional feeling theorist did. However, if my argument so far is right, it can be said that judgmentalism fails to unify the emotions, such as fear and depression, so as to encompass the concept of emotion in general. On the other hand, in order to solve the judgmentalist's problem, the hybrid theory that adopts physiological changes or bodily sensation is also implausible, in that it admits of crucial counter-examples or leads to serious misunderstanding of the intentionality of human action related to the emotions.

Chapter 3: How do (bodily) feelings and thoughts combine to constitute different emotions?

Preliminary Remarks

In the previous Chapter, I have argued that the dominant philosophical view of emotion during the last thirty years has been cognitivism. There are variations of cognitivism and I have discussed a spectrum of views. There has been agreement that the pure cognitive theory, which identifies emotion with evaluative judgment or belief, is inadequate. I have called the pure cognitive theory 'judgmentalism.' As we have seen, the main objection to the standard pure cognitive theory is that it cannot explain the affective aspect of emotion. Judgmentalism is in opposition to the feeling theory, according to which emotions are identified with feelings of bodily change. The previous Chapter has also discussed a hybrid theory, according to which emotions include both cognitive and non-cognitive constituents. The hybrid cognitivist attempts to avoid the difficulty of judgmentalism by adopting a non-cognitive element from traditional feeling theory in the analysis of the concept of emotion. Thus the reason why we call this view the hybrid theory is that it is a combination of the traditional feeling theory and the cognitivist view. On the one hand, it draws on the traditional feeling theory in the sense that it reintroduces the element of feeling, that is, 'bodily feeling', in order to analyze the emotions. On the other hand, it follows judgmentalism in the sense that it individuates the emotions in accordance with types of judgments or beliefs.

In the previous Chapter, I have also raised several objections to cognitivism, especially a modified version of judgmentalism, and hybrid cognitivism. My main objection to the modified version of judgmentalism was: although the judgmentalist tries to distinguish emotional from unemotional judgment by adding a *specific additional characteristic* of the cognitive element to the combination of judgment and desire, that is, '*a focused*

attention' or *'serious concern'*, they cannot distinguish some emotions- for example, pity or sympathy- from unemotional judgment. For these emotions are other-regarding, not serious self-concern or overvaluation. On the other hand, my main objection to hybrid cognitivism concerns the inadequacy of the concepts of physiological changes or bodily changes, which the hybrid cognitivists adopt in order to solve the judgmentalist's difficulty.

In this Chapter I shall argue that both judgmentalism and hybrid cognitivism have difficulties to be admitted. When the judgmentalist stresses the role of cognition in emotion they presuppose that cognition and bodily feeling are distinct. The hybrid theorist suggests that emotions incorporate both cognitive and non-cognitive constituents. Although the hybrid theorist admits emotions are essentially link with bodily feelings, they also think that cognition and affect are distinct. Then what is at issue for an adequate theory of emotion, it seems to me, is to explain what holds cognition and bodily feeling together. Before looking at my alternative, in what follows, I shall focus on more recent discussion, which mainly addresses the following issues: *an emotion is a state incorporating both bodily feeling and cognitive states.* This leads to the following question, namely, *how do bodily feelings and thoughts combine to constitute different emotions?*

3.1. Qualified Cognitivism

In the previous Chapters, I have discussed two crude views of emotions of which attempt to reduce emotions via a basic element. The one is that emotions are essentially bodily feelings. The other is that emotions are essentially cognitive-evaluative states. We have seen that this kind of reduction is inadequate. In order to provide an alternative to these two reductionist views, in the previous Chapter, we have considered a hybrid theory. However, as I have argued, the hybrid theory has difficulties to be admitted, since although the theory suggests that emotions incorporate both cognitive and non-cognitive constituents, it also presupposes that cognition and affect are distinct. Many contemporary philosophers have proposed componential views. Patricia Greenspan takes emotions to be compounds of affective states of comfort and discomfort and evaluative propositions.[1] Recently Peter Goldie proposes new versions of cognitive theories, defending the irreducibility of the affective element.[2] In what follows, I shall call these componential views, for example, Goldie's position, a *qualified cognitivism*. I shall differentiate the qualified from hybrid cognitivism, since for the qualified cognitivism, the feeling is a necessary condition of emotion, and the cognitive element is a secondary addition which is not always present, whereas according to the hybrid cognitivist emotions essentially include both cognitive and non-cognitive constituents.

3.1.1. Feeling theory leaves out the intentionality of emotion and normative evaluation of emotion.

[1] Greenspan (1988).
[2] See Goldie (2000, 2002).

As we have seen before, the main objections to the feeling-centred view are: first, it cannot satisfactorily explain the intentionality of emotions. Second, it cannot satisfactorily represent emotions as proper objects of rational assessment. According to the feeling theory, emotions, on the one hand, assimilate the phenomena to bodily sensation, especially the sensations of pleasure and pain. Emotions, on the other hand, differ from bodily sensation in that they are located in the soul itself. In this respect, the modern philosophers, for example, Descartes, Locke and Hume take emotions to be internal sensations of pleasure and pain. Yet, this modern view, as I have argued before, leaves out the intentionality of emotion. Since according to the traditional feeling theorist, bodily sensations of pleasure and pain are not directed at or toward anyone or anything. In what follows, I shall show a different view from the traditional feeling theory, according to which a bodily feeling is also intentional, since when we feel something from the inside of the condition of our body, the feeling is directed toward an object, that is, our body as undergoing certain change.[3] The feeling-centred theorists have overlooked this fact. William James, however, sometimes wanders away from his account. As I have argued before, according to James, on the one hand, an emotion is simply the complexity of bodily sensation having a certain pattern. One the other hand, James sometimes, implies that emotions are always consciously felt. In another words, on the one hand, for James, emotions are turbulent states. To experience emotion is to be in some state of agitation, commotion, excitation, etc. In this respect, James seems to identify emotions with certain feelings of bodily changes. On the other hand, for James, emotions are intentional phenomena, since when we have emotional experience, we are directed our attention to our environment, our lives, or ourselves, and orient our

[3] C.f. Goldie (2000, 2002, 2004).

thoughts and touches our sensibilities.

Although Jamesian theory is capable of accommodating the intentionality from bodily feeling in explaining emotion, he should be criticized for his failure to acknowledge the fact that emotions are proper rational assessment. This criticism can also be attributed to the traditional feeling centred theory. As I have argued before, it is believed that a certain emotional response is in some respects 'appropriate' to the circumstances; in other cases a response can be seen as 'irrational.' In short, an emotion has a 'normative' character in terms of rationality, appropriateness, or reasonableness. The traditional feeling-centred theory and Jamesian does not allow the notions of reasonableness in our emotional experiences.

3.1.2. Emotions are not mere judgments of values: Ambivalent Emotions and Recalcitrant Emotions

As I have argued before, the judgmentalist holds that an emotional state is a combination of some cognitive element – evaluative judgment or belief – with some desire. This theory individuates specific emotions in terms of differences in their constitutive belief or judgment. Although not every philosophical theory of emotions adopts this picture, it is dominant in the philosophical study of emotion. However, there has been agreement that judgmentalism is inadequate, due primarily to P. Greenspan. Thus a substantially modified version of cognitivism appeared. Greenspan's main objection to judgmentalism, as I have argued before, is that the judgmentalist cannot explain emotions that conflict with judgment, since for judgmentalism, emotion can be unified by judgment. Yet, according to Greenspan such conflict exists, for the 'logic of

emotion' differs from the 'logic of judgment'.

Greenspan argues that a rational person can have ambivalent emotions with respect to one and the same object (being happy that p and unhappy that p, where p is that someone who is both my rival and my friend who wins the lottery instead of me). If this is right, she claims that emotions are different from evaluative judgments – it would not be rational to believe that p is good and that p is bad, for instance. Since two contrary emotions can both be appropriate but contrary judgments cannot both be true, she argues that being appropriate is different from being true. Now in order for Greenspan's argument to be adequate, she must show that emotions are not judgments or perceptions of values.

Indeed, the evaluation of value is a complicated matter. Something can be good in one way, while bad in another way. Thus, it could well be the case that being happy that p is appropriate just if p is good in some way. But this would be perfectly compatible with the claim that unhappiness about the same p is also appropriate: something good in some way can also be bad in some other respect – since he is my friend, it's good in a way that he won the lottery, but at the same time it's disappointing that he got the money instead of me.

Greenspan's solution to the problem of judgmentalism is that the logic of emotion and the logic of judgment are different. Hence in her recent work, Greenspan suggests that "we should distinguish between rational appropriateness as a kind of evidential warrant, and social or moral appropriateness as fitting or failing to fit social or moral norms."[4] And she seems to assume that the appropriateness of emotion refers to the latter.

[4] Greenspan (2003), p. 120.

Although Greenspan is quite right to reject judgmentalism about emotions, the argument against judgmentalism, as Christine Tappolet notes, should not be based on ambivalent feelings. Rather Tappolet suggests that it should be based on cases of so-called irrational emotions, such as when one fears something while judging that there is absolutely no danger involved.[5] I call these emotions 'recalcitrant emotions.'[6] A recalcitrant fear, for example, is one where the man is afraid of the snake despite believing that it poses little or no danger. Or someone may feel fear toward the cliff, despite believing the cliff not to be dangerous. If the judgmentalist were right, they should say that the man, who does not believe that the snake is particularly dangerous, cannot be afraid of it, since according to judgmentalism, the relevant belief is a necessary element of the emotion. But judgmentalism cannot explain the above two cases which show us the emotional recalcitrance. If the judgmentalist adopts the possibility of emotional recalcitrance, they should admit that the relation between judgment and emotional response is inconsistent and the agent's behaviour is incoherent.

Now a question arises: if 'recalcitrant emotion' is not based on belief or evaluative judgment, on what are such emotions based? Some qualified cognitivists claim that recalcitrant emotion is derived not from belief but from concern-based 'construal', or 'feeling-toward.'

3.1.3. Concern based Construal

[5] Tappolet (2005), p. 229.

[6] I adopt this terminology from Susan James. See James (2003). In a similar context, Peter Goldie uses another concept: 'cognitively impenetrability.' According to him, our emotional experiences are cognitively penetrable only if they can be affected by our beliefs. On the other hand, our emotions are cognitively impenetrable when they are not affected by our persistent beliefs. See, Goldie (2000), pp. 74-78. But I prefer the term 'recalcitrant emotion' on the ground that it gives us a more vivid idea of emotion's passivity.

What then is construal? Construal, according to Robert Roberts, is "*a characterization of the object*, a way the object presents itself." We often see these cases: "when I see the duck-rabbit as a duck, the figure itself takes on a ducky look."[7] In order to characterize the appearance of the object, my concern is committed to the construal. The way in which my concern enters into the construal is that I pick options up from my psychological repertoire by direct or indirect control. But sometimes I am faced with situations in which I do not have options to pick up in order to construe them, due to lack of control. There are several senses of being unable to control an emotion: 1) being unable to helping feeling it. (e. g., feeling sorry for a burglar); 2) being unable to help expressing it (e.g., the witness in court bursting into tears); 3) being afraid of a grass snake, despite believing that it poses no danger. These cases are possible because to some extent our emotions and emotional responses are passive, and cannot be controlled. The third case is one of recalcitrant emotion, which is the emotion over which I have no control. Now what then is the phobic person's fear about? This case gets the judgmentalist into trouble. If I believe that there is nothing to be afraid of in the presence of a perception with content that *not p* (e.g., the rabbit is not harmful), then I should not fear it. If I do, it can be said that it is a mistake to fear *p* when *not p*, and the fear should be abandoned. But it is possible that fear persists, even when I believe that the rabbit is harmless. If the judgmentalists were right, they should say that the man or woman who does not believe that the rabbit is particularly dangerous cannot be afraid of it, since according to judgmentalism, the relevant belief is a necessary element of the emotion. However, judgmentalism cannot explain the above cases which are examples of emotional recalcitrance. Hence it follows that emotional recalcitrance gets

[7] Roberts (1988), p. 192.

judgmentalism into trouble.

One way of handling the case of a phobic person's emotional reaction, following Locke would be: the state of 'uneasiness' when one sees a spider determines the will to run away from the situation. Now a question arises: why the phobic person, unlike normal person, is in the state of 'uneasiness' about the spider? It seems to me, that Locke cannot answer this question. We can say that the phobic person's fear is similar to the following case: the Muller-Lyre lines continue to appear to be of different lengths while they are known to be equal lengths. In regard to this idea, many people try to develop it to show how, at times, our thoughts, feelings and actions seem to behave with a mind of their own despite our best efforts in trying to control them. In this respect, we can call phobic fear a recalcitrant emotion. This is because the emotional subject's attitude toward the state of affairs is very complex and the desire in which the emotion is involved depends on all kinds of things: the circumstances, character-traits and moods that the person has. In this respect many philosophers argue that these emotions are passions in the sense that "we are often passive in the sense of being unable to control them."[8] As we have seen, the traditional feeling theorists, for example, Descartes and Hume, explain the passivity of emotion in terms of something shared with sensation. Yet, as I have said in Chapter 1, emotions are "more agitating and disturbing than sensory perception of bodily sensation."[9] Ignoring this fact, the judgmentalists, notably, Solomon, tend to regard all emotions as voluntary actions, whereas a qualified cognitivist, for example, Roberts, takes the view that "the subject of an emotion is both a) sometimes able to exercise voluntary control over it and b) sometimes unable to do

[8] James, S. (1998), pp. 21-2.
[9] Ibid.

so." Perhaps the feeling theorist takes emotions as events over which, by definition, we have no control.

If my argument so far is right, we can say that the emotions' unreasonableness depends not on the belief but on the construal made beyond our control in an inappropriate way, due to the emotions' passivity. However, the judgmentalist may think that the appropriateness of emotion depends on the accuracy of the construal. The judgmentalist might say that to judge an emotion as reasonable is to think its evaluation to be correct. They might say, as Roberts notes, "if I am a rational person, judgments are almost entirely forced upon me, in any given situation, being required by evidence and logic."[10] However, I shall argue that the reasonableness of emotion does not amount to our assessment of the accuracy of the construal, as the judgmentalist might say, for it does not depend on true or false belief, but is contextually determined. I shall discuss the problem of the appropriateness of emotion in the next Chapter (Chapter 4).

3.1.4. 'Feeling-Toward'

I have argued that recalcitrant emotion is derived not from belief or evaluative judgment, but from 'concern-based construal.' Now most people are afraid of (harmless) grass snakes. Is that because they construe them as (harmful) adders? If so, that seems to return us to judgmentalism; do they just react to them as if they were dangerous when they perhaps know that they aren't? If we don't want to return to judgmentalism, we should consider construal as an activity that is different from judgment or belief. In this respect, P. Goldie defines construal as 'feeling toward.'

[10] Roberts (1988), p. 200.

In attempting to explain how feeling can be integrated into the world-directedness of emotion, Goldie distinguishes two kinds of feelings: bodily-feelings and feelings-towards.[11] These two sorts of feelings are both directed towards an object. "Bodily feelings", Goldie writes, "are directed towards the condition of one's body, although they can reveal truths about the world beyond the bounds of one's body – that, for example, there is something dangerous nearby." On the other hand "feelings towards are directed towards the object of the emotion – a thing or a person, a state of affairs, an action or an event; such emotional feelings involve a special way of thinking of the object of the emotion."[12]

The idea of 'feeling towards' can best be understood in terms of comparing the states of beliefs and desires. Beliefs and desires are a person's attitudes towards propositional contents. We can distinguish these two states in terms of 'direction of fit.' Beliefs 'aim at' 'truth', whereas desires 'aim at' 'realization' or 'satisfaction.' But 'feeling towards' is, as Goldie remarks, "not an attitude which is defined in terms of any particular 'direction of fit'," for it differs from belief in that "it is not an attitude but is related to the will."[13] If this is so, 'feeling towards' can be cashed out by the idea that we 'try to think of X as Y,' since trying to think of is sometimes subject to the will. This means that in 'feeling towards', we try to think of X as Y. However, we cannot try to believe something, for believing at will is impossible and a belief aims at truth by depending on evidence. We cannot try to believe, for example, "the pudding to be disgusting, whereas we can try to feel disgust toward it - by trying to think of it as disgusting (trying to think

11 Goldie (2000), Ch. 3; see also Goldie (2002), pp. 235-245.
12 Goldie (2002), p. 235.
13 Goldie (2000), p. 72.

of it as vomit perhaps)."[14] Another difference between belief and 'feeling towards' lies in the fact that "beliefs, once rationally arrived at on the basis of evidence, tend to persist until further evidence becomes available which should lead to their revision; whereas feeling towards something is much more episodic."[15] The third difference between 'feeling towards' and belief is that feeling towards "tends to have a perceptual quality of a sort which is, at least usually, lacking in belief," since it involves perception and imagination, and it is 'frequently' imagistic. Such feelings "come and go and vary in intensity." [16] The famous example which Goldie presents as the imagistic characteristic of feeling towards is: one can be in the grip of the sorts of visual imagination involved in sexual jealousy.[17] He argues that the imagistic feature of emotion can be applied to emotions whose objects are of a less visually imaginable nature. According to him, in the cases of emotion which lack a visual character, we are disposed to have episodes of feeling towards with a visual or imagistic character.[18] Yet, his famous example in his later work shows that the imagistic characteristic is not actually an important element of feeling towards. Let's consider the example:

> You are in an audience at a conference and a new speaker takes the stand. A friend next to you observes that you are becoming increasingly restless; your fingers are drumming on your notepad, your foot is tapping, and your lips and jaw are tense. Your friend surmises, rightly, that you are becoming irritated by something about the speaker: his manner, what he is saying, or something. But you are not aware of this. You have not noticed that you are feeling irritated by the speaker, yet you do have feelings of irritation toward him. Then your friend passes you a note, asking what is irritating you; and *then* you notice, or become aware, that you are feeling this emotion. Before seeing the note, you had feelings of irritation toward the

14. I owe this example to Goldie (2000), p. 19.
15. Goldie, Ibid., p. 73.
16. Ibid.
17. Ibid., p. 74.
18. Ibid.

speaker but were not aware that this was so.[19]

In this example, when the audience member has feelings of irritation towards the speaker, do his episodes of feelings towards have an imagistic character? If he feels irritation towards the speaker's gestures and grimaces, we can say that he may have feelings of irritation towards the speaker. However, if the audience member has irritation towards what the speaker is saying, does it involve visualizing? I think the answer is no. Let us consider another example. Golfers are sometimes told to think of the club as pendulum. This may involve visualizing, but it need not; the advice may be acted on by simply handling the club in a certain way, refraining from trying to control how it swings. I assume this is not imagistic, yet it is what the teacher, who cares nothing for your mental images, is hoping for.

Now one might resist the terminology, 'feeling towards', since it is not necessarily either propositional or evaluative, both of which are necessary for explaining the connection between evaluative judgement and response. But if this were true, one might argue that Goldie cannot explain the intimate tie between evaluation and action in terms of 'feeling towards', for in trying to think of X as Y with feeling, there seems to be no propositional content. This is because many believe that in order to explain action, we need to posit evaluative judgements, for example beliefs and desires, which have propositional content. Confronted with this objection, Goldie replies that 'feeling towards' can in fact explain the tie between our evaluation and action, since feeling towards has intentionality, which is directed towards the object of emotion in the world. Suppose that you visit a zoo. Looking at a lion in the cage, you are thinking of the lion as dangerous, but you do not feel fear, because it appears to be safely behind bars. Then

19.Goldie (2002), p. 242; see also Goldie (2004), pp. 96-7.

you see that the door to the cage has been left wide open. When you see this, suddenly you put your thoughts – the lion is dangerous and the cage is open – together. Your way of thinking of the lion as dangerous is different from how it was prior to your noticing the open door. Once you see that the door to the cage is left open, dangerousness is emotionally relevant to you. If this is so we can say that the earlier thought before you saw the open door, 'That lion is dangerous', differs in content from the new thought in that the new thought occurs with a feeling. Now in feeling fear towards the lion you are emotionally involved with the world. You are poised for action due to the emotion.[20] Hence you may run away. In this way, Goldie writes, feeling toward can explain the tie between recognition and response.

Although Goldie attempts to show with this example that feeling towards has intentionality, it is not obvious in what sense feeling towards has the intentionality ascribed to it, because thinking of lion as dangerous might simply be a matter of bodily behaviour, such as locking the door, rather than entertaining propositions. One might raise a further question: what is the difference between the earlier and later way of thinking the thought, 'that lion is dangerous'? One might say that what is new is a set of new beliefs or desires, such as the belief that the door is open and realisation that the lion could get you, which makes it reasonable to close it or hide. However, Goldie would say that the difference is not a new set of beliefs, but the fact that we are now emotionally involved with the world, and poised for action.

If this is so, another question arises: does the fact that you are emotionally engaged with the world always dispose you to move to act? Suppose that you read about what

20. A similar example is presented by Goldie. See Goldie (2000). p. 61.

Caligula did and feel angry toward the injustice he perpetuated. The propositional content of your emotion, anger, might be that Caligula is an awful person. However, seeing Caligula as an awful person does not by itself provide an end for action, and it won't dispose you to move to act. Since in this case there is nothing you can do in order to change the world in such a way that Caligula no longer appears awful to you. You may think what he did ought not to have happened, but realise that you cannot do anything about it. Let's apply this idea to another example: my hoping my favourite team will win their upcoming match. If, as Goldie argues, the fact that you are emotionally engaged with the world dispose you to move to act, how then does my emotion, in this case, –hope- move me towards action? I may not do anything to put it right. Rather, my desire involves an attitude toward (a pro-attitude) the winning: I want my team ought to win, but realise that I cannot do anything about it. If this is right, Goldie's idea that in feeling toward something you are emotionally involved with the world and you are poised for action due to the emotion is problematic. These considerations seem to show that feeling towards does not necessarily entail our being emotionally involved with the world, nor our being poised to act.

How then can feeling towards explain cases of recalcitrant emotion?: believing that something is not dangerous, yet at the same time, feeling afraid of it. According to Goldie, recalcitrant emotion is possible, when the tie between recognition-response is weaker. If someone fears grass snakes, Goldie could say that via his or her fear a person is unreflectively emotionally engaged with the world. In this type of case we can say that the tie between recognition and response is weak or broken. Thus we can say that the breaking or weakening of the tie between recognition and response when one is unreflectively emotionally involved in the world is what makes recalcitrant emotion

possible. Why then does the breaking or weakening of the tie between recognition and response happen? This is because of the alleged passivity of feeling. Many philosophers argue that we are sometimes passive in the sense of being unable to control our emotions. Goldie's answer to the above question is that we should consider the circumstances, character-traits and moods that the person has. He writes: "someone can be educated to recognise situations as emotion-invoking, and to respond with appropriate motivations, and can still, on an occasion, fail to be so motivated."[21] Thus Goldie claims that "our emotions, moods, and character traits, broadly conceived, can interweave, overlap, and mutually affect each other."[22] What then about the standard of the appropriateness of those emotions? The appropriateness of a person's emotions can be found "through considering the history of his psychological development and not every time a situation presents itself." [23] Goldie seems to argue that emotional appropriateness differs from truth. However, in his recent work, he seems to betray some ambiguity on this point by claiming that 'our emotional dispositions can, so to speak, *attune* us to the world around us' and appropriate emotions 'enable us to get things right.'[24] The lack of clarity arises due to the fact that such idiom as 'getting things right' is usually used in the case in which the truth predicate is available, while emotions are not capable of being true or false. Thus, in order to explain the appropriateness of emotion, he should show a way to explain the idea of 'getting things right' without being committed to having the epistemic warrant. However, he remains rather unclear about this matter.

21. Ibid.
22. Goldie (2000), p. 235.
23. Ibid., p. 37.
24. See Goldie (2004), p. 99.

3.1.5. Bodily Feeling

One might raise a question: how can Goldie handle bodily feeling? When we say that emotions involve feeling, Goldie writes, this means both 'feeling towards and bodily feeling. For example, when I say that I *feel* fear of a snake, first, there is the bodily feeling: my body shudders and my hands sweat. Second, I have a feeling of fear towards the snake, which is inseparable from bodily feelings, but not identical with them. What then is it that holds bodily feeling and feeling towards together? Goldie argues that both are inextricably linked by virtue of their phenomenology and intentionality. In order to appreciate this, let us consider an example.

> A man, while walking down a path sees a snake. He suddenly realizes that the place is inhabited by snakes. Then the place stimulates his fear.

Here we can say that there are two types of mental state: a phenomenological and intentional one. The man evaluates the situation as dangerous. So he says to himself, "I'd better be very careful: watch my step; and keep my eyes peeled." The intentional state, that is, the thought, gives itself as a maxim of behaviour which, if the man follows it, will lead him to avoid the snake. The other mental state, the phenomenological, also contributes to the evaluation in virtue of its particular 'feel.' The moment would be one of terror. In the grip of terror the man trembles and quivers, and he will get out of the wood as fast as his feet can carry him.

Now to answer the above question, namely, what it is that holds bodily feeling and feeling towards together in our minds when we have emotional experience, Goldie maintains that in the phenomenology of emotion, we experience bodily feelings and feeling towards almost as one. Goldie locates the root of this idea in Descartes, since

Descartes remarks that in the emotions, mind and body are "very closely joined and, as it were, intermingled."[25] The way in which bodily feeling and feeling towards are intimately connected is as follows: when we have bodily feeling, "it is essentially phenomenological, but it is also essentially intentional, being directed towards a part of one's body as being in a certain condition."[26]

Goldie calls the bodily feeling 'introspective knowledge' in the sense that it reveals "our thoughts, emotions, and the condition of our bodies," whereas he calls feeling towards 'extraspective knowledge' in the sense that it reveals "things about the world beyond the bounds of our bodies."[27] He goes onto argue that a bodily feeling is intentional in the sense that "the feeling is directed toward an object, one's body, as being a certain way or as undergoing certain changes."[28] Furthermore, he claims that the bodily feelings can provide prima facie reasons for believing that one is experiencing a certain sort of emotion, and for believing that there is something in the environment that has the related emotion-proper property. However, bodily feelings alone do not reveal much about the object of one's emotion.[29] For bodily feeling can only tell you that "there is *something* in the environment that has a certain property, such as the property of *being frightening.*"[30] It does not provide an epistemic route for us, like belief, to the object of the emotion *as such.* Suppose that you have heard some strange sounds while you were sleeping, and it suddenly wakes you up. You may feel that the hairs on the back of your neck are standing up, and this feeling can tell you that there is something frightening nearby. But this feeling cannot tell you that this something is a *burglar*. Now in order to

[25]. Descartes, *Meditation,* VI, CSM II, p. 56.
[26]. Goldie (2002), p. 249.
[27]. Ibid., p. 92.
[28]. Ibid., p. 93.
[29]. Ibid., p. 96.
[30]. Ibid. p. 94.

explain the object of fear in this case, Goldie holds that there must be the other kind of emotional feeling, which is directed at the specific object itself, in this case, for example, the fear directed at the burglar. The other feeling, which is directed toward the object of one's emotion *as such* is what Goldie calls 'feeling towards.'[31]

3.1.6. Goldie's Puzzle

With regard to the example which has presented the above we need to ask whether feeling towards is simply another type of bodily feeling, which arises as a consequence of the more basic revelation of the condition of one's body, as opposed to being a distinct kind of state. Furthermore, one might wonder if feeling towards can provide an epistemic route from bodily feeling to a belief about the object of our emotion as such. Goldie does not offer a clear answer to these questions. Instead, he presents a rather puzzling remark: "feeling toward is unreflective emotional engagement with the world beyond the body; it is not a consciousness of oneself, either of one's bodily condition or of oneself as experiencing an emotion."[32] Goldie himself presents conference example as we have seen before, in order to illustrate this characteristic of feeling towards. He provides the example to show that one can be unreflectively emotionally engaged with the world. It shows someone undergoing an emotion without being reflectively aware of having that emotion.

If in undergoing feelings towards we are unreflectively emotionally engaged with the world, Goldie writes, this has the advantage of enabling us to explain recalcitrant emotions. That is, how is it one can, without being irrational in the sense of having

[31]. Ibid. p. 96.
[32]. Ibid.

conflicting beliefs, feel afraid of something and yet, at the same time, believe that it is not dangerous. The person's recalcitrant fear is possible through his or her unreflectively regarding something as dangerous, due to his or her childhood trauma or character-traits, for example. If so, we can say that the difference between bodily feeling and feeling towards, if there are any differences, is as follows: as far as phenomenology is concerned, bodily feelings are reflectively emotionally engaged with the condition of one's body, whereas feelings towards are unreflectively emotionally involved in the world beyond one's body. Now as far as intentionality is concerned, feelings towards are feelings which are directed towards the object of our emotion as such. On the other hand, Goldie writes that 'bodily feelings' have a restricted form of intentionality in that they are directed only toward the condition of the body. The bodily feelings 'borrow' their intentionality from the intentional states, such as beliefs, that they occur in conjunction with. Thus another way of specifying the difference is as follows. Bodily feelings are intentional – they are directed at the condition of the body – and they have a cognitive element. Feelings towards are intentional – they are directed at the world - but they lack a cognitive element. These two sorts of feelings, according to Goldie, are "'united in consciousness' in being directed towards its object: united 'body and soul,' 'heart and mind.'"[33] Even if we accept this view, we can say that there need not be a conceptual connection other than the fact that both are needed for something to be an emotion.

I am sympathetic to Goldie's view in that he notes that certain feelings are inextricable from experience of the world and have directedness towards things. However, I do not accept his claim that there are two distinct kinds of feeling, one of which is merely

[33]. Goldie (2000), p. 55.

bodily and has a restricted intentionality. Goldie himself recognizes that these are potentially problematic issues, and claims that "the bodily feeling is thoroughly infused with the intentionality of emotion; and, in turn, the feeling towards is infused with a bodily characterization."[34] These characterizations of bodily feeling and feeling towards support the claim that feelings of the body and feelings towards objects in the world are two sides of the same coin.[35] Yet he still holds a distinction between two kinds of feelings when he argues that bodily feelings have only a borrowed intentionality, while feelings towards have intentionality intrinsically. He argues that the claim that all emotional feelings are of the same type is dogma.[36] However, as Rattcliffe points out, we can subsume this distinction under the single category of feeling towards, some of which have parts of the body as their object, and some of which have parts of the world as their object.[37]

Pace Goldie, if bodily feeling and feeling towards are not distinct, that is, if they are two sides of the same coin, then the real problem of explaining emotion is how we can develop a thorough account of how emotions dissolve the distinction between thoughts and feelings, or cognition and affect. Hence the discussion which follows focuses on a view which addresses this problem well.

[34]. Ibid., pp. 56-7.
[35]. A similar point has been made by M. Rattcliffe. See M. Rattcliffe (2005b), p. 47.
[36]. Goldie, Ibid., p. 57.
[37]. c.f. Rattcliffe (2005b), p. 47.

3.2. Embodied Appraisals

I have argued in Chapter 1 that for William James, emotions, on the one hand, are inner responses to bodily changes, and on the other hand they represent matters of concern; Prinz, following James, takes emotions to be embodied in the sense that they are inner response to bodily changes. I have also claimed that Prinz has developed the latter idea – that emotions represent matters of concern - into one holding that emotions represent core relational themes, such things as danger and loss. Now if we combine the above two elements, we have a so-called 'an embodied appraisal theory.' According to this view, emotions are, on the one hand, the internal states that register bodily changes, that is, 'embodied states.' On the other hand, "the embodied states represent core themes because they have the function of being reliably caused by core themes."[38] Prinz suggests that we should take the term 'an appraisal' in a different way from judgmentalists. According to Prinz, "if we define an appraisal as any mental state that represents an organism-environment relation that bears on well-being, then the embodied states in question qualify as appraisals."[39] Unlike judgmentalists who define an appraisal as an evaluative judgment, Prinz regards it as "any representation of an organism environment relation that bears on well-being."[40] Thus, if we understand an appraisal in this way, Prinz argues, the embodied states are 'embodied appraisals.' Now cognitivists might say that emotions can reliably co-occur with core relational themes only if they contain judgments. In order to appreciate this point, let us consider the snake example again. According to the embodied appraisal account, our story goes as follows: "when we see a snake, our bodies enter into characteristic patterns that are

[38] Prinz (2003), p. 80.
[39] Ibid.
[40] Prinz (2004a) p. 57.

registered by the embodied states that I have identified with emotions. There is no judgment involved, just a snake image in the early visual system."[41] We also have the same body pattern triggered by several other kinds of images, such as bugs, darkness, threatening faces, and blood. These things have a common effect in the sense that they instantiate a common property, that is, danger. "The embodied state is a danger detector, because danger is the property that gets the items in this hodgepodge to have an impact." Prinz appeals to an evolutionary story to explain the origin of such bodily patterns. Things like snakes and spiders were, he claims, a threat to our ancestors and so now these things cause our hearts to beat faster and cause our palms to sweat.[42]

Now one might wonder whether for embodied appraisal theory, judgments do or do not play a role in our emotional experience. According to Prinz, in the initial state, "judgments play no role," since an embodied state such as "danger detection is entirely noncognitive." We may have a danger concept, once we have experienced bodily patterns occurring in dangerous situations, and may apply the concept to the situations that would rarely be recognized by causal observation. Even if we can apply the cognitive element, for example, the concept of 'danger' to the causally unobservable cases (e.g., judging that the newly elected politician is dangerous) the emotion, in this case fear, which is triggered by a cognitive appraisal, is derivative, since "explicit judgments come after a meaningful emotion already exists."[43] In this respect, Prinz argues that the judgments are not something that we have in the initial state, but one that we have later, once we experience the bodily patterns. As a result, he argues that judgments are not an essential element of emotions, even in those cases where

[41] Prinz (2003), p. 80.
[42] Ibid., p. 80.
[43] Ibid., p. 81.

judgments in fact occur alongside emotions. Judgments, are simply caused, not constitutive.

Let us look at how the embodied appraisal account explains our problem: dissolving the dichotomy between feelings and judgments. The embodied appraisal theory suggests a solution to our problem as follows. The theory, on the one hand, agrees with feeling theory in the sense that emotions are embodied, and disagrees with feeling theory in the sense that judgments are needed for emotion elicitation. The theory, on the other hand, agrees with cognitive theories in the sense that emotions represent core relational themes. According to Prinz, this explains why emotions interact with thinking. "If emotions represent core themes, thoughts pertaining to those themes will be rationally tied to emotions." "Thoughts that provide evidence that one is in danger warrant fear, and fear warrants thoughts about strategies for coping with danger."[44]

Now one might wonder how the embodied appraisal theory can explain the problem of intentionality. In explaining this problem, two ideas are important. Firstly, how emotions have formal objects. Secondly, how emotions attain particular objects. The former can be explained in terms of the core relational themes that emotions have the function of reliably detecting. The latter can be explained in terms of the fact that emotions get linked to representations of particular objects. According to Prinz, formal objects are 'core relational themes' – dangers for fear, losses for sadness, insult for anger, and so on. In this respect, we can say that the formal object is not merely a perceptual property but the norm by which each emotion type defines its own standard of fittingness, as De Sousa points out.[45] Particular objects, according to Prinz, elicit the

[44] Ibid., p. 81.
[45] See De Sousa (2002).

emotion because they can be appraised as exhibiting the core theme expressed in the emotion. If we adopt these two ideas, we can explain how feelings and thoughts combine to constitute different emotions. For example, when we stand in fear on the edge of a tall cliff, we can say that the perception of that bodily state represents danger, since "it is under the reliable causal control of dangerousness." "Danger is the property in virtue of which these highly desperate eliciting conditions have come to perturb our bodies."[46] In this case, our mental state "contains a representation of the particular object (high place) as well as a representation of the property that makes it fearful (the danger of falling)." Hence Prinz claims the object represented by an emotion falls under a concept representing a core relational theme. A particular emotion monitors the body and detects salient phenomena such as danger, threat and loss. "Emotions are gut reactions; they use our bodies to tell how we are faring in the world."[47]

In short, according to embodied appraisal theory, emotions are "structurally simple embodied states, but they carry the kind of information that full-blown cognitions can carry." Hence Prinz argues that "cognitive theories have been right about content, and noncognitive theories have been right about form."[48] Now I want to raise an objection to this embodied appraisal theory. The typical objection to the theory is that the embodied appraisal theory cannot explain the full range of emotion we experience. The point has been raised by cognitivists that embodied appraisal theory cannot explain disembodied emotions such as loneliness or aesthetic appreciation, and long-standing emotions such as love. Furthermore, it has been argued that the embodied appraisal theory cannot explain cases of moral emotions such as guilt, regret, and shame. If this is right, we can

[46] Prinz (2004 a), p. 55.
[47] Prinz (2004 b), p. 69.
[48] Prinz (2003), p. 82.

say that embodied appraisal theory does not generalize, across the entire spectrum of emotional phenomena. In order to meet this objection, Prinz draws a distinction between dispositional and occurrent states and he takes long-standing emotions to be dispositions. As we have seen in Chapter 1, if I say that I love my husband all the time, my love in this case, is an enduring disposition to have occurrent states of love. An occurrent state of love, according to Prinz, "is an embodied reaction of the kind one has when one encounters the object of one's love."[49] The reason why this state is embodied is because being in such an occurrent state shows that "brain activation in limb areas associated with bodily response when subjects viewed pictures of their lovers."[50] In this respect, Prinz suggests that we should take long-standing love to be such an embodied state, since when I say that I love my husband without experiencing cuddly tenderness in relation to him, for example, it seems to be nonsense. Although the embodied appraisal theory explains long-standing emotion by distinguishing dispositional and occurrent states, one might still doubt whether the theory can explain calm passions, such as aesthetic appreciation or moral emotions. Confronted with this objection, as I have shown before in Chapter 1, Prinz appeals to the basic emotion theory, according to which some emotions are basic and others blend these basic emotions together, for example, contempt may be anger plus disgust; thrills may be joy plus fear. Guilt may be sadness plus the thoughts about my transgressions. Obviously some 'basic' emotions – fear, hate, anger, wonder – are fundamental as an equipment of survival in a vulnerable species. However, as we have seen in Chapter 1, the history of philosophy shows that the examples of the basic emotions do not remain constant. Hence I suggest that in order to explain moral emotions, we should take account of them in the socio-cultural,

[49] Prinz (2003), p.83.
[50] Ibid.

political, and economic activities. In Chapter 5, I shall discuss this issue with regards to socio-cultural context, without, however, addressing political and economic aspects. Before providing my alternative, in what follows I shall discuss an alternative view to solve this difficulty of the embodied appraisal theory by considering a Humean naturalism.

3.3. A Humean Naturalism

I argue in this section that the Humean account can solve the difficulty that the embodied appraisal theory encounters, namely, explaining calm passions and moral emotion, by placing the passions in their very concrete natural, biological and social contexts. According to A. Baier, naturalist philosophers are "those who take umbilical cords, navels, and other basic anatomical reminders of our mammalian condition more seriously than they take our immortal and sinful souls."[51] Baier defends naturalism by this remark, but she does not defend reductionism, that is, naturalism which identifies emotions with those basic or primitive ones reserved for Darwinist zoologists and biologists. Now in order to be a non-reductionist naturalist one must, I argue, take seriously not only our animal nature but also our social and cultural condition, considering them as essential constituents of the natural facts that should be taken into account.

When we read Hume's theory of passions as naturalistic, the important moral is twofold. According to Hume, our will-power does not control our passions, since for him the will, like passions, is the impression of reflection. The reason why for Hume the will is impression of reflection is because: first, it is not derived from sensation of the external object but occurs in us; second, the will for Hume, like all other impressions does not have independent active power. The will for Hume, as Baier notes, "is always determined by felt motive or 'impressed' sensory evidence, its apparent freedom in some cases being no more than ones' consciousness of internal causes (of one's own passion and concurring beliefs) without any current consciousness of external

[51] Baier (1990), p.439.

determinants."[52] Hence Hume says: reason is powerless and should be governed by the passions.[53] Hume argues that the way in which our emotions or passions are controlled is in terms of our emotional sensitivity and reactive responses.

Now Hume also recognizes that there are calm passions, which are more constant and fit to govern than others. Calm passions produce little emotion in the mind and are "certain instinct originally implanted in our natures, such as benevolence, the love of life, and kindness to children; or general appetite to good, and aversion to evil."[54] These calm passions are, according to Hume, the superior principles, that "cause no disorder in the soul." "When our passions are calm," as A. Baier points out, "there will be no internal 'combat'," since calm passions motivate us without any discernable turbulence in the soul or in the body. However calm passions may become violent if there are the violence-inducing factors among them.[55] Violent passions such as anger are ones that do cause emotion[56] and disorder.[57] Calm passions are often confounded with reason,

---

[52] Baier (1986), p. 53.
[53] Hume T. 415.
[54] T. 417
[55] See a wonderful analysis of calm and violent passions in Baier (1991), pp. 167-71.
[56] Hume takes the term "emotion" as a bodily disturbance, like some pleasures and pains. To that extent he seems to regard it as an "impressions of sensation" rather than as an "impressions of reflexion." Now question arises as to the relation between emotions and desires. Confusion about the nature of emotions is, according to Baier, reflected in the historical vicissitudes of the words used to speak of it. As Baier notes it, while 'emotion' used to mean 'violent passion', we now seem to use 'passion' to mean 'violent emotion.' See Baier (1990), p.1.
[57] The passions are next divided into the direct and the indirect. The direct passions arise directly from good, evil, pleasure, or pain. Examples are "desire, aversion, grief, joy, hope, fear, despair and security."(T. 277) Here what is interesting is that he classifies desires as direct passions in that they "arise from a natural impulse and instinct which is perfectly unaccountable": "these passions strictly speaking produce good and evil, and proceed not from them like other affections," (T. 439). Indirect passions arise from good or evil plus other qualities or ideas. That is, with the indirect passions, "the conjunction of other qualities" (T. 276) is also needed. The other quality is the relation of the cause to the object. What then are indirect passions? Indirect passions are listed "pride, humility, ambition, vanity, love, hatred, envy, pity, malice, generosity, with their dependents" (T. 276f). What is interesting is, in the case of the indirect passions, like pride, there is a 'foreign object' (T. 287) as well as oneself, the proper object; while in the case of the direct passions, the same thing plays both roles. For example, in joy, the enjoyed thing is both cause and object. So the only quality it needs to possess is being a good. The indirect passions not only have a cause (an impression which excites them), but also an object (to which the passion directs our view when excited) (T. 276).

since they produce little sensible emotion. What Hume calls strength of mind however is not due to reason, but merely to the prevalence of one kind of passions (calm ones) over another (violent ones).[58]

Now a question arises: why for Hume reason should obey passions, and which passions it should serve?[59] In reply to this question on behalf of Hume, Baier presents two points: first, in order for passions to take over a ruling role from reason, they must be lasting and second, in order to last, some "dependable order must be established among them."[60] Now in order to lend support to this idea, Baier notes two features of Hume's theory of passions. The first feature that Baier focuses is "the genetic feature of being a 'corrected' sentiment." The second one is what Hume calls 'purity'. Let us first consider the second one. Pride and humility are for Hume "pure emotions in the soul, unattended by any desire, and not immediately exciting us to any action"(T 307). According to Hume, pride is pure emotion in that it is free from accompanying desire. It has its own impetus or desire. Furthermore, pride is self-sustaining in existence by a self-directed desire, that is one directed at oneself. In this respect, Hume says that unlike love and hate they are "completed in themselves," they "rest in the emotion they produce." In order to show the mechanisms whereby pride achieves self-sustaining, Hume presents an analogy between the instruments of music and the passions. Hume argues that the mind is "not of the nature of a wind instrument of music, which in running over the notes immediately loses the sound after each stroke the vibrations still retain the same sound, which gradually and insensibly decays"(T. 440-1). According to Hume, some

[58] Hume, T. 417-8.
[59] This question has been raised by Baier in order to defend Hume's thesis that reason must be slave of passion. See Baier (1980), pp. 406-7.
[60] Baier (1980), pp. 406-7.

passions, like pride, are reverberated and strengthen. They are sustained, they increase and grow when other repeat, reflect or echo them, like instruments of orchestra joining in and strengthening melody played by the soloist. What Hume means by this musical metaphor is that the proud man sustains his pride by the recognition of his worth which he finds in others. His self-esteem is reinforced by other people, when it is mirrored by them. Hume writes:

> In general we may remark, that the minds of men are mirrors to one another, not only because they reflect each others emotions, but also because those rays of passions, sentiments and opinions may often be reverberated, and may decay away by insensible degrees. Thus the pleasure, which a rich man receives from his possessions, being thrown upon perceiv'd and sympathiz'd with, increase the pleasure of the possessor; and being once more reflected, become a new foundation for pleasure and esteem in the beholder.[61]

The proud person, hence, can sustain his pride only when others let him have it – either in terms of their approval - admiration and recognition - or in terms of their disapproval – envy, or in some other way.

Now we can say that these 'approval and disapproval' are moral 'sentiments' which can take over the ruling role from reason. Now let us consider another feature of Hume's theory of passions which Baire notes: the genetic feature of being a 'corrected sentiment.' If emotions can be revised at all, it is only through the influence of the approving or disapproving responses of those who are affected by them and whose reactions matter to one. Our behaviours cause other's emotional reactions, and the reaction make us reflect on and revise them. This is because as Lilli Alanen points out "we are sensitive to the ways our fellow human beings respond to our own emotional

[61] Hume, T. 365.

reactions, and this mutual sensitivity is a pre-condition for moral agency."[62] But it is not modified by the power of our own rational judgment or will alone, but changed only through "attuning our emotional responses and behaviour reactions."[63]

According to Hume, our moral sentiments inevitably "vary according to the distance or contiguity of the objects," even when our moral judgments do not. Hence in order for morality to be impartial, Hume must say what is appropriate for impartial others to feel about an acknowledged wrong. We can say that it makes no sense for you to be angry about killing people in the Iraq war. But this would be absurd. Hume would agree that how angry it makes sense for me depends on what the war has to do with me. But this does not mean what happens far away is not morally important. By contrast, war is equally wrong no matter where it occurs. At this stage, contemporary expressivists argue that moral norms are norms for what it makes sense to feel from 'an *impartial* but *fully engaged* standpoint.' Hence we can say that though it makes no sense to get angry at the distant wrongdoing, anger would be warranted were we fully engaged. This view is similar to Hume's. Hume suggests that in order to correct for the effects of immediacy on our responses and to prevent what he calls "continual contradictions," in moral reflection "we fix on some steady and general point of view; and always, in our thoughts, place ourselves in them, whatever may be our present situation."[64]

However, as I shall argue in Chapter 5, to take the moral point of view, namely, that of a fully impartial and fully engaged spectator seems to be impossible for mere humans due to a number of distorting factors, such as prejudice and partiality to oneself. Yet,

[62] Alanen (2003), p. 332.
[63] Alanen, Ibid.
[64] Hume, T. 581-2.

Hume implicitly says that a fully impartial and fully engaged observer is possible by our emotional changes. But it is not modified by the power of our own rational judgment or will alone, but changed only through the influence of approving or disapproving responses of those who are affected by our emotion. He claims, for example, that though a beautiful face may not provide so much pleasure when seen at a distance, Hume says we "correct its momentary appearance" and judge it no less beautiful, because we know what effects it would have on us if brought closer. He goes on to argue that "such corrections are common with regard to all the senses." When Hume talks about matters of empirical fact, the favourable conditions for making judgments are the clear-cut observation of the empirical fact. Yet when he talks about corrections, Hume argues that it is possible in terms of other's disapproval. If this is so, we can say that to judge an act is wrong is to judge that it calls for disapproval from others. But the trouble, as we shall see in Chapter 5, is whether other's disapproval has objectivity or not. I shall argue that although contemporary expressivist, for example, Gibbard's 'others' anger' may not be impartial, Hume's 'approval' or 'disapproval' are close to impartial. Yet, I shall argue that if the Humean appeals to a vacuous notion of an impartial observer, then it can be said that they admit the following point: to make an evaluative judgment is not to '*have*' but to '*endorse*' a sentiment, since appeal to an ideal observer is to accept rationality of the sentiments. If this is right, the Humean should show how to justify a sentiment when one makes an evaluative judgment. I shall discuss this issue in Chapter 5. Before discussing this issue, in the next Chapter (Chapter 4), I shall show another kind of crude view which reduces emotion to belief or desire. I shall demonstrate that this kind of crude view is implausible, using Michael Smith's analysis of the concept of direction of fit.

Part II: Emotion, Value, and Personal Identity

Introduction

In Chapters 1 and 2, I have challenged two crude views: the *feeling-centred theory* in Chapter 1 and *judgmentalism* in Chapter 2. Chapter 4 aims to demonstrate that a crude view of emotion, namely, *a Humean functionalism*, which reduces emotion to belief or desire is implausible. In order to avoid such reductionist views, I have suggested two approaches to emotion. The first alternative I have considered, was a qualified cognitivism (Chapter 3). However, I have argued that qualified cognitivism does not provide a satisfactory account of emotions. This is because, Goldie, in distinguishing two sorts of feelings, fails to resolve the problem of how bodily feelings and thoughts combine to constitute different emotions. If bodily feeling and feeling towards are not distinct, as I have argued, then the real problem of explaining emotion is how we can develop an account of how emotions dissolve the distinction between thoughts and feelings, or cognition and affect. Thus, I have demonstrated a view which explicitly addresses this problem.

I have then demonstrated how an embodied appraisal account overcomes our problem, that is, how bodily feelings and thoughts combine to constitute different emotions. The embodied appraisal theory suggests a solution to our problem as follows: the theory, on the one hand, agrees with feeling theory in the sense that emotions are embodied, and disagrees with feeling theory in the sense that judgments are needed for emotion elicitation. The theory, on the other hand, agrees with cognitive theories in the sense that emotions represent core relational themes. According to Prinz, this explains why emotions interact with thinking. If emotions represent things like danger and loss, that is, core relational themes, thoughts pertaining to those themes will be rationally tied to

emotions.

However I have raised an objection to this embodied appraisal theory. The typical objection to this theory is that it cannot explain the full range of emotion we experience. It has been raised by cognitivists that embodied appraisal theory cannot explain disembodied emotions such as loneliness or aesthetic appreciation, and long-standing emotions such as love. Furthermore, it has been argued that the embodied appraisal theory cannot explain cases of moral emotions such as guilt, regret, and shame. If this is right, we can say that embodied appraisal theory does not generalize, across the entire spectrum of emotional phenomena.

In Chapter 3, I have shown that Hume's theory of passions can explain long-standing emotions such as pride and love. Furthermore, Hume's theory can account for cases of moral emotions. My interpretation of Hume, following Baier's lead, is naturalistic, since I believe that Hume is a naturalist who places the passions in their very concrete natural, biological, and social contexts.

In Chapter 5, I shall also show that the emotions play an important role in our lives as moral agents. Descartes, as many understand it, is a cognitivist about emotion, in spite of his emphasis on the role of bodily feeling or sensation. It has been argued that Descartes ignores the role of emotions in their biological and social contexts, by analyzing the passions as atomistic, individual mental states, and by stressing their effects on the individual will and their expressions in the individual body.[1] For Descartes, the ultimate aim for well-being concerns the God who has created all things, including eternal truths. It follows that the standards of truth and goodness are not ones

[1] A more recent writer who makes this point. See Alanen (2003).

we can interact with in any way.[2] Hence it can be said that the individual agents are alone responsible for where they go, alone in their choices and alone also in their joys and miseries. Yet Hume places our faculties in the context of the biological and social facts. We can say that this is a turn from Descartes's conception of passions as atomistic cognitive states of individual souls to passions as socially sustained, socially reinforced and corrigible mental and behavioural phenomena.

I have claimed in Chapter 3 that the social aspect of emotion provides an alternative to the affect and cognition accounts. I shall attempt in Chapter 5 to demonstrate this account in terms of the response-dependent theory of emotional properties. In Chapter 5, I shall also present my alternative to meet the difficulty that the embodied appraisal theory encounters. As I have mentioned in Chapter 3, my alternative takes moral phenomena to be natural and socio-cultural phenomena. Chapter 5 shall offer a naturalistic picture of moral evaluation by associating it with certain emotions, namely, guilt and anger.

In Chapter 6, I shall explicate the social aspect of emotion in terms of the problem of personal identity by considering Hume's theory of personal identity as regard in Book Two of his *Treatise.* According to Hume, the necessity of the transition from Book One's solitary self to Book Two and Three's social self is concerned with his thought that the puzzle raised in Book One cannot be resolved by thought. Seen in this light, it can be said that Book Two's transition from solitary reason to social passions provides a way to dissolve the intellectual puzzles of Book One concerning the self.

In Book Two of the *Treatise,* Hume attempts to solve the problem of personal identity

[2] As argued by Alanen (2003), p. 333

raised in Book One. He does this by attempting to unify the diachronic expansion of a person's life in terms of its influence on emotion, and by addressing the self from the perspective of social world. However, I argue that this view is also problematic. Hence, in Chapter 6, I suggest an alternative way to solve Hume's problem. The deep problem, it seems to me, is rooted in Hume's view of memory, since his view in Book Two presupposes the view of memory as a mental image which lends support to the episodic self. In order to address the self which is extended through past, present and future, I suggest that Hume must take notice of the notion of 'the experiential content of memory' which is embedded in certain *feelings*.

The fact that experiential memory plays a central role in explaining personal identity has been neglected by many philosophers in the mental connectedness tradition. Hence in Chapter 6, I attempt to show the role of experiential memory for explaining the unity of a person's life. I shall show that experiential memory is forward-looking as well as backward-looking in orientation. In order to do this I will challenge the view of memory as retrospective only with reference to a temporal direction.

I shall argue that if we adopt the notion of experiential memory to identify mental connectedness with personal identity, we bring the person under the influence of his past, such that the past is recreated and then influences the future. I shall then show that in reconstructing our past life with experiential memory, *emotions*, especially moral emotions - for example, pride, self-love, conceit, shame, regret, remorse and guilt - are interconnected. This will prove that, as we have seen in the previous Chapter, especially Chapters 3, and as we shall see in Chapter 5, emotions play an important role in our lives as well as in our self-understanding as moral agents.

Chapter 4: Do emotions have direction of fit?: Against Reductionism

Preliminary Remarks

In Chapters 1 and 2, I have shown that reductionism which reduces emotion to feeling or evaluative judgement is implausible. This Chapter aims to show that a reductionism which reduces emotion to belief or desire is problematic. In order to do this I shall present a series of arguments against reductionism in terms of the notion of direction of fit. I argue against a conception of emotion which attempts to view it essentially in cognitive terms, in particular in terms of belief and desire.

There are two interpretations of the direction of fit understanding of beliefs and desires: a descriptive and a normative interpretation. In this Chapter, I, first of all, resist the descriptive account on the grounds that it does not explain other intentional states, especially, emotions. Then I consider whether a normative account may prove more successful and suggest that the normative alternative to the descriptive account fares better in that it shows us why emotions have neither direction of fit. Then I argue that the normative account is not promising either in that it cannot explain how to distinguish between truth and satisfaction. Finally, I argue that the norm of belief and desire is rationality, whereas the norm of emotion is response-dependent one. If so, it is a mistake to attempt to give an account of all these, namely, belief, desire and emotion, in terms of rationality.

4.1. Belief/Desire and a Humean Functionalism

In explaining emotion, there are three kinds of crude view, which reduce emotion to feeling, evaluative judgment/ belief or desire respectively. In this Chapter I shall argue against only the latter two types: a conception of emotion which attempts to view it essentially in cognitive terms, in particular in terms of belief and desire. The first position I want to discuss in this section is a Humean functionalism, according to which emotion is reducible to desire. In order to illuminate this, in what follows, I shall utilize the concept of 'direction of fit'. The concept of direction of fit is often attributed to Anscombe, for example by Platts.[3] In Anscombe's story, there are two men, a man going shopping with a shopping list, and a private detective who tries to write down what the man buys. Anscombe contrasts a mistake in the performance of someone buying from a shopping list, with a mistake in a record which a detective is keeping of the shopping purchased. Then she asks "What is the difference between the man's shopping list and the detective's list?" Her answer is as follows:

> It is precisely this: if the list and the things that the man actually buys do not agree, and if this and this alone constitutes a *mistake,* then the mistake is not in the list but in the man's performance... whereas if the detective's record and what the man actually buys do not agree, then the mistake is in the record.[4]

In this passage, Anscombe suggests that the difference between belief and desire is similar to that between the detective's list and the shopping list. As Danto remarks, the shopper may make a mistake by putting in the basket something that is not on the list – beer when the list says 'beef'; the detective may make a mistake in writing 'beef' when

[3] See Anscombe (1957). See also Platts (1979), p. 256. L. L. Humberston suggests that it originate with J. L. Austin, See Humberston (1992).
[4] Anscombe (1957), p. 56.

the basket contains beer. The shopper corrects the mistake by replacing the beer with beef; the detective by writing 'beer' instead of 'beef'.[5] This way of contrasting beliefs and desires in terms of 'direction of fit' is, as Platts rightly claims, highly metaphorical,[6] but many critics believe that the metaphor is very useful in illuminating the role of beliefs and desires in explaining action, especially, reasons for action'.[7] According to this idea, desire is aimed at some end, whereas belief plays a role in achieving that end. This is what Smith calls teleological explanation of action, according to which an agent's reasons are explained by the agent's attempt to achieve some purpose or goal. Smith argues that this is so because desires have 'world-to-mind' direction of fit.[8]

What then is 'world-to-mind' direction of fit? Smith believes that 'having a motivating reason just is, *inter alia,* having a goal' and 'the having a goal is being in a state with the direction of fit of desire'.[9] But the trouble with this account is that it faces a circularity problem. For in order to explain having a motivational reason, it appeals to the idea of 'having a goal'. Then having a goal is defined as 'being in a state with which the world must fit'. Now in order to explain the concept of direction of fit, one appeals to 'having a goal'. So the argument seems to be circular. In order to avoid this difficulty, Smith appeals to a functional account of direction of fit, dropping the idea of goals or purposes.[10] He writes:

---

[5] See Danto (1999), p.56.
[6] See Platts (1979), p.257.
[7] See e.g., Smith (1994), Ch.4; see also Searle (2001).
[8] Smith discusses this in his earlier article: 'The Humean Theory of Motivation'(1987, pp. 36-61). In this paper I focus on his later discussion (1994).
[9] Smith 1994, 116.
[10] The circularity problem has been pointed out by G. F. Schueler (1991) and Nick Zangwill (1998). In order to avoid the circularity problem, L.L. Humberston (1992)has appealed to second-order background intentions.

> For the difference between beliefs and desires in terms of directions of fit can be seen to amount to a difference in the functional roles of belief and desire. Very roughly, and simplifying somewhat, it amounts, *inter alia*, to a difference in the counterfactual dependence of a belief that p and a desire that p on a perception with the content that not p: a belief that p tends to go out of existence in the presence of a perception with the content that not p, whereas a desire that p tends to endure, disposing the subject in that state to bring it about that p.[11]

There have been detailed criticisms of Smith's account.[12] However, in this Chapter I focus on some difficulties arising from applying this formulation to *emotions*. Although Platts glosses Anscombe's view as that only desire is 'a prime exemplar' of those states having world-to-mind direction of fit, Smith is invested in thinking that the account does cover emotions, since he says that states that include having a goal will count as desires on his broad construal.[13] In what follows, I shall argue against a conception of emotion which attempts to view it in cognitive terms, in particular in terms of belief and desire.

4.2. Emotions and Direction of Fit

4.2.1. a Humean Functionalism

It might be thought that *hoping* that p counts as a desire for Smith and is an emotional state. If so, one might say, hope will have to have, with desire, a world-to-mind direction of fit. The reason why hope, like desire, could have world-to-mind direction of fit is that 'he hopes the team wins but he doesn't want them to' is contradictory.

[11] Smith (1994), p. 115.
[12] E.g., See Schueler (1991); L. L. Humberston (1992); N. Zangwill (1998); David Sobel and David Copp, (2001); Robert Dunn (2004). Schueler and Humberston's target of criticism has been Smith's former discussion -'The Humean Theory of Motivation,' (1987) - whereas recent writers focus on his later version-*The Moral Problem* (1994). I too focus on his later discussion.
[13] Smith (1994), p. 117.

Another reason that hope could have the same direction of fit as desire, one might argue, is because hopes aim at satisfaction rather than at truth. If my hope fails to 'fit' the world, that is not any defect in my hope but in the world. However, the trouble is that if we take hope as having world-to-mind direction of fit like desires, it cannot cover some kinds of hope close to cognitive states. For example, if I know that it is raining, then I can't hope that the road is not wet, though I can wish it were not. When we hope, we also believe that it is possible that something we can do will 'make it more likely that this hope (or other pro-attitude) gets realized.' We cannot have a hope or desire that cannot possibly be satisfied, unless we are committed to irrationality. If this is right, then according to Smith's functional account of direction of fit, 'all hopes will have, like beliefs, a mind-to-world direction of fit,' since 'I can't hope that *p* once I discover that *not p*'. [14] If this is the case, we can say that hope that p, like beliefs, tends to go out of existence when we perceive that not p.

But the problem with seeing hope as having a mind-to-world direction of fit like belief is shown by cases of 'recalcitrant' belief. We can compare the recalcitrant mental state with optic illusions of the Muller- Lyre lines: the Muller- Lyre lines continue to appear to be of different lengths while they are known to be equal lengths. In this respect, many people try to develop this idea to show how, at times, our beliefs, feelings and actions seem to behave with a mind of their own despite our best efforts in trying to control them. This is because, S. Döring argues, the emotional content "resembles the content of sense-perception in that both kinds of representational content need not be revised in the light of belief and better knowledge."[15] She attempts to show this, using an analogy

[14] As argued by Schueler (1991), pp. 279-80.
[15] S. Döring (2003), p. 223.

emotion with optic illusion of the Muller-Lyre lines as follows: just as our perception that the two lines differ in length persists in the face of our belief that the lines are the same length, so emotions may persist even though we have relevant and countervailing knowledge. Thus, we can say that a person can have a sense-perception or can feel emotion that p, and at the same time believe that –p. However, this is not the case in believing that p. If this is true, it follows that an emotion, like a sense perception, is not an attitude that can be considered as true or false as such. The point that Döring makes here is that in such cases of sense-perceptions and emotions we are able to be in a contradictory or ambivalent state, while in the case of belief, this is impossible.[16]

Having elaborated on the analogy between emotions and sense-perceptions, now I want to argue against a cognitive conception of emotion which attempts to view it essentially in terms of belief. The crude cognitivist, as we have seen in Chapter 2, for example, Solomon and Nussbaum, might say that emotions have the same direction of fit as beliefs, namely, that an emotion has a mind-to-world direction of fit. Now if we apply this idea to Smith's the above formulation, we can gloss it in functional terms as follows: an emotion that *p* is a mental state that tends to go out of existence in the presence of a perception with the content that not *p*. If I believe that there is nothing to be afraid of in the presence of a perception with content that not *p* (e. g., seeing a rabbit is not harmful), then I should not fear it. If I do, it follows that it is a mistake to fear *p* when not *p*, and the fear should be abandoned. But it is possible that fear persists, even when I consciously believe that the rabbit is harmless. If emotions imply, like Solomon and Nussbaum say, evaluative belief, we would have to be said to hold inconsistent beliefs as follows: an evaluative belief (fear) tends to *endure* in the presence of

[16] See Döring (2004), pp. 264-9.

perception that not *p* (the rabbit is harmless). If this is true, is it a mistake, and should it be abandoned? But in actual life we have those emotions; consider the case of phobia. If a crude view which reduces emotion to belief were right, it could be said that the man or woman, who does not believe that the rabbit is particularly dangerous, cannot be afraid of it, since according to the view, the relevant belief is a necessary element of the emotion. However, the crude cognitivism cannot explain the above cases which are examples of emotional recalcitrance. Hence it follows that emotional recalcitrance gets strong cognitivism into trouble.

Let's apply recalcitrance to the case of hope which is supposed to have mind-to-world direction of fit, like belief. I might still hope, for example, that my team will win even though all the evidence is to the contrary. The Labours genuinely hope they will win the next election. Should the Labours give up their hope that they will win? I'm afraid, the answer is no, because if they did, they would be sure not to win. This raises difficult issues about rationality, which are beyond the scope of this Chapter. The point of this objection is to ask whether the cognitive element involved in the emotions is necessarily a state of *belief.* Indeed, cognitive theorists do not always claim that the cognitive element, which is essential to the emotions, should be 'belief.' Belief, among the many cognitive attitudes, is somehow a 'very strong' kind of attitude, for belief itself is necessarily connected with evidential rationality or justification. I do not have a belief when I have no or insufficient relevant evidence. In other words, if cognitivists claim that the cognitive element, which is essential to the emotions, is belief, there is a risk of ruling out a number of 'irrational emotions.' In this respect, cognitivists substitute 'judgment', 'evaluation', 'appraisal,' 'apprehension', or simply 'construal' for the belief. In rejecting the crudely cognitivist approach by assessing whether emotions have the

kinds of directions of fit found in belief and desire respectively, I hold the cognitive aspect of emotion to be construal rather than belief - the latter is 'cognitively penetrable'[17] unlike the former and unlike emotion.

Another difficulty with Smith's functionalist account of the 'direction of fit' concerns the latter part of the formulation of desire: 'a desire that p tends to endure, *disposing the subject in that state to bring it about that p.*' Suppose that you read about what Caligula did and became angry at the injustice of it all. The propositional content of your emotion, in this case, might be that Caligula is an awful person. But seeing Caligula as an awful person does not by itself provide an end for action, and it won't dispose you to move to act, since in this case there is nothing you can do in order to change the world in such a way that Caligula no longer appears awful to you. I think what he did ought not to have happened, but realise that I cannot do anything about it. Let's apply this idea to another example; my hoping my team will win. If, as Smith argues, an emotion can be reduced to a merely functional state [that is a desire] that disposes the subject toward action, this might involve a desire to put it right. How then does my emotion –hope- move me towards action? I may not do anything to put it right. Rather, my desire involves an attitude toward (a pro-attitude) the winning: I want my team ought to win, but realise that I cannot do anything about it. If this is right, Smith's view, according to which emotion can be reducible to desire gets into trouble.

Now one might argue that the latter part of Smith's formulation - 'disposing the subject in that state to bring it about that p' - has another difficulty when we apply it to the case

[17] I owe this terminology to Peter Goldie. I think 'cognitive penetrability' and 'cognitive impenetrability' are helpful to explain our emotional experience. Our emotional experiences are cognitively penetrable only if they can be affected by our beliefs. On the other hand, our emotions are cognitively impenetrable when they are not affected by our persistent beliefs. Goldie uses the terminology 'cognitive impenetrability' as the same meaning as 'recalcitrance.' See Goldie (2000), pp. 74-8.

of hope about the past. Imagine a case of hope about the past: I hope that I did not behave idiotically at last night's party. In this case, although my hope tends to endure when I know that I behaved idiotically, it won't dispose me to bring it about that I behaved charmingly. For I can do nothing *now* to make it the case. If I regret my behaviour, in this case, it may dispose me to bring it about that p; I may try to apologize to the host and to avoid the people who I met at the party. But in the case of my hope that I behaved charmingly, I won't be disposed to make that which I hope for the case, even though my hope does not tend to go out of existence. Hence in this case we cannot say that hope has the same direction of fit as desire. In response to this kind of objection, Smith might say that these can be looked upon as (future directed) hopes about what will turn out about the past. If this is so, then Smith's test still seems to work in this case.

I have demonstrated so far that a direction of fit account cannot deal with emotion, for example, hope. The best way to handle hopes on a direction of fit view would be to say that they involve having a desire for *p* plus the belief that it is still possible that *p*. In general, I have attempted to show, using Smith's analysis that it is implausible that all emotions have the same direction of fit. A fear is inappropriate if the mind doesn't fit the world, and a hope is wrong if the world doesn't fit the mind. For some emotions, we strive to bring about their objects and, for others, we try to avoid their objects. In what follows, I shall use the idea of recalcitrant emotions to show why emotions have neither mind-to-world, nor have world-to-mind direction of fit.

4.2.2. The Normative Account

So far we have seen the difficulties that a Humean functionalism faces. Now let us

consider whether the normative account proves more successful. According to Searle, "all intentionality has a normative structure," and "all intentional states which have a normative structure have a motivational force."[18] He writes:

> If an animal has a belief, the belief is subject to the norms of truth, rationality, and consistency. If an animal has intentions, those intentions can succeed or fail. If an animal has perceptions, those perceptions either succeed or fail in giving it accurate information about the world.[19]

If emotional evaluations were understood as evaluative beliefs, as the crude cognitivist understands them, the normative understanding of our emotions is the way we take things to be. Then the responsibility of our emotions to match the world constitutes their mind-to-world direction of fit. If we take emotions to be reducible to desire, on the other hand, we can make this claim in normative terms as follows: these emotions set an onus of match on the world by being part of desire, since the emotions as intentional states, as Searle notes, 'presuppose that the propositional content of the emotion is already satisfied.'[20] In order to appreciate this, let us take Searle's example.

> If I am overjoyed that France won the World Cup, I simply take it for granted that France won the World Cup. My joy has as its propositional content that France won the World Cup, and I presuppose that the propositional content matches reality.[21]

In this passage, Searle seems to suggest that although emotions presuppose that their propositional content is already satisfied, they in themselves represent neither how I believe the world to be nor how I want it to be. Instead it can be said that the state being had is a result of the propositional content matching reality. If this is true, emotion's

[18] See Searle (2001), p.182.
[19] Ibid.
[20] Ibid., p. 38
[21] Ibid.

having direction of fit is epiphenomenal. In this spirit, Searle regards the intentional state of emotion as having 'the null or zero direction of fit,'[22] since one's joy or grief, love or hatred, is neither true or false, nor satisfied or frustrated. If this, in turn, is true, it can be said that emotions neither have mind-to-world direction of fit, nor have world-to-mind direction of fit. In what follows I shall use the idea of recalcitrant emotions to show why emotions have neither mind-to-world, nor world-to-mind direction of fit. I shall discuss cases of pride and jealousy which will provide vivid illustrations of this.[23]

[22] See Searle (2001), p. 38; also (1983), pp. 8-9

[23] I am not trying to discuss jealousy and pride in general. I have picked out instances of a kind that are likely to be rated as counter examples to the view that emotions have either mind-to-world (belief) or world-to mind (desire) direction of fit.

4.3. Emotional Recalcitrance and Personal Perspective

One might say that emotions may have either mind-to-world or world-to-mind direction of fit, due to their involving either belief or desire. There are cases in which emotions sometimes have belief-like directions of fit. For example, 'pride' sometimes does not involve desire but belief. My pride at my daughter's achievements depends on the belief, evaluable as true or false, that she has had them. When I am proud of my beautiful house it is because of my belief that the object is mine. This is 'cognitively penetrable.' In being proud of my beautiful house, I first of all must believe that it is valuable; secondly, in order for the feeling to play a role I must believe the house to be in some way connected with me. G. Taylor calls those two beliefs 'explanatory' and 'indentificatory' belief, respectively.[24] The 'explanatory' beliefs just explain the relation between the valuable things and the person, whereas 'identificatory' belief refers to something 'closely' related to the person who feels pride. Thus, according to Taylor, 'a person may hold the requisite explanatory beliefs and yet not feel proud.' 'She may regard her beautiful house as a most desirable possession but may not regard this as reflecting on her own worth.'[25] Thus in order to feel pride there must be indentificatory belief that 'the agent regards the desirable as something she herself has brought about.' That is, she must regard the information given by explanatory beliefs as her worth.[26] But if we accept this view we cannot explain the following case: in the case of the triumph of the team which I support, pride may involve 'explanatory belief,' but not

[24] Taylor (1985), p.27.
[25] Ibid., p.34.
[26] This is, according to Taylor, a sufficient condition for pride.

involve 'identificatory belief,' since I cannot regard the team's victory as one that I myself brought about. Thus in my view, the pride in the triumph of the team does not derive from belief but from my trying to think of the team's victory as mine. In trying to think of the teams' victory as mine, there is no 'fit' or 'directionality, since trying to thinking of X as Y is subject to one's will. This is one of the difference between 'thinking of X as Y' and 'belief', since as Goldie notes, 'believing at will is, as is generally accepted, impossible; one cannot directly try to believe something; at best one can indirectly (albeit irrationally) try to come to believe that thing by, for example, partly oneself in an environment where one is likely to do so'.[27] If this is so, we can say that the pride in the triumph of the team is derived from my attitude of thinking X as if Y. If this is true, in this case, we can say that pride does not have belief-like direction of fit.

My pride, one might argue, presupposes a desire that my daughter has achievements, a desire with satisfaction-conditions and the world-to-mind direction of fit. But that pride does not itself have satisfaction conditions, for it does not itself set goals for action. We could say that pride sometimes involves mind-to-world (belief), sometimes world-to-mind (desire), or sometimes neither directions of fit. It follows that it is difficult to say that emotions have either direction of fit. The reason why emotions sometimes have neither direction of fit is because we have the personal point of view when we experience emotion. We have different standards of fittingness when we experience emotion. Hence the appropriateness of the emotions can be stated from the agent's perspective. Now in what follows, I shall discuss a case of jealousy which will provide

[27] Goldie (2000), p. 72.

one vivid kind of illustration of this.[28]

If we take jealousy as having the same direction of fit as desire, that is, world-to-mind direction of fit, then it could be said that jealousy aims at satisfaction. What then is satisfaction of jealousy? For example, I might just wish my rival ill or dead. I may choose the way in which I confront my rival: a fight or to work hard. If I have low-esteem, I may be depressed or self-destructive. But this emotional reaction cannot give me ultimate satisfaction. One might say that the ultimate satisfaction of one's desire involved in jealousy extinguishes that desire. One way to terminate the desire might be to get rid of a belief that the rival is superior to the agent. If this were true, then one might argue that jealousy has the same direction of fit as belief, for the desire involved in jealousy rests on belief.[29] In what follows I want to show why we cannot say that jealousy has belief-like direction of fit.

If we take emotions as having a belief-like direction of fit, we cannot explain kinds of malicious envy, or jealousy, for these cases are sometimes 'cognitively impenetrable.' Suppose that Kate and Lucy have been friends and they assess themselves by their success in some competition. They apply for a job at the same place. But Lucy is accepted and Kate fails in the final interview. So Kate is jealous of Lucy's success. In being jealous of Lucy, Kate thinks of Lucy as an enemy to her self-esteem. So despite their friendship, Kate wishes Lucy is ill. Kate wants Lucy to be absent from the interview due to the illness.[30] This is because, Taylor would say, Kate's jealousy involves false belief, and is morally wrong belief that rests on irrationality and cannot

[28] In this discussion I owe much to Peter Goldie (2000), pp. 224-40.
[29] Taylor, for example, argues that jealousy is unjustified, since this emotion must include a false or unjustified belief. See Taylor (1975), pp. 401-2; see also Ben-Ze'ev (1990), p. 506.
[30] R. Robert presents a similar case to mine. See R. Robert (1991).

be justified. Now the presumption of this kind of strong cognitivist's inference, it seems to me, that there is an analogy between our emotional response and the world, and a true belief and the world. According to them, the former relation can be described as 'fittingness', whereas the latter one is maintained by truth. Then the cognitivists argue that these attitudes are both rational. Since they take both to be rational, they argue that emotion can be justified or criticized. Furthermore they claim that justification or criticism of the emotions is always committed to morality or prudence. Taylor claims that 'an emotional reaction is unjustified if it rests on irrationally mistaken beliefs or when it is disproportionate to a given situation'.[31] But this is, we can say, what some call a 'moralistic fallacy.'[32]

In my view, the reason why Kate feels jealousy is that she construes Lucy's success as diminishing herself, not that she believes it is diminishing. Thus, we can say that Kate's jealousy is derived from her attitude of trying to think of Lucy's success as diminishing herself. But Taylor may say that Kate believes that it is diminishing, and she takes the proposition [that Lucy's success is diminishing] to be true or false. Then Taylor goes on to argue that Kate's belief is a false, unjustified and morally wrong one. Now the presumption of this view, it seems to me, is that moral evaluations of an emotion are concerned with its fittingness, and emotion's unfittingness means it is morally wrong to feel an emotion. If this were true, we could think that morally impermissible feeling amounts to unfitting feeling. But this is a philosophical error, since we can never argue from an emotion's rightness to its fittingness, even though such an argument sometimes

[31] Taylor (1975), p. 393.

[32] See D'Arms and Jacobson (2000a), p. 73. 'The moralistic fallacy', according to D'Arms and Jacobson, is 'simply to infer, from the claim that it would be morally objectionable to feel F toward X, that therefore F is not a fitting response to X.' The reason why this inference is fallacious, as they say, is that 'an emotion can be fitting despite being wrong to feel.' (D'Arms and Jacobson 2000a, pp. 73-4). If this is right, we go wrong when we take evaluative sentences as entailing imperatives.

is valid, since as D'Arms and Jacobson notes, 'the wrongness of feeling an emotion never, in itself, constitutes a reason that the emotion doesn't fit.'[33] For example, when we feel that something is funny, the feeling may differ among us. Something that appears to me funny does not seem to amuse others. I may laugh at an offensive joke, but it cannot be said that it is morally wrong, and that it is never funny. If we infer from the fact that to laugh at an offensive joke is morally inappropriate to its unfittingness, then it is, as D'Arms and Jacobson say, a 'moralistic fallacy.'

We can apply this idea to our jealousy case: Kate is jealous of Lucy's achievement. I may say that Kate's jealousy is inappropriate, since Lucy made more of an effort, and deserved her success, while Kate did nothing. If my evaluation were right, I could say that it is morally wrong for Kate to be jealous of Lucy. Although my judgment is right, it cannot be right to assess the appropriateness of Kate's jealousy morally, since jealousy does not represent its object as bad or unjust. In my view, jealousy is comparable to the following cases: the snake is not dangerous and at the same time I fear it; we believe that the Muller-Lyre lines continue to appear to be of different lengths while they are known to be of equal lengths. As I have said before, this implies that emotion, like sense-perception can present in the face of relevant and countervailing knowledge. In this respect, I argue that emotion which do not vanish in the face of better knowledge is recalcitrant emotion. This is because to some extent our emotions and emotional responses are passive, and cannot be controlled. There are several senses of being unable to control an emotion: 1) being unable to helping feeling it. (e. g., feeling sorry for a burglar); 2) being unable to help expressing it (e.g., the witness in court bursting into tears); 3) being afraid of a grass snake, despite believing

[33] D'Arms and Jacobson (2000a), p. 74.

that it poses no danger. The third case is one of recalcitrant emotion, which is the emotion over which the agents have no control, even though they believe that p. In this respect, we can call jealousy recalcitrant emotion.

Although one might doubt that the jealousy is always recalcitrant, we can say that Kate's case is recalcitrant in the sense that her emotion and emotional response are passive and cannot be controlled in the face of relevant and countervailing knowledge.[34] This is because the subject's attitude toward those involved in jealousy is very complex and the desire in which jealousy is involved depends on all kinds of things: the circumstances, character-traits and moods that the person has.

[34] The familiar example for emotional recalcitrance is phobia: the phobic who is afraid of something despite believing that it poses little or no danger. In support of my view that emotions are not easily reducible to belief, desire or feeling and we cannot say that they have either directions of fit, I choose pride and jealousy which are tricky of saying whether they have one of directions of fit.

4.4. Conclusion

In this Chapter I have rejected the idea that there is a correspondence between our emotions and the world. If we accept this idea, then we are conceding to the view that emotions are capable of having epistemic warrant. Many contemporary writers regard this kind of warrant as the concept of appropriateness or fittingness that is taken to be analogous to truth in the emotional realm.[35] Yet, if we allow an analogy between appropriateness and truth, it would seem to allow that emotions are capable of being true or false.[36] However, many philosophers have circumvented the concept of truth in the emotional realm, for there are some emotions that cannot be reduced to propositional attitudes which are eligible for being truth-apt, unlike beliefs, thoughts, and judgments. I have demonstrated so far these cases in terms of, for example, recalcitrant emotions, phobia, pride and jealousy. Especially, I have shown these cases by utilizing the notion of 'direction of fit'.

I have also presented a series of argument against a crude view of emotion, namely, a Humean functionalism. My reason for rejecting the crude view rests on a kind of holism, since as Goldie notes, 'our emotions, moods, and character traits, broadly conceived, can interweave, overlap, and mutually affect each other'.[37] Hence, when we apply this holism to our case of jealousy, the evaluation of a person's disposition to be jealous can be made relative to that person's narrative: the situations, his or her other emotions, his moods and his character. In evaluating a person's emotion, the person's character and

[35] See e.g. Goldie (2000); D'Arms & Jacobson (2000a); Nussbaum (2001).
[36] In this respect many contemporary emotion theorists tend to be committed themselves to the claim that emotions are able to be true or false, when they say such things as the appropriate emotions 'enable us to get things right'(e.g., Goldie 2004, p. 99).
[37] Goldie (2000), p. 235.

mood may play an important role. When you laugh at an offensive joke, the funniness of the joke might be understood by you in the light of your delighted mood after a delightful day, or in the light of your general disposition to be cheerful. If this is right, the emotions' appropriateness or inappropriateness can be said to depend not on the belief which is true or false, but on a construal depends on personal perspective.

We have specifically seen that Smith's descriptive account of direction of fit fails to explain mental states other than belief and desire, especially, emotions. Furthermore, Smith's descriptive account cannot explain whether an emotion is fitting, for this is a normative question. When we believe p, our attitude is regulated by epistemic norms. On the other hand, I have argued that emotions' fittingness or unfittingness depends neither on the belief which is true or false, nor on the desire which is satisfied or frustrated. Instead, I have argued that emotions' fittingness or unfittingness depends on the construal which depends on personal perspective. Hence, I argue, it is a mistake to give an account of all these, namely, belief, desire and emotion in terms of rationality.

Chapter 5: Moral Emotions and Qualified Sentimentalism

Preliminary Remarks

Recently, emotion has attracted much attention in moral philosophy. Many people think that human values depend on human valuing, which is inextricably connected with affect, sentiment, and emotion. We may call this approach sentimentalism. The most notable exponent of sentimentalism is Hume. Hume argues that "a) moral distinctions are not deriv'd from reason, but b) are in some way a matter of 'sentiment.'" I shall argue that sentimentalism is not an epistemic doctrine; rather the basic idea is metaethical. Thus the discussion in this Chapter focuses on metaethical claims.

Following Hume's lead, simple subjectivists argue that one's emotions are essentially involved in evaluation. I argue that simple subjectivists ignore Hume's remarks that "in order to pave the way for a sentiment,...it is often necessary...that much reasoning should precede" (E. 173). If this is so, those who want to take Hume's view should allow reasoning, as well as feeling to play a role in evaluative judgment.

Having established this claim, I argue that the central challenge for sentimentalism is to hold that there are both sentimental and rational aspects of evaluation. I call this Neo-sentimentalism. Neo-sentimentalists take value to be dispositional by analogy with colour. If we take colour concepts to refer to dispositional properties, we can say that an object is red, for example, iff it looks red to normal observers under standard conditions. If we apply this view to value we can say that for some X to be φ (shameful, funny, enviable, etc.) would be for normal observers under standard conditions to feel F

(shame, amusement, envy) at X.[1]

The problem for this dispositional account is that neither one's own dispositions, nor the dispositions of the majority capture the normativity of evaluative judgments, for it suggests that our attitudes are defective and can be revised by the proper use of knowledge, real capacity for sympathy, and so on. I take this dispositional analyst to be revisionist. Furthermore, proponents argue that this improvement constitutes an increased sensitivity to φ. I shall discuss some objections to this revisionary thesis and attempt to meet those objections.

I shall then argue that value judgments are not apt for truth and also argue that evaluative properties are not features of the world independent of human capacities and sensibilities. I shall demonstrate this in terms of Hume's metaphor of 'gilding and staining.' We shall then discuss how contemporary sentimentalists have developed this view. According to neo-sentimentalists, the phenomenology of valuing is such that sentiments purport to be sensitivities to evaluative properties. I characterize this claim as the problem of ascribing evaluative properties. On this view, the fearsome or the funny, for example, are not intrinsic properties of the world, independent of our sensibilities, but are sensitivities to features of the world. I shall call this view the response-dependence theory of emotional properties. According to this view, for an evaluative property ϕ to be manifested by X is for an associated response A to be appropriately held toward X. Given this, what is at issue is the problem of both ascribing evaluative properties and the correctness and incorrectness of responses.

I shall then discuss how to restrict considerations of warrant for an emotion.

[1] I borrow this formulation from D'Arms and Jacobson (2000 b), p. 726.

Furthermore, I shall present non-cognitivist approach of what it could be to think an emotion warranted. It has been argued among many philosophers that 'warrant' is an epistemic notion with reference to the evidential support of one's beliefs. If we apply this notion to emotion, we should be able to tell whether one has evidence for shame or amusement, for example. Yet, since I have supported a noncognitivist approach so far, holding that emotions are not even partly constituted by beliefs, I shall not claim that an emotion is warranted whenever some constitutive belief is warranted. Instead, I shall argue that emotional appropriateness grounds for feeling a given emotion. I shall then argue that if we apply this idea to moral emotion, the problem that remains is illuminating what gives rise to the right feeling. Given this, I shall argue that the real issue in explaining appropriateness of moral emotion is not to '*have*' but to '*endorse*' a sentiment. I shall present the views which meet this challenge, which I call qualified sentimentalism as an advanced view of Hume's sentimentalism.

5.1. Sentiment and Value: Phenomenology of Value

In Chapter 1 and 4, I have rejected the idea that there is a correspondence between our emotions and the world. If we concede to this idea, then we allow the claim that emotions are capable of having an epistemic warrant. Many contemporary writers regard this kind of warrant as the concept of appropriateness or fittingness that is taken to be analogous to truth in the emotional realm.[2] Yet, if we allow an analogy between appropriateness and truth, it would seem to allow that emotions are capable of being true or false. In the previous Chapter (Chapter 4), I have demonstrated that this claim is untenable. Many contemporary emotion theorists tend to be rather unclear on the idea that emotions are able to be true or false when they say that appropriate emotions 'enable us to get things right'[3] by 'properly tracking those properties [funny, shameful, etc.] of which they purport to be perceptions.'[4] The lack of clarity arises due to the fact that such idioms as 'getting things right' and 'tracking' are usually used in the case in which the truth predicate is available, while as I have discussed before, emotions are not capable of being true or false. Hence, in this Chapter I attempt to show a way to explain the idea of 'getting things right' without being committed to having the epistemic warrant. In order to illustrate this idea, in this section, I shall focus on 'sentimentalism.' Sentimentalism is not an epistemic doctrine; rather the fundamental claim is metaethical. A. J. Ayer claims that "in saying that a certain type of action is right or wrong, I am not making any factual statement, not even a statement about my own state of mind. I am merely expressing certain moral sentiments."[5] Now, the defenders of his expressivism

[2] e.g. Greenspan (1988);Goldie (2000); D'Arms & Jacobson (2000b); Nussbaum (2001).
[3] Goldie (2004), p. 99.
[4] D' Arms & Jacobson (2000a), p. 69.
[5] Ayer (1952), p. 107.

have attempted to provide a phenomenological argument concerning the defensibility of values and sentiments. Although I try to defend a phenomenological account of sentiments and value, I shall argue that there is no 'purely phenomenological' account of the sentiments, since the sentiments are not mere feelings. In the following sections, I shall support the argument of the most notable recent proponents of expressivism of values.[6] In order to present a phenomenological argument concerning the defensibility of values, I begin with Hume's view of values.

5.1.1. The Analogy between the Secondary Qualities and Values

Hume wrote of moral value:

> It lies in yourself, not in the object. So that when you pronounce any action or character to be vicious, you mean nothing, but that from the constitution of your nature you have a feeling or sentiment of blame from the contemplation of it.[7]

He introduces an analogy between moral approval/disapproval and the perception of so-called secondary qualities as traditionally understood:

> Vice and virtue, therefore, may be compared to sounds, colours, heat and cold, which, according to modern philosophy, are not qualities in objects, but perceptions in the mind: and this discovery in morals, like that other in physics, is to be regarded as a considerable advancement of the speculative sciences; though, like that too, it has little or no influence in practice [8]

Is Hume in this passage saying that moral sentiments are to be compared to the experiences of secondary qualities in that each at least seems to provide us with the

[6] There are two proponents of similar views. Simon Blackburn and Allan Gibbard to meet these challenges. I shall mainly support Allan Gibbard's view. See Blackburn (1984); and Gibbard (1990).
[7] Hume, T. 469.
[8] T. 469.

presence of objective qualities?[9] There seems to be a parallel between the way secondary qualities appear objectively in the world and how moral qualities appear objectively in the world.[10] Hume seems to have in mind John Locke's theory of secondary qualities. According to Locke, secondary qualities are not qualities in objects, like shape, size and solidity, but "powers which objects have to give rise to certain perceptions in the mind of human observers."[11] For example, according to Locke, the redness of the object is a power which the object has to give rise to certain perceptions in normal perceivers under certain standard viewing conditions. And Hume tries to apply this theory to morality. According to him, virtue is the power that certain actions have. This is because, firstly, it has certain characteristics involving actions. Secondly it has a power to give rise to feelings of approval in the minds of those *normal people* who observe or contemplate those actions under certain *standard conditions.*[12] Now on the basis of the above consideration, in what follows I take up the ways of reading Hume's view of moral value in recent philosophy.

Hume's view of moral value is subject to varying interpretation by recent philosophers. John Mackie, for example, believes that Hume holds 'an error theory' of secondary quality, whereas S. Blackburn takes Hume's view to be projectivism. Moreover, some neo-sentimentalist read Hume's view as 'rational sentimentalism' by holding the view

---

[9] It seems that many people agree with the view that objective qualities are usually meant mind-independent ones. However, some people among supporter of moral realism deny that the objective means 'mind-independent'. See Dancy (1986).

[10] The contemporary moral philosophers have also proposed the comparison of which Hume speaks. See: McDowell (1981) and (1985). Also Wiggins (1987); Blackburn; Crispin Wright also well address this issue. See Wright (1988), pp. 1-26

[11] Locke, *Essay*, II. viii.

[12] Here it should be noticed that the term 'power' is not paralleled with Locke's idea of natural necessary connection. In the *Enquiry* Hume rejects the Lockean conception of power (*Enquires concerning Human Understanding and concerning the Principles of Morals.* Hereafter *E.* 63-73). If we take Hume's view of the concept of power to be contingent constant conjunction, we can understand the idea that virtue is the power that certain actions have.

that there are both sentimental and rational aspects of evaluation.[13] On the other hand, other neo-sentimentalists take value to be dispositional by analogy with colour.[14] My intention in this section is to show the difficulties of applying all theses readings to Hume's sentimentalism. Let us consider these views in more details.

5.1.1.1. 'Error Theory' or 'Projectivism'

Hume's analogy between moral approval/ disapproval and the perception of so-called secondary qualities would be welcome to those who favour a projective view of moral judgment, or 'error theory.' J. Mackie holds 'an error theory' of secondary qualities and employs the analogy to promote a parallel theory of moral evaluation. Mackie observes that we mistakenly objectify moral properties, since in moral evaluation Hume attributes a property to an object that is distinct from its natural features.[15] The error theorist agrees with this Humean picture; but it adds that, since no such ethical properties exist, ethical judgments are always false. In other words, error theorists believe that when we make moral judgments we inevitably *project* a moral property onto the world, but this is a mistake. Let us take an example, i.e., Willful murder presented by Hume. When we think about this case, we have a response – outrage or condemnation. According to Hume the response is in us. Error theorists agree with this, but they add that we have a propensity to *objectify* it by taking it to be of a moral property in the world (the outrageousness or contemptibleness of the killing). We think of our response as the recognition of some further property that is independent of it and that our response registers. Mackie challenges any theory that postulates the existence

[13] See D'Arms and Jacobson (2000b), pp. 722-48.
[14] See McDowell (1985).
[15] See Mackie (1977).

of moral properties. According to him, if a theory postulates the existence of a property of wrongness, for example, no such account will be as comprehensible and as simple, and as plausible psychologically, as the idea that we are not perceiving a property of wrongness at all, but rather simply responding subjectively and negatively to the natural features in question.[16]

However, the problem for the error theorist, as S. Darwell nicely pointed out, will then be: "one might continue to have the illusory responses that one thinks ethical thought and conviction mistakenly objectify, just as a straight stick can continue to *look* bent in a pond even to a person who knows this appearance is mistaken." "But one can no more continue rationally to believe that anything has value or disvalue, or is right or wrong, and believe also that these beliefs are false, than one can coherently think that the bent-ness of the stick is really bent."[17] Hence if we believe the error theory, we can no longer rationally hold beliefs about value and disvalue, good and evil, right and wrong.

A projectivist allows the possibility of fallibility. Yet the projectivist goes beyond the error theorist. Simon Blackburn writes: "I think that X is good, but I may be wrong. Thus a projectivist can go beyond saying of our moral sensibility that it might change, to saying that it might improve and not only because of improving knowledge, but also because of improving reactions to whatever information we have."[18] The reason why we *erroneously* do so, according to projectivist, is that "we speak and think 'as if' the world contained a certain kind of fact, whereas the true explanation of what we are doing is that we have certain reactions, habits or sentiments, which we voice and discuss by such

[16] Ibid., Ch. 1, sec. 9.
[17] Darwell (1998), p. 64.
[18] Blackburn (1981a), p. 175.

talk." [19]

Moral realists, like McDowell, try to undermine the distinctive claim of projectivism about value, namely that value is not a *genuine* feature of persons, acts, states of affairs, etc., but only appears so because we mistake features of our evaluative responses for features of such things. For McDowell, projectivism is 'spurious, ' and he presents an alternative account.[20] According to him, if we think of secondary qualities in a broadly dispositional way, then we have to think of the colour, for example, as consisting in a disposition to bring about a certain sort of visual experience. "The moral quality of a situation is a certain sort of disposition to induce a certain sort of distinctive experience." For example, when we say or believe that 'X is good,' it can be said that being good consists in propensities on the part of good things to elicit observer's reactions of moral approval. Now let us look closely at the Dispositional Analysis.

5.1.1.2. Dispositional Analysis

McDowell thinks of the projectivist picture as a 'thin conception' of genuine reality. And he suggests that the metaphor of projection should be replaced by a comparison between our knowledge of ethics and our knowledge of other things, such as colours. He claims that evaluative 'attitudes', or state of will, are like colour experience. The idea of value experience is essentially subjective in much the same way as an experience of redness. We can understand it in terms of the appropriate modification of human sensibility. What, then, is the aim of the analogy between secondary qualities and colour in particular? When the analogists employ the conception of colour it is a so-

[19] Blackburn (1981).
[20] McDowell (1985), p.118ff.

called dispositional conception. The dispositional thesis of colour says that being red consists in looking red to some perceiver(s). We can formulate it as follows:

> "X is red iff X is disposed to look such and so (ostended) to standard perceivers as they actually are under standard conditions as they actually are." [21]

Colours, tastes, sounds, smells and textures are proper objects of our five respective senses. If we think of secondary qualities in the broadly dispositional way suggested by McDowell, then we have to think of the colour, e.g., of an object as consisting in a disposition to induce a certain sort of visual experience which, ceteris paribus, constitutes a perception of that very quality. And some, including McDowell, who have wanted to defend - at least- an analogy between moral and secondary qualities, seem to have had it in mind to recommend precisely the transposition of this combination to the moral case. According to the dispositional analysts, our faculty of moral judgment may be regarded as a *perceptual* capacity with moral qualities as its proper objects. If this view is right, we can say that a moral property, for example being good, consists in propensities on the part of good things to elicit observer's reactions of moral approval, that is, emotions. Justice or injustice applies to something if and only if it produces 'emotions', that is, 'sentiments of approbation' in observers. In the next section, I shall argue that emotions have the response-dependent properties. The response-dependent theory implies that emotions are elicited by things as they are related to us. I shall show some objections to the response-dependence theory of emotional properties, according to which the theory is viciously circular if those properties represent anything that causes fear, anger, disgust, and elation in us because it then looks as if those emotions can never occur in error. However, I argue that if we take emotions to be secondary

[21] I employ this formulation from Mark Johnston. See his paper 'Dispositional Theories of Value', *Proceedings of the Aristotelian Society* Supplementary Volume 13, p. 74.

qualities, we can escape this circle. Red things are things that cause us to have red experiences. This is not circular, because a red experience can be characterized by its distinctive feel, without mention of red things. If my argument is right, it can be said, as McGinn notes, that "just as secondary qualities predicated have sensory states as part of their satisfaction conditions, so moral predicates have certain affective or valuational psychological states as part of their satisfaction conditions: any account of what it is for an object to have a moral property will therefore make essential reference to certain psychological reactions in appropriately constituted beings." [22]

We can say that red is anything that causes normal humans to have red experiences in normal viewing conditions. If we have a red experiences as a result of pressing my finger against my eye for a few seconds, it doesn't mean that my finger is red, because finger-pressing is a not a normal viewing condition.[23] In a similar vein, we might say that fear represent secondary qualities. We might say fear represent the property of having the power to cause fear experiences in normal human beings under normal conditions. If we agree with the dispositional analyst that our faculty of moral judgment is analogous to a perceptual capacity, then we can say that just as an object is red if it would look red to a normal human observer in standard lighting conditions, so moral properties are analogous to them. That is, when we make moral judgments, we suppose that other people will share our moral attitudes, if they are in normal conditions and are making their judgment in normal circumstances. What then are the notions of 'normal conditions' and 'normal circumstances'? It might be said that just as we cannot see the true colours of things unless we have the appropriate sensory capacity, i.e., colour

[22] See, McGinn (1983), p. 147.
[23] As argued by Prinz (2004b), p. 61.

vision, so we cannot be sensitive to the moral qualities of people and their actions unless we have the appropriate moral capacity – a capacity to feel sympathy for the concerns of other humans. However, the problem with this account is that it permits error. In section 3 of this Chapter, I shall show how the response-dependent theorist handles this problem. Before looking at this issue, let us consider the disanalogy between emotion and the secondary qualities, for example, fear and colour. There is, as Prinz points out, an important difference between fear experiences and colour experiences. "The conscious feelings associated with mental states that represent colors are projected out into the world. When we experience redness, we experience it as if it were out there on the surfaces of objects."[24] But opponents of the dispositional analyst claim that this is not the case with emotions. As Prinz puts it, the feeling of rage, for example, is not projected onto the object of rage; it is experienced as a state within us.[25] In what follows, I shall show how the feeling of rage, for example, which is experienced as a state within us is projected onto the world, namely, the object of rage, for example, in terms of the metaphor of 'gilding' and 'staining.'

5.1.2. The Metaphor of 'Gilding' and 'Staining'

Hume argues that when we call an action virtuous we are saying that it arouses a certain feeling in us, that it pleases us in a certain way. He maintains that the vice and virtue of an action is not a matter of fact about the action that can be inferred from anything, nor a matter of fact that can be discovered by direct perception of an action, but that, rather, the vice and virtue of an action is a matter of moral sentiment. And he argues that the idea of vice or viciousness does not mark anything in the 'object.' Hume puts the case

[24] Prinz, Ibid., p. 63.
[25] Ibid.

starkly:

> Examine it in all its lights, and see if you can find that matter of fact, or real existence, which you call *vice.* In which-ever way you take it, you find only certain passions, motives, volitions and thoughts. There is no other matter of fact in the case. The vice entirely escapes you, as long as you consider the object.[26]

Here, Hume is denying that moral properties form part of the fabric of the world. Elsewhere, he asks us to imagine a case of willful murder. There is nothing in an act of willful murder, for example, like a 'property' of its being vicious. But we *think of* the action *as* being vicious. Then a question arises: how do we come to have such a thought? Here Hume's answer is that we can have the thought that an action is vicious when we *feel* a certain 'sentiment.' He writes:

> You never can find it, till you turn your reflexion into you own breast, and find a sentiment of disapprobation, which arises in you, towards this action. Here is a matter of fact; 'tis the object of feeling, not of reason. It lies in yourself, not in the object.[27]

If we never had such a 'sentiment,' according to Hume, we would never 'pronounce' anything to be valuable, desirable, or beautiful. We could describe this view in more concrete form in the following way: "It is a sentiment that occurs in the mind whenever we observe or contemplate actions or characters that have certain characteristics, and if we never got such feelings there would be no such thing as morality."[28] We may call this view projectivism.

According to the projectivist, if we feel that an action is vicious, we project the feeling onto reality. We see the world as though it had the 'objectified feeling,' and assert that it

[26] Hume, T. 468.
[27] Hume, T. 468-9.
[28] Stroud (1977), p. 179.

does. This projectivist reading of Hume, as many point out, is not only able to avoid the subjectivist reading of the emotivist, but also avoids depriving moral judgments of their truth-values. Although Stroud accepts that moral judgments have truth-values, he does not regard moral judgment as something that could be arrived at by reasoning, since sentiment, when we have the thought that some action is vicious, is not discovered by reasoning, but by being felt. Thus Stroud claims that moral judgments are judgments that ascribe objective moral characteristics such as goodness and rightness to agents and actions. The projectivist goes on to say that these objective moral properties do not really exist as properties of actions and agents, but are only products of our moral feelings. When we make moral judgments, we project onto actions and agents, thinking of these projections *as if* they were objective features of actions and agents, even though they are no such thing in reality. Here one might raise a question: "How [does] the appearance of one of those things in the mind have the effect of giving us 'ideas' of vice or virtue, of beauty or ugliness, of causation, or of any other qualities and relations we ascribe to objects, when according to Hume those qualities and relations do not and cannot actually belong to 'objects as they really stand in nature'?"[29] Hume's answer would be that "it is only because we naturally get certain feelings or impressions, and, even more importantly, only because of the mind's 'productive faculty' in 'gilding or staining' the world with what those feelings give us, that we ever come to think in those ways at all." The discussion which will follow focuses on the matter of why we need this mental operation, namely, projection. In order to do this, let us look at Hume's *Enquiry.*[30] Hume writes, of reason and taste respectively:

[29] This question is raised by stroud. Stroud, Ibid., p. 258.
[30] In his recent paper, Stroud (1993) supports his projection theory interpretation with Hume's metaphors of 'gilding and staining.' B. Stroud "Gilding or Staining" the world with "sentiments" and "Phantasms",

> The one discovers objects as they really stand in nature, without addition or diminution: the other has a productive faculty and gilding or staining all natural objects with the colours, borrowed from internal sentiment, raises in a manner a new creation.[31]

Now, we can say that the above passage could mean that the mind 'spreads' features of its internal workings or contents to an external world which does not really contain them. Furthermore, in the case of morals, our moral evaluations 'gild' the human world that they evaluate 'with the colours, borrowed from internal sentiment,' and 'raise in a manner a new creation.' What Hume wants to say with this metaphor, it seems to me, is that in our thoughts we tend to ascribe to objects something 'new,' certain qualities or relations which they do not possess as they really are in nature, 'without addition or diminution' by us. In other words, "we take something mental and see it as external."[32] Stroud calls this mental operation 'projection.'[33] Now the reason why for Hume the mental operation of 'gilding' or 'staining' the world is needed is because according to his view of the world, we cannot find the source of any thoughts that we might have, for example, the idea of necessary connection, the beauty of any objects, or the virtuousness or viciousness of any action or characters. Then the question is: how can we have the ideas or thoughts which do not represent 'any thing, that does or can belong' to the way things 'really stand in nature.' As long as Hume retains his theory of ideas, he cannot answer this question. If we adopt his theory of ideas, we cannot explain, as Stroud points out, "the notion of 'the world' or of the way things 'really stand in nature' which is supposed to exclude beauty and causal connections and the virtuousness or

---

Hume Studies, Vol. XIX, No.2, pp.253-272.

[31] Appendix 1 of the Enquiry concerning the Principle of Morals, p. 294.

[32] Stroud (1993), p. 260.

[33] Stroud (1977), pp. 86-7 and pp. 185- 86. See also Stroud (1993).

viciousness of actions and characters."[34] For Hume the mental operation of gilding or staining is only supposed to produce the idea (thought) of something that is not there.

What then is 'a new creation'? Hume would answer that it is *something* in the mind. That is, an impression – a 'sentiment' or feeling or an impression of reflection. "The source of that extra 'stain' or 'gilding,'" as Stroud observes, "could only be the mind itself, or its contents, so it is from there that we must 'borrow' whatever materials are used in the 'spreading' or 'transforming' operation."[35] Now a question arises: what do we 'borrow' from our internal impressions, and what do we ascribe to the external objects we 'gild or stain'? It is not the impression itself, but the perceivable or felt qualities of the impressions themselves that we ascribe to the external object. When we say that the act of willful murder is disgusting, we try to ascribe a feeling or impression to an act of murder. But here what we should take notice of is that in order for our feeling or thinking to involve the operation of projecting or spreading, the feeling or thinking should not be the mere feeling that an action pleases or displeases us in a certain way, but should be feeling or thinking that action is *virtuous or vicious*. If this is so, we can say that the operation of 'gilding or staining' is not initiated by a mere feeling but by "a feeling or impression which is 'of' something, or has an object, in the 'intentional sense.' " [36] What then is the intentional object of the feeling? Stroud argues that "it cannot be 'of' any object or quality or relation which could be part of the way things 'really stand in nature,' " since "if it were, no 'gilding or staining' would be necessary."[37] Thus, we can say that the intentional object of the feeling which the

[34] Stroud (1993), p. 269.
[35] Ibid., p. 260.
[36] Ibid., p. 267.
[37] Ibid.

mental operation of gilding or staining produces, are good and evil actions, admirable and contemptible characters, and beautiful and ugly objects. That is, intentional objects of the feeling that the mental operations of gilding or staining produce are objects being valued by evaluators. In the next section I shall show that this view has been developed for the thesis that emotions are the response-dependent properties.

The essential features of the projectivist's reading of the metaphor of 'gilding or staining' are as follows: reason first discovers and judges the relevant matters of fact, for example, willful murder; second, "the mind, from the contemplation of the whole, feels some new impression of affection or disgust, esteem or contempt, approbation or blame."[38] Now I want to raise an objection to Stoud's projectivism. Although Stroud's 'projection' theory is quite right as a reading of Hume's intention, one might object, arguing that there is a missing element in his interpretation of Hume's metaphors of 'gilding or staining.' The missing element is the fact that the thing stained or gilded is a real psychological trait, as real as any other quality.[39] This could be confirmed by Hume's undoubted maxim: Hume writes:

> It may be establish'd as an undoubted maxim, that no action can be virtuous or morally good, unless there be in human nature some motive to produce it, distinct from the sense of its morality".[40]

Here we can say that 'some motive' which produce actions are character-traits which constitute human nature as virtuous or vicious. If so, it might reasonably be said that virtues and vices are likened to "colours gilded onto some thing, furniture or natural

[38] Hume, E. 290.
[39] Baier (1991), p. 194.
[40] Hume, T. 479.

object that presumably has some colour already, before the staining or gilding."[41]

In short, we could read Hume's metaphors of gilding or staining as follows: "The human mind does have a great propensity to spread itself onto external objects, but in moral evaluation the 'internal sentiment' with which it gilds is spread on *internal,* not external object, since it is human minds and characters that are 'gilded' with an impartial human mind's approbation of them."[42] This view still might be regarded as projectivism. However it is not an 'error' theory of moral value. Since "there need be," as Baier remarks, "no illusion of something 'out there' when in fact it is really 'in here.'"[43] The way in which moral value, i.e., virtue, could *be 'out there'* is in the people in whom they are found. "The 'guilt,' or 'approval', for example, is in the approvers, but their actual approbation makes a real difference to the approved persons – it gives them a particular status in a moral community, and it may give them pride, or encouragement in conserving or strengthening the approved character traits."[44] In the next section, I shall demonstrate that Gibbard's 'adaptive syndrome theory' is a more sophisticated version of Hume's view.

[41] Baier, Ibid. 194.
[42] Ibid., pp.194-5.
[43] Ibid., p.195.
[44] Baier attributes this reading to 'internal realism.' See Baier (1991), p. 195.

5.2. Moral Emotion and the Notion of Warrant: Non-Cognitivism

5.2.1. Guilt as Moral Emotion

In order to show the relation between Gibbard's and Hume's view, in what follows, I shall provide a more detailed account of Hume's claim that x is good or virtuous if and only if X is such as to arouse the sentiment of approbation. In order to do this, I attempt to apply the notion of 'warrant' to emotions in the naturalistic sense.[45] There are many notions which show that morality engages the emotions, for example, guilt, remorse, compunction, and regret.[46] If we use the notion of 'warrant' in the naturalistic sense, for example, warrant of moral feeling, it neither concerns what to feel, all things considered, nor what it would be right to feel. Now, if our feelings can fall under moral evaluations, how can our naturalistic approach handle questions about the rightness of feeling guilt? This feeling can be said to be naturalistic in the following sense: a person is guilty if guilt over this thing, is warranted, on his part. According to this analysis, there are two elements: the feeling of guilt, and the warrant. To call a feeling warranted, as Gibbard suggests, is not to ascribe some special property to the feeling, for there is no one property that everyone who uses the term proficiently must have in mind. Rather, to call a feeling warranted is to express one's acceptance of norms that allow having that feeling. What then is the norm which makes sense of an agent's feeling guilt for doing something, and for other's resentment of him? In order to avoid judgmentalism which I have rejected in Chapter 2, let's take the view that this normative judgment is not

[45] Allan Gibbard (1990) applies the notion of 'warrant' broadly. He assesses the warrant of feelings as well as beliefs in the naturalistic sense. In using the term 'warrant,' I am indebted to Gibbard's argument.
[46] Gibbard tries to avoid making a distinction between guilt and remorse, and uses the terms interchangeably, whereas some philosophers distinguish between them. See Gibbard (1990), pp. 296-97.

supposed to be propositional. The reason why I reject the view that the normative judgment is propositional is because in expressing our emotions, such value judgments are not apt for truth or falsity. But in order to describe the notion of warrant in non-cognitivist terms, my approach to warrant does not allow recourse to the notion of normative judgments either. What then does 'norm' mean for this account? In order to answer this question, I suggest that we should first of all take notice of how morality actually involves the emotions; second, we should take a notion of wrongness or rightness to be to some degree culturally specific:

> What a person does is *morally wrong* if and only if it is rational for him to feel guilty for doing it, and for others to resent him.[47]

Thus in order to understand the rightness of actually feeling guilty, we can ask: Would it make sense to feel guilty, and for others to resent you, for not feeling guilty in a particular situation? In order to answer this question I take a sense of 'rational' suggested by Gibbard. Gibbard's use of the sense of rational is one of 'direct and flavorless endorsement,' which can be applied broadly to emotions, beliefs, desires, and actions. On this view, in each case, to call a thing rational is to think it 'makes sense' – which is to express one's acceptance of norms that permit it. Thus, according to him, emotions can be more or less justified, that is, they are manageable to norms of rationality. Although he uses 'rational' and 'make sense' interchangeably, his use of rationality does not mean advantageousness but means warrant. He writes, "the judgments I want us to consider are ones of warrant....I reserve 'rational' and 'make sense' for these uses."[48] On this view, moral judgments of right and wrong are connected with the rationality of what he identifies as the moral emotions guilt and

[47] I borrow this formulation from Gibbard (1990), p. 42.
[48] Gibbard (1990), p. 36.

impartial anger. According to Gibbard, to think some action is wrong is to think it rational for the agent to feel guilty, and others to resent him for so acting. Thus, Gibbard writes, "an act is shameful if it makes sense for the agent to feel ashamed for having done it, and for others to disdain him for having done it. An act is dangerous if it makes sense for the agent to feel afraid on account of having done it."[49]

This proposal is consistent with the view that emotions are the represent response-dependent properties. According to the response-dependent theory of emotional properties, something cannot be valued without being the object of a mental state, namely, being valued. "Something can be dangerous only *to* some creature or other, and whether or not something is dangerous depends on the creature in question."[50] However, this does not mean that being dangerous depends on being represented as dangerous. Radiation would be dangerous even if we didn't know that it is. The person might not be afraid of the radiation in a particular area due to a lack of the information needed to recognize the danger. Then, although he would not be warranted in being afraid, the situation is still dangerous. If this is the case, then dangerousness would not seem to be a response-dependent property, because many cases, such as radiation and high voltage, are dangerous for humans even if we don't represent them as such. However, many cases, such as amusement, happiness, envy, guilt, and shame, confirm that they are response-dependent emotional properties, because we are amused by funny things, angered by offensive things, disgusted by disgusting things, and elated by pleased things. The response-dependent theory implies that emotions are elicited by things as they relate *to us*. In this respect, some theorists have suggested that the formal object of

[49] Ibid., p. 51.
[50] See Prinz (2004 b), p. 63.

fear is frightening or fearful rather than dangerous.[51] But Prinz maintains that a response-dependent account is uncomfortably circular. For "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."[52] So it is viciously circular. What's worse, according to Prinz, is that if those properties represent anything that causes fear, anger, disgust, and elation in us then it looks as if those emotions can never occur in error.[53] However, in my view, Prinz misunderstands the response-dependence view of emotional properties, because according to the response-dependent theorist, all particular objects that cause fear in us are represented as frightening in our experiences of fear but those objects are actually frightening only if they *merit* fear. My reason for agreeing with the response-dependent theory of emotional properties is to show that the formal object is not merely a perceptual property, as I have discussed in Chapter 1, but the *norm* by which each emotion type defines its own standard of fittingness, as de Sousa points out.[54]

5.2.2. Adaptive Syndrome Theory

In defence of a response-dependent theory of emotional properties, in what follows, I, following Gibbard, do not focus on such concepts as shameful, funny, or dangerous but on ethical concepts such as right and wrong. According to Gibbard, the judgment that an act is wrong is sensitive to the agent's evidence in a way that almost allows the warrant of guilt to capture the rightness or wrongness of the act. If this is so, we have the use of wrongness or rightness in the subjective sense. In order to understand this, let

[51] See Goldie (2004); McDowell (1998).
[52] Prinz, Ibid., p. 60.
[53] Ibid., p. 61.
[54] See de Sousa (2002).

us consider an example: suppose that Kate's partner is an alcoholic. He abuses his family psychologically and physically. Kate knows he wants to stay with the family, but she is not sure whether she can be patient and can care for him. Her children are restless. If she can find reasons to avoid staying with him and not caring for him, she can make a decision to put him into the clinic. Once she is sure that this is the right thing to do, she can go ahead and do it, even if she knows she will feel guilt for doing so. But does Kate's guilt make sense? In this case, we may say that Kate's feeling guilt is unwarranted, for she has reasons to put him into the clinic.

The cognitivist might respond to this example as follows: the emotion - feeling guilt- is unwarranted when its belief component is unjustified. The difficulty the judgmentalist faces is to explain how it makes sense to feel guilt, for example, about an action which one does not believe to be wrong. As the above case show, someone can feel guilt for something that she has done and yet believe that she was right to do it. Now in order to avoid this difficulty, we should reject cognitivism or judgmentalism. Instead, let us take an account of the prescriptivity of value judgment on the basis of evolutionary theory. One speculation which addresses this idea well is Gibbard's 'adaptive syndrome theory'. On this view, we have various emotional states as a result of different, highly specific, and adaptive mechanisms. The way in which we can individuate emotions is in terms of disparate adaptive mechanisms underlying the various emotional states. If this picture is plausible, we can say in the case of guilt, guilt refers to a mechanism that consists in humans for feeling guilty was historically an adaptive response to anger; it makes for reconciliation, bringing about mutual aid between social groups rather than conflict. Whereas other animals have a comparatively restricted array of responses to hostility (fight, combat), humans needed to be able to moderate anger in order to reconcile the

conflict. Guilt evolved in the course of those responses. Given this, we can say that the standard of rightness or wrongness for feeling guilt is a response dependent ones, since in many circumstances we could feel guilt where others are reasonably be justified in being angry with us. According to Gibbard, guilt and anger are supposed to be mutual emotions, in that "guilt aims to placate anger, and it is governed by the same norms as govern anger."[55] Hence, to judge an act wrong is to think guilt and anger warranted. To renounce this reciprocity would make Gibbard's adaptive syndrome theory vulnerable.

But the trouble with this account is that it is not clear why the emotions, namely, guilt and anger, should be reciprocal. If we accept Gibbard's suggestion, our case shows at most that Kate's guilt can be warranted when her partner's anger is warranted. Let us take the following scenario: Kate's partner has sacrificed much for her, he wants to stay with the family, and Kate understands this. Then both guilt on her part and anger on his make sense. Given this, can we say that this is all about the normativity? Is to judge an act wrong to think guilt and anger are warranted? There is more to say: perhaps Kate's feeling guilt makes sense if she *believes* that she has deliberately desecrated the strongest wish of someone completely dependent on her. On the other hand, her partner's anger may make sense if he *believes* that his partner hurts him toward whom she has a special responsibility. Yet, as I have said before, our case tells us that Kate is convinced that she did the right thing, but feels guilt. In order to solve our difficulty, let's take the thesis that 'ought implies can.' The point of this thesis is that it is absurd to say that one ought to do something if he or she does not have the ability to perform the action. Many tragic cases such as ours show that there are cases in which the agent cannot help but to do wrong. If we cannot follow the norms which we are supposed to

[55] Gibbard (1990), p. 139

keep due to the limit of our ability, those norms cannot be established as the appropriate ones. Hence, many people suggest that the contents of norms should be restricted within the limit of human capability, and the norms should be replaced by descriptive ones. Thus, in order to offer a purely descriptive account, avoiding normative judgment, Gibbard suggests the norms in terms of the following question: "Are there situations in which, no matter what the agent does, it will make sense for him to feel guilty for having done it, and for others to be angry at him for having done it?"[56] In defence of the non-cognitivism of warrant for feeling guilt, avoiding the subjective sense of those emotions, we should say that the appropriateness of Kate's feeling guilt in the above case neither depends on Kate's blaming herself, nor on her partner's anger in being put into the clinic, but depends on whether anger on the part of *others* is warranted. We need, therefore, to discuss the status of these 'others'. Consider first Kate's partner's family and his closest friends. They may be angry, as it were 'on his behalf.' But if this is so, we can say that how angry we get rests strongly on what the situation is *to us*. If this is right, as D'Arms and Jacobson point out, "feeling anger depends on aspects that seem morally irrelevant",[57] because it lacks objectivity which morality requires. Hence, we can say that your feeling in the situation where there is moral import depends very strongly on where you're placed.

5.2.3. Anger and Disapprobation: Are they impartial?

Hume acknowledged this problem and regarded it as a difficulty with his sentimentalist account of moral judgment. According to Hume, our moral sentiments "whence-ever they are deriv'd, must vary according to the distance or contiguity of the objects."

[56] Gibbard, Ibid., p. 53
[57] D'Arms and Jacobson (1994), p. 756.

However, Hume claims that our esteem- our moral judgment- does not vary.[58] If this is so, in order for morality to be impartial, Hume must say what is appropriate for impartial others to feel about an acknowledged wrong. We may say that it makes no sense for you to be angry about killing people in the Iraq war. I have argued in Chapter 4 that evaluation of our emotional response depends on our interests and purposes. Thus one might say that an official three-minute silence for tsunami disaster victims is inappropriate, for the three-minute silence diminishes the significance of the British annual two-minute commemoration of those who fell in the world wars. On the other hand, some might argue that this way of expressing our emotion is appropriate, since the shock inflicted upon us by the tsunami is all the greater and we should grieve for people to whom terrible things happen and acknowledge that man is at the mercy of forces outside his control. Tony Blair, who has launched an unjustified war in Iraq and has to fight the coming election, might say that such emotional expression is appropriate, in order to appear to be embracing people's concern. But this is not the case when we apply it to moral appropriateness. Thus, in the case of moral emotion, it is strange to say that it makes no sense for you to be angry about the killing of people in the Iraq war. Hume would agree that how angry it makes sense for me to be depends on what the war has to do with me.[59] But this does not mean what happens far away is not morally important. By contrast, war is equally wrong no matter where it occurs. At this stage the non-cognitivist should qualify his argument. Moral norms are norms for what it makes sense to feel from what Gibbard calls an *impartial* but *fully engaged* standpoint. Hence, while it is understandable why people do not get angry at distant wrongdoing, they

[58] Hume, T. 581.

[59] In Chapter 4, I have argued that in the case of emotion, it is unintelligible if I say that I am proud of someone's ancestors who are completely unrelated to me. But this is not the case when we apply it to moral appropriateness.

would be angry were they fully engaged. This view is similar to Hume's. Hume suggests that "in order to present those continual *contradictions,* and arrive at a more *stable* judgment of things, we fix on some steady and general point of view; and always, in our thoughts, place ourselves in them, whatever may be our present situation."[60]

However, to take the moral point of view, namely, that of a fully impartial and fully engaged spectator seems to be impossible for mere humans due to a number of distorting factors, such as prejudice and partiality to oneself. In the previous section, I have argued that to answer what the morally correct course is, is more difficult to answer than what the correct colour of things is. To appreciate this difficulty let us consider the example provided by McGinn.

> Two people could in principle be capable of making equally fine-grained moral distinctions, and yet one of them be right and the other wrong with respect to a case which they can both equally discriminate from other cases. Moral blindness is not (or is not principally) a matter of perceiving two situations as indistinguishable in value which others perceive as having different values; it consists rather in assigning the *wrong* value to a situation which the judger can morally distinguish from other situations as well as the next man.

Another example of McGinn's shows us this more clearly. He says:

> Consider a moral analogue to the case of human and Martian colour perception: the moral ascriptions of the two communities are equally complex and fine grained, but different values are assigned to the same state of affairs. In the colour case we should say that no error is involved on either side; and this is predicted by the dispositional theory. But in the moral case the analogous prediction is borne out: we want to say that one side (at least) must be morally wrong.[61]

It seems that this is a way that we might be morally mistaken about the matter. Yet

[60] Hume, T. 581-2.
[61] McGinn (1982), p. 152

Hume says that correction of our mistake is possible. When he talks about corrections, Hume argues that it is possible in terms of others' *disapproval.* If this is so, we can say that to judge an act is wrong is to judge that it calls for disapproval from others. But the trouble, as we have seen so far, is whether others' disapproval has objectivity or not. In this respect, we can also raise a question for Gibbard: does his 'others' anger have objectivity? Do we really need to appeal to other's anger to judge that an act is wrong? Confronted with this objection, Gibbard appeals to 'outrage' or 'indignation,' an attitude that seems to suggest fully fledged moral judgment and one that has more cognitive character than does anger. But if he adopts this kind of notion, he appeals to a cognitive element which he has attempted to avoid. In this respect, some people say that although Gibbard's 'others' anger' may not be impartial, Hume's 'approval' or 'disapproval' are close to impartial, since as D'Arms and Jacobson point out, "it's more tenable to say that to think an act wrong is to think it would make sense to *disapprove* of it from an impartial, informed position."[62] If we agree with this claim, then we are back to the original difficulty of sentimentalism: is disapprobation a feeling or a judgment?' In replying to this question, neo-sentimentalists suggest that the disapprobation has an aspect of rationality, since in order to correct our sentiments "there is some position from which our sentiments are more likely to get things right."[63]

Now I want to raise another objection to Hume's and Gibbard's non-cognitivism in order to support a qualified sentimentalism. How then, on Hume's view, could a fully informed impartial moral being be possible? If we identify moral judgments or opinions with beliefs about the responses of an impartial or otherwise *ideal observer*, it is clear

[62] D'Arms and Jacobson (1994), p. 760
[63] See D'Arms and Jacobson (1994), p. 758; A similar view appears in 'A Sensible Subjectivism', David Wiggins (1987)

that it cannot capture Hume's thinking. Such beliefs must be amongst the targets of his arguments concerning moral reasons for action, since on Hume's non-cognitivist point of view, to judge that an action is right is not to have a belief that an ideal spectator would approve of it.[64] Now a question arises for Gibbard: is full impartial engagement when we feel anger from this standpoint warranted in all cases or only those in which we judge an act wrong?[65] In reply this question, Gibbard describes full engagement as "vivid awareness of everything generic that would affect one's feelings toward a situation." "It requires, moreover, unrestricted contemplative leisure, for in responding to an emergency I might be vividly aware of all that was involved but have no time for feelings."[66] Although I agree with Gibbard in that full engagement may endorse the feelings they give rise to, but I doubt that they give rise to the right feeling. Faced with this objection, Gibbard replies that "moral norms tell how to feel only given a special standpoint." Anger is the standpoint from which one is in full and impartial engagement; to feel guilt, on the other hand, is to be situated in a partial standpoint, that

---

[64] In this respect, J. Bricke (1996) remarks that "Hume has the conception of an impartial desirer, as also the complementary conception of an individual whose consequent affective responses are impartial." Bricke believes that these ideas are essential constituents of Hume's expanded conativist theory. Bricke goes on to say:

> [Hume] can have no objection to the suggestion that, in adopting the vantage-point of conative and affective impartiality, an individual may, or may typically, have an accompanying belief that others who share that vantage-point will thus far want, and be satisfied by, what she herself wants and finds satisfying. He can surely countenance the suggestion that, in having that accompanying belief, the individual will think of herself and such others symmetrically, as each approximating to an ideal, and ideally reflective and informed, impartiality in desire and affection. He must insist, however, that such accompanying beliefs not be confounded with the impartial desires and affections they accompany and presuppose. (p.157)

[65] This question is raised by D'Arms and Jacobson (1994), p. 759.
[66] Gibbard, p. 127

is, influenced by aspects of your own life. "Moral norms coordinate the anger and the guilt it makes sense to feel from these special standpoints, and in turn these standpoints will have something to do with the feelings it makes sense to have in the flux of life."[67] Gibbard is recommending that we should try to vividly consider the circumstances of everybody's lives. However, can we sympathize with everyone equally? Since Hume says: "We sympathize more with persons contiguous to us, than with persons remote from us: with our acquaintance, than with strangers; with our countrymen, than with foreigners."[68] In this respect, a universalist may say that sympathy can be extended beyond personal contacts.[69] For example, John Rawls thinks of Hume's sympathy based moral theory as a universalist view when he addresses Hume's theory of justice.[70] If so, it can be said that all normal moral beings who take up the common standpoint will make similar moral judgments, just as in the case of agreement about colour perceptions. However it is difficult to hold the moral point of view, and therefore to arrive at the ultimate goal of complete agreement, since as much research shows, sympathy is biased[71] and in our actual state, there seems to be limits to our sympathetic power.[72]

---

[67] Gibbard, Ibid.

[68] Hume, T. 581

[69] Universalism requires the moral agent to consider everyone's pleasure or pain equally and impartially. By contrast, particularism allows the agent to display some degree of partiality toward individuals with whom one has a personal affinity such as family members, friends, students, or comrades. In this respect, some people argue that if we consider the prospects of an empathy-based view of morality, it might seem natural for it to tilt toward particularism. See Lawrence Blum (1980) and (1977), pp. 306-37.

[70] Rawls(1977), pp. 185-86

[71] Martin Hofmann shows us that there are research evidences that sympathy tends to be biased. According to him, first, observers are more empathic to victims who are familiar and similar to themselves than to victims who are different. Secondly, people are more apt to be empathically aroused by someone's distress in the immediate situation than by distress they know is being experienced by someone elsewhere. See Martin Hoffman, "The Contribution of Empathy to Justice and Moral Judgment," in *Empathy and Its Development,* ed. Nancy Eisenberg and Janet Strayer (Cambridge University Press), pp. 47-80.

[72] Hume remarks that institutions of justice would not be necessary at all if the human mind were "so replete with friendship and generosity, that every man has the utmost tenderness for every man, and feels

I agree with Gibbard's analysis that guilt and anger provide the foundation for our culture's morality, since as Gibbard notes, "if the very terms in which different cultures express emotions resist translation, then it is hard to see why our own ways of thinking should be get things right."[73] However, I would raise the following question. If we identify moral considerations with anger in particular, it overly emphasises one seemingly arbitrary emotion. In other words, in my view, anger is the only emotion suitable for explaining Gibbard's example, namely, judging an act to be wrong. If so, Gibbard's expressivism refers only to a notion of wrongness which is at any rate culturally specific. On this view, moral reasons are connected with the emotions of guilt and anger. We have moral reasons of what to do indirectly when we feel these emotions.

Finally I raise a question to both Hume and Gibbard's non-cognitivism. If Hume and Gibbard appeal to a vacuous notion of an ideal observer, they could be said to admit the following point: to make an evaluative judgment is not to *have* but to *endorse* a sentiment, since to appeal to an ideal observer is to accept the rationality of the sentiments. If this is right, the Humean should show how to justify a sentiment when one makes an evaluative judgment. In order to do this, in what follows, I shall show that contemporary philosophers who want to combine cognitivism and sentimentalism put forward the view that values are grounded in the sentiments, and at the same time try to make sense of the rational aspects of evaluation. We shall see this picture in the dispositional analysis such as that presented by McDowell and Wiggins. If we adopt this view we can say that moral judgments are grounded in our human dispositions, and

---

no more concern for his own interest than for that of his fellows." (E. 146) In our actual state, there seem to be limits to our benevolence, and our empathic powers. These limits create the need for legal and political institutions that will be sufficiently constraining to shape the conduct of creatures like us.

[73] Gibbard, Ibid., 128

yet maintain their objectivity.[74] I shall call the view qualified sentimentalism.

[74] See Wiggins (1987).

5.3. Qualified Sentimentalism

5.3.1. Response-Dependent Properties of Emotions and Revisionists

In section 1, I have argued that the dispositional analysts attempt to develop Hume's sentimentalism by combining cognitivism with sentimentalism. In doing so, I have argued that they use the analogy between colour and value. They compare the perception of value to the account of an object being red. For example, something is red iff it is such as to look red to a normal observer under standard conditions. Accordingly, something, X has the value Q (shameful, dangerous, funny, enviable, etc.) iff normal observers under standard condition would feel F (shame, fear, amusement, envy, etc.). But they face difficulties in specifying what 'normal' observer under 'standard conditions' means. Neo-sentimentalists, most notably McDowell and Blackburn, agree that those conditions are our own dispositions, corrected for occasional anomalies and the dispositions of a majority. "My attitude," Blackburn writes, "ought to be formed from qualities I admire – the proper use of knowledge, real capacity of sympathy, and so on. If they are not, and if the use of those capacities and the avoidance of the inferior determinants of opinion would lead me to change, then the resulting attitudes would not only be different, but better."[75] Then this proposal can be called revisionist and one might object that revisionists cannot provide the normative aspect of evaluative judgments. The revisionists try to solve the normative question by appeal to those whose response is more normal. But they do not provide an answer for what 'normal' means. This is a compelling objection to the theory.

[75] Blackburn (1993), p. 79.

Confronted with this objection McDowell suggests a normative locution, that is, 'merit.' He tries to define sentimentalism in terms of this normative notion as follows: "a virtue (say) is conceived to be not merely such as to elicit the appropriate 'attitude' (as a colour is merely such as to cause the appropriate experiences), but rather as to *merit* it."[76] But critics argue that the term 'merit' does not much help to distinguish the proper sense of appropriateness of moral emotions.[77] Suppose that someone is very cowardly due to some severe childhood trauma. Can we say in this case that he merits shame, even though his cowardice is no fault of his own? Dispositional analysts might reply that the case doesn't make his cowardice any less shameful, since his cowardice is his own disposition and it can be corrected by 'the proper use of knowledge' or pathology, and so on.

According to revisionists, the normativity of evaluative judgment can be attained in terms of our refinement of our own dispositions. Furthermore, they argue the refinement of our dispositions leads to improvement and this improvement constitutes an increased sensitivity to φ. If we agree with this idea, then we will support a response-dependent theory of evaluative properties. This theory can be formulated as follows: for an evaluative property Φ- what φ is relative to - to be instantiated by X- where X is a situation or circumstance - is for an associated response A- where A is our response - to be appropriately held toward X.[78] But the trouble with this view, as D'Arms and Jacobson point out, is that I may revise my defective attitude, for example, in favour of not laughing at an offensive joke, from an increase of sympathy. It might be said that this is moral progress, but it constitutes a decrease in sensitivity to the funny. Thus

[76] McDowell (1985), p.207, his emphasis.
[77] For example, see D'Arms and Jacobson (2000 b), pp 734-736.
[78] See, for example, McDowell and Blackburn

although it can be said that our refined sense of humor is morally better, it cannot be said that the joke is not funny.[79]

The opponents of the response-dependent theory complain that the revisionists do not present ways to differentiate moral and prudential reasons for feeling a sentiment: they do not provide a way to distinguish, for example, the moral reason that the joke was offensive, from the prudential reasons bearing on whether X is φ – where X is a state of affairs and where Φ is a moral/emotional property. Furthermore, opponents contend that morally admirable qualities, for example, kindness, charity, and love, are not granted by attitudes that are more sensitive to any given property.[80] Opponents also claim that even if someone changes his attitude by increasing his sympathy, if this refinement is grounded on the wrong kind of reason, it is not an improvement to 'earn talk of truth.'[81] Thus, opponents hold that revisionists fail to distinguish the relevant notion of appropriateness of evaluative judgments. I have argued in the previous Chapter (Chapter 4), our emotional responses cannot be assessed by a logical or conceptual 'norm'. What then are the norms applying to emotions? There might be, one might argue, several distinct norms of emotion. There may be a prudential norm; 'whether it is good for you to feel F'; or a moral norm, 'whether it is right to feel F'; or 'whether F is what to feel, all things considered.'[82] But our emotional responses cannot be assessed by a logical or conceptual 'norm,' since they are evaluated for their appropriateness.[83] If we say that

[79] As argued by D'Arms and Jacobson (2000b), p. 734

[80] D'Arms and Jacobson, Ibid., 734

[81] Response dependent theorists argue that given the subjective basis of values in human responses, talk of the truth of evaluative judgment must be earned, rather than simply assumed. See McDowell and Blackburn.

[82] D'Arms and Jacobson criticize those who try to apply these normative idioms to the question of whether F is fitting. See D'Arms and Jacobson (2000a), pp. 69-86.

[83] We can trace this idea from Hutcheson and Hume. The idea that our emotional responses can be assessed for their appropriateness has encouraged the sentimentalist tradition in ethics.

emotion's appropriateness is a matter of fitting, this is a normative question: to judge an emotion as fitting is to endorse its evaluation as correct. This is a higher-order attitude toward the emotion, which we re-establish by using some regulative terms such as *fearsome, shameful,* and *funny.* However, as I have argued in Chapter 4, the appropriateness of emotion is not the same as our assessment of the accuracy of their appraisals. Considerations of the fittingness of emotion, I shall argue, can only apply to morality, since it can ground the specific force of 'should.'

5.3.2. Appropriateness of Emotions: 'Getting Things Right'

David Wiggins provides a similar suggestion to mine. According to him, evaluative properties and their associated responses are 'made for' each other.[84] On this picture, we regard things that actually disgust us as 'disgusting', things we are ashamed of as 'shameful,' and so on. Now the reason why these properties and their responses are associated is because, these responses matter to us. Wiggins calls this view 'subjectivism' in that the properties in question are explained by reference to the reaction of human subjects.[85] But one might say your being angry about a theft happening halfway around the world isn't appropriate, since the event does not matter to you. But this does not mean what happens far way is of no moral import. By contrast, Wiggins might say that thefts are equally wrong no matter where they occur. In this respect, Wiggins maintains that moral properties have salient features of things which he calls the 'marks' of such properties. Hence he argues that we have to look for the features. Furthermore, according to Wiggins, our responses are not 'mere' responses. They are responses that are correct when and only when they are occasioned by what

[84] Wiggins (1987), p. 198.
[85] Ibid., p. 195.

has the corresponding property ϕ and are occasioned by it because it *is* ϕ.[86] In other words, we can say that the property retains an essential connection to the response. In this respect, he claims that his subjectivism is a subjectivism of subjects and properties *mutually* adjusted.[87] Thus, he argues that his subjectivism does not give up the idea of achieving any simple or single statement of the standard of correction,[88] since we can revise what is shameful, for example, partly by appeal to the marks of the property and partly by appeal to the nature of our 'shared way of talking, acting, and reacting.'[89] According to this view, the property ϕ is only really ϕ if it is such as to evoke and make appropriate the response A in any those who are sensitive to ϕ-ness.[90] The property of being funny, for instance, is specified by reference to an individual, being amused. Thus as far as evaluative properties are concerned we may call this revisionist view the 'response-dependent thesis.'

But the trouble with this view is how to determine what a given sentiment, amusement for example, is appropriate to. In answering this question, Wiggins remains unclear. His explicit answer to this question is simply that we don't need complete agreement about the marks of the property nor do we need to be concerned with the same attitude directed towards it.[91] He suggests that "by hypothesis, the linked properties and responses... are arrived at by a historical process."[92] In other words, what a sentiment is made for depends upon the history of the evaluative practices of the people whose sentiment it is. Different cultures find different things funny, disgusting, shameful, and

[86] Ibid., p. 204-5.
[87] Ibid., p. 199.
[88] Ibid.
[89] Ibid., p. 205.
[90] Ibid.
[91] Ibid., p. 212.
[92] Ibid., p. 195.

so on. Their differing histories of reflection and refinement, driven by social pressures imposed by the feelings and judgments of the community, inevitably establish disparate standards of what is, for example, shameful. Now one might object to this view on the ground that although the standards of our community exert a deep influence upon us, one can sensibly reject them. Confronted with this objection Wiggins claims that in order for an evaluative concept to play a central role in normative discourse, its marks must be 'essentially contestable' – whether, for instance, social class, wealth, and sexual preference are appropriate sources of shame and disdain.[93] Hence in order for Wiggins's neo-sentimentalism to be considered a powerful alternative to competing views, it must present a more specific argument for the idea of 'essentially contestable.' However, Wiggins does not fully explicate the idea of essential contestability. In the previous sections, I have shown the attempt to meet the 'essential contestability' objection in terms of Gibbard's 'others' anger' and Hume's disapprobation. I have also shown that these accounts have difficulty, since as Wiggins claims, "a sentiment of approbation cannot be identified except by its association with the thought or feeling that x is good (or right or beautiful) and with the various considerations in which that thought can be grounded, given some particular item and context, *in situ*."[94] In this respect, Wiggins suggests that we should abandon Hume's aspiration to secure the standard of correctness in valuation from outside the domain of values. If we so do, the criterion for a good judge is that he is apt to get things right.[95] As I have suggested before, if we take the view that the disapprobation has an aspect of rationality, there is the possibility that our sentiments are more likely to get things right. Hence, I suggest that if we accept this

[93] Ibid., pp. 198, 212.
[94] Ibid., p. 187.
[95] Ibid., p. 194.

idea, we can apply the consideration we came to, regarding the fittingness of emotions, to morality.

5.4. Conclusion

In this Chapter, I have shown that the emotions play an important role in our lives as moral agents. I have argued that our emotional reactions and responses play an important role for our moral evaluations. Furthermore, I have argued that emotions can be changed. Yet I have claimed that emotions are not modified by the power of our own rational judgment but changed only through the influence of approving or disapproving responses of those who are affected by our emotion. We pursue other's approval, since their company and judgments are important for us, and we can change our emotions and improve them only in the course of attuning our emotional responses and behaviour reactions to theirs. Our emotional reactions and behaviour reactions are judged by others, and at the same time we evaluate these other people's judgments. I have demonstrated that this view lends support to a 'response dependent theory of emotional properties', according to which all particular objects that cause an emotion in us, for example fear, are represented as frightening only if they *merit* fear. Then, also in this Chapter, I have shown the difficulties that this theory encounters. The objections which have been made was that the response dependent theory of emotional properties cannot explain the normative aspect of evaluative judgments, since the theory begs the question concerning whose emotional response is normal or appropriate. As many argue, the normal isn't normative. In order to meet this objection I have argued that our emotional responses cannot be assessed by a logical or conceptual norm, since they are evaluated in terms of their appropriateness, something which depends on a person's perspective or their historical-cultural location.

I have claimed in Chapter 3 that the social aspect of emotion provides an alternative to

the affect and cognition accounts. I have also attempted in this Chapter to demonstrate it in terms of the response-dependent theory of emotional properties. This Chapter has also presented my alternative to meet the difficulty that the embodied appraisal theory encounters. As I have mentioned in Chapter 3, my alternative takes moral phenomena to be natural and socio-cultural phenomena. This Chapter have offered a naturalistic picture of moral evaluation by associating it with certain emotions, namely, guilt and anger. In the next Chapter (Chapter 6), I shall explicate the social aspect of emotion in terms of the problem of personal identity.

Chapter 6: Emotion, Memory, and Personal Identity

Preliminary Remarks

What are the conditions under which a girl *A* at time *t* and 10 years later, a woman *B*, are the very same person? To answer this question, I begin with the Hume's theory of personal identity. Although many philosophers in the mental connectedness tradition begin with assuming Locke's view in explaining personal identity, Hume's theory of passions and the self has been neglected by them. Hence, in this Chapter I shall begin by Hume's view of emotion and personal identity.

Hume's official view, according to which the self is a bundle of perceptions, cannot explain the unity of a person's life. Hume addresses personal identity in Book One of the *Treatise* almost exclusively in terms of the relationship of the present self to its past, primarily in terms of causation and memory so that he faces difficulties explaining the unity of a person's life. But it should be noted that Hume also holds that the self is extended through time and that a person's life is a projection over time with regard to emotion. In Book Two of the *Treatise* Hume attempts to explain why past perceptions, thoughts, and actions affect my present *feelings*, and therefore why these feelings influence my future. He introduces a further causal relation into the account of the self—the relation between intention and action. He explains why I act from concern for a future that will bear to me, now, the relation I bear to my past. Hence he argues that in thinking of *myself* in the future I am thinking of actions that follow from my motives, intentions and character.

Although his view in Book Two fares better than Book One in the sense that he tries to unify the diachronic expansion of a person's life in terms of its influence on emotion, I

argue that this view has defects as well. It seems to me that in order for Hume's view of personal identity in terms of emotion and tense to be plausible, Hume must take notice of the notion of 'the experiential content of memory' which is embedded in certain *feelings*. That experiential memory plays a central role in explaining personal identity has been neglected by many philosophers in the mental connectedness tradition. In order to understand memory as experiential, we need to understand the affective tendency attached to some memory. I argue that memory affects not only my past thought but also my past emotions and those emotions deriving from the past stay on to affect my whole being and my future. Thus for example the shame or regret that I experience when I remember my past wrongdoing represent the whole of me up to now as well as at that time.

If we adopt the notion of experiential memory to identify mental connectedness with personal identity, we bring the person under the influence of his past, such that the past is reconstructed and then influences the future. Hence, a life- the diachronic expansion of a person- is unified. If this is right, I argue, the problem of personal identity can be resolved in terms of personal history understood from a person's own perspective (e.g. emotion).

6.1. An Episodic Outlook on the Self

The first position I want to discuss in this section is the episodic self. An episodic view, according to which the self is nothing but fleeting inner mental states, is an anti-narrative view. An episodic view denies the kind of self of long-term concern or self-created and creating narrative self. On the other hand, a diachronic view, according to which a person's life is something existing in the past and future, can accommodate the narrative view. In this Chapter, I argue that although the episodic outlook on the self in Book One of the *Treatise* has difficulty in explaining personal identity, the diachronic outlook on the self in Book Two leaves open a solution by paving the way to the narrative outlook.

6.1.1. Identity through Time

In what follows I take up Hume's two main and distinctive theses about the self: One is his negative view of the self, namely, his opposition to the Cartesian view, and the other is his positive conception. For Hume the polemical target of personal identity is not the existence of a conception of the self, but rather the theory of the identity and simplicity of a substantial soul. Hume is writing against

> Some philosophers, who ...are certain, beyond the evidence of a demonstration, both of its perfect identity and simplicity.[1]

Here Hume unquestionably has the Cartesians in mind. What Hume is intending by the above remark is that the perfectly identical and simple self is a philosopher's invention, without basis in everyday self-awareness. He criticizes the Cartesians by virtue of his

[1] Hume, T. 251.

first principle of empiricism:

> From what impression cou'd this idea be deriv'd?...it must be some one impression, that gives rise to every real idea.[2]

The upshot of this investigation shows the indefensibility of the theory of the self as an identical simple substance. There is no idea of ourselves as something distinct from the particular impressions and ideas. So the question concerning the substance of the soul is unintelligible. Hence Hume writes:

> When I enter most intimately into what I call *myself,* I always stumble on some particular perception or other ... I never catch *myself* at any time without a perception, and can never observe anything but the perception... they are nothing but a bundle or collection of different perceptions, which succeed each other with inconceivable rapidity, and are in a perpetual flux and movement[3] .
>
> The mind is a kind of theatre, where several perceptions successively make their appearance...The comparison of the theatre must not mislead us. They are the successive perceptions only, that constitute the mind; nor have we the most distant notion of the place, where these scenes are represented, or of the materials, of which it is compos'd[4].

Here by emphasizing serial identity Hume shows that he is concerned with the conditions for identity of *xs* through time, not with the criteria for regarding a perception as a member of this set rather than that.[5] The reason why Hume is concerned with identity through time is that he aims at removing conceptual confusion stemming from unreflective common sense that every one retains their identity throughout their lifetime. Hume does so by tracing the idea of the self to its source in experience. He

---

[2] Hume, T. 251.
[3] T. 252.
[4] T. 253.
[5] I take this to mean that by 'identity' Hume does not mean the identity of any entity with itself. Hume's identity is rather identity through time.

finds the source of confusion in the psychological reasons: we have "so great a propension to ascribe an identity to these successive perceptions, and to suppose ourselves possest of an invariable and uninterrupted existence thro' the whole course of our lives"[6]. Now Hume tries to explain the various ways in which "our propension to confound identity with relation" makes us "to attribute identity, in an improper sense, to variable or interrupted objects"[7]. This tendency is so great, that 'our mistake' is usually also accompanied by the fiction of

> something unknown and mysterious, connecting the parts, beside their relation; and this I take to be the case with regard to the identity we ascribe to plants and vegetables.[8]

What Hume in the above passage is trying to show is that it is possible to establish an analogy between the identity we ascribe to ourselves and that we ascribe to plants and animals. We can find the most apparent situation of these in the relation of cause and effect: the '*sympathy* of parts to their *common end'*, which appears in plants and animals. For Hume the terminology of a 'person' first appears in this connection, and in a comparison with a plant. He writes:

> tho' every one must allow, that in a very few years both vegetables and animals endure a *total* changes, yet we still attribute identity to them, while their form, size, and substance are entirely alter'd. An oak that grows from a small plant to a large tree, is still the same oak; tho' there be not one particle of matter, or figure of its parts the same. An infant becomes a man, and is sometimes fat, sometimes lean, without any change in his identity.[9]

In the case of a living being, that is, animals and vegetables, "not only the several parts

[6] T. 253.
[7] T. 254-5.
[8] Ibid.
[9] T. 257.

have a reference to some general purpose"[10] but among the parts themselves there is a "mutual dependence on, and connexion with each other," amounting to "*sympathy* of parts to *their* common end"[11]. What then does Hume mean by analogy between the identity we attribute to plants and animals and identity we attribute to ourselves? Hume's answer is:

> The identity, which we ascribe to the mind of man, is only a fictitious one, and of a like kind with that which we ascribe to vegetable and animal bodies.[12]

For their changeability of parts is a 'fictitious one.' Given that what Hume is seeking by 'identity' is the identity through time, what he in effect claims is that the identity of individuals through time is not perfect. What we can call a thing at one time, and 'the very same thing' at a different time is not perfectly identical. They are wholly distinct temporal parts. 'Identity through time' is the relation which welds these temporal parts into a whole. And for Hume the idea of a self or a person is essentially a temporal one. The problem is then by what connection the idea of a self (a person) is bound into a whole. Hence, in what follows, we shall see Hume's answer for this question in his Book One of the *Treatise.*

6.1.2. The Unity of the Mind

The question now arises: given that we are just bundles of perceptions, why do we think that these bundles constitute individual minds? And what unites perceptions into what is or is thought to be a single thing? In order to answer this question Hume provides another analogy.

---

[10] T. 257.
[11] Ibid.
[12] T. 259.

> ...I cannot compare the soul more properly to any thing than to a republic or commonwealth, in which the several members are united by the reciprocal ties of government and subordination, and give rise to other persons, who propagate the same republic in the incessant changes of its parts. And as the same individual republic may not only change its members but also its laws and constitutions; in like manner the same person may vary his character and disposition, as well as his impressions and ideas, without losing his identity. Whatever changes he endures, his several parts are still connected by the relation of causation.[13]

A republic survives because there is some very general pattern of organization - some 'reciprocal ties of government and subordination' - which enables the same republic to persist despite the constant change of the citizens which make it up. And it is possible for its particular laws and constitution to be changed without losing its identity. "In like manner," according to Hume, "the same person may vary his character and disposition, as well as his impressions and ideas, without losing his identity"[14]. If this is the case, as Nathan Brett (1972) points it out, "there are some basic conditions which are constitutive of the identity of such complex objects as states."[15] What are then these basic conditions? According to Hume, they are the principles of association, that is, "uniting principles of the ideal world"[16]. Hence, it can be said that it is these principles that give the organization an identity. Without them such complex objects as states or perceptions would be a mere heap, bundle, or collection of perceptions. If this is the case, as Brett remarks, "the associations are precisely what create a system out of our varied and interrupted succession of perceptions."(p. 121) Given this, we can say that the Cartesian view, according to which the problem of personal identity is the problem of the simplicity of a substantial soul, is replaced by the view that the identity (unity) of

[13] T. 261.
[14] Ibid.
[15] Brett (1972), pp. 120-1.
[16] T. 260.

the mind consists in the principles of association among ideas.

Although there is a similarity between the self and a republic or commonwealth, there is also a disanalogy. According to Hume, the way in which the perceptions making up a self constitute a unity is not exactly the same as the way of those relations binding together the members, etc., of a commonwealth. In this respect, J. Biro (1976) claims that "the relations uniting the perceptions which constitute the self are radically different from those involved in accounting for the unity of a commonwealth."[17] This is because the perceptions making up a self include memory perceptions which play 'both revelatory and productive' role, while in the case of the constituents of a commonwealth these roles are neither required nor possible.[18]

If this interpretation is right, we can say that the role of memory perceptions is the glue that holds the self together. Hence memory, Hume says, "acquaints us with the continuance and extent of [the] succession of perceptions"[19]: it is through memory that we are made aware of how perceptions succeed each other in time. "Had we no memory, we never shou'd have any notion of causation, nor consequently of that chain of causes and effects, which constitute our self or person"[20]. Thus, it is the presence of memories among my perceptions that is the ultimate source of the idea that I am a temporally extended being. But we do not remember all our past actions or experiences. Yet we think that the present self is the same person as the self of that time. Hence, there must be something else which enable us to think of our identity as extending beyond our memory. What, then, in addition to memory, would guarantee one's identity over time?

[17] Biro (1976), p. 26.
[18] Ibid.
[19] Hume, T. 261
[20] T. 261-2

In order to do this Hume appeals to *causality and resemblance*.

6.1.3. Causality and Resemblance

Hume then says that we, in the absence of memory, can guarantee our identity over time as follows:

> Having once acquir'd this notion of causation from the memory, we can extend the same chain of causes, and consequently the identity of our persons beyond our memory, and comprehend times, and circumstances, and the actions, which we have entirely forgot, but suppose in general to have existed[21].

However, why is causation essential in the production of the identity of a temporally extended self that learns of its own existence by way of that relation? The answer is that causation is the only relation "that can be trac'd beyond our senses, and informs us of existences and objects, which we do not [now] see or feel."[22] It is the only thing that can yield a "conclusion beyond the [present] impressions of our senses."[23] Hence Hume can say that

> The true idea of the human mind, is to consider it as a system of different perceptions or different existences, which are link'd together by the relation of cause and effect, and mutually produce, destroy, influence, and modify each other.[24]

Now if this is the true story of the self, why does then Hume confess his labyrinth difficulty in the Appendix? I shall offer my interpretation of Hume's dissatisfaction with the problem of personal identity later in this Chapter. Before providing the answer, in what follow I shall consider a provisional answer to Hume's difficulty which we can

[21] Hume, T. 262.
[22] T. 74.
[23] Ibid.
[24] T. 261.

solve in his theory of passion in Book Two.

6.1.4. The Problem of Individuation

Hume's claim has been accepted by all proponents of the psychological continuity criteria of personal identity. They agree that memory alone cannot constitute personal identity, but must make up only one of various causal links between the earlier and later psychological states of a person which constitute his identity. Before I suggest what Hume's real difficulty is, I present a source of difficulty in relation to the episodic self.

One of the difficulties that Hume faces is that the relation of causation and resemblance are insufficient to explain personal identity. For any two qualitatively identical perceptions in the minds of A and B respectively can differ in their causal relations only by differing in their temporal priority or spatial contiguity to other perceptions. But according to Hume there are spatially non-locatable perceptions (for example, sounds, tastes, smell and even passions). He says:

> An object may exist and yet be nowhere, and I assert, this is not only possible, but that the greatest part of beings do and must exist after this manner.... This is evidently the case with all our perceptions, except those of sight and feeling. A moral reflection cannot be placed on the right or on the left hand of a passion, nor can a smell or sound be either of a circular or a square figure. These objects and perceptions, so far from requiring any particular place, are absolutely incompatible with it, and even the imagination cannot attribute it to them.[25]

Now if those qualitatively identical, but spatially non-locatable perceptions, for example, smell could occur in distinct minds *simultaneously,* how, on Hume's theory, are we to assign these two perceptions to different minds? Confronted with this objection, Hume

[25] T. 235-6

would answer that: it is not by resemblance, for we have assumed that they are qualitatively identical. If resemblance cannot be a candidate, the only alternative would be causality. But this doesn't seem to work either, if distinct causes can be separated only on the basis of spatial relations, which these perceptions lack, or temporal relations, which these two perceptions share. On this view of causation, Hume's appeal to resemblance and causation in explaining the unity of mind gets into trouble.[26] In brief: as Don Garrett argues, if A and B both simultaneously *feel*, Hume would not be able to explain the fact that one of these feelings is A's, the other B's.[27] I think that this is a trenchant criticism. And Hume may have in mind this kind of difficulty when he pronounces himself arrested in a 'labyrinth.'[28] But I am not convinced they are the whole of his problem, for Hume has resources available to deal with the apparent counterexamples.

We can find the resources in the bodily criterion in Book Two. In explaining the distinction between the self in Book One and Book Two of the *Treatise*, the body is an important clue. Marina F. Spada emphasizes its importance. According to her, "the relation between the self and the body marks a difference between Book One and Book Two of the *Treatise*."[29] The vulgar tend to think that we are a composite of mind and body. But in Book One, Hume's point is this:

> 'Tis not our body we perceive, when we regard our limbs and members, but

[26] We can call this kind of difficulty so-called 'the problem of individuation.' Many people including Hume scholars point out that this is Hume's real difficulty. E. g., See Stroud (1977); Garrett (1981); Noonan(1989).
[27] Don Garrett, (1981) pp. 350-354.
[28] Many different interpretations have been given to Hume's dissatisfaction with his view on personal identity in the Appendix of the *Treatise*. It seems to me that we cannot give an exact reason for his dissatisfaction with this matter, since Hume never says explicitly what is bothering him. However, we could suggest what Hume ought to say about his difficulty. I will propose an alternative way of this matter later in this Chapter.
[29] See Spada (1998), pp. 196.

> certain impressions, which enter by the senses; so that the ascribing a real and corporeal existence to these impressions, or to their objects, is an act of the mind as difficult to explain, as that [of the external existence of our perceptions].[30]

In Book Two Hume states the point in a different way:

> Bodily pains and pleasures are the source of many passions, both when felt and consider'd by the mind; but arise originally in the soul, or in the body, whichever you please to call it, without any preceding thought or perception. A fit of gout produces a long train of passions, as grief, hope, fear; but is not deriv'd immediately from any affection or idea. [31]

In the above passage[32] which appears in Book Two, Hume seems to have in mind ordinary people's view that the self is composed of body and mind. He writes: "pride and humility have the qualities of our mind and body, that is self, for their natural and more immediate causes"[33]. Hence, as John Bricke remarks, "[a] theory of the self that is adequate to the understanding of the emotions must... treat the self as a compound of mind and body."[34] In Book One, the mind has itself as its only objects, and postpones a sense of self that makes reference to the body and to other people. But once Hume recognizes the metaphysical problem with the self in Book One, as shown in the Appendix, he finds in Book One's *conclusion* that this agonizing difficulty is a consequence of philosophers' emphasis on reason, the rational system of his predecessors and contemporaries. Hence, setting aside the metaphysical problem with the self in Book One, Hume tries to strike a balance between thought and passion. Hume hints at this resolution towards the end of the section 'Of Personal Identity' in Book One when he notes that passions "corroborate" the work of the imagination, "by

---

[30] T. 190-1.
[31] T. 276. I owe this quotation to Spada (1998), pp. 196-197
[32] T.276.
[33] T. 303
[34] Bricke (1980), p. 99.

the making our distinct perceptions influence each other, and by giving us a present concern for our past or future pains or pleasures."[35] Hence, in Part One of Book Two he writes that "'Tis evident, that the idea, or rather impression of ourselves is always intimately present with us, and that our consciousness gives us so lively a conception of our own person," that "'tis not possible to imagine, that any thing can in this particular go beyond it."[36] It can be said that it is our *passions* that make us the persons we are, not our perceptions and their fainter copies. Hume's scepticism concerns primarily our cognitive, not our emotional lives.

[35] T. 261.
[36] T. 317.

6.2. Emotion and the Diachronic Outlook of the Self

6.2.1. Passions and the Idea of the Self

So far I have shown that the episodic view of the self in Book One cannot explain the unity of a person's life. In what follows I shall show that the diachronic view in Book Two fares better in explaining it. According to this diachronic view, a person lives his or her life by naturally experiencing himself or herself as something existing in the past and future. We can find a hint at this idea in Hume's view in Book Two.

In Book One, Hume remarks that "the true idea of the human mind is to consider it as a system of different perceptions, or different existences, which are linked together by the relation of cause and effect."[37] It is clear that this theory of the self cannot lend support to the diachronic view. Since Hume denies that any substance identical to me, now, existed in the past or will exist in the future, one might argue that there is only my present self to be concerned about. If this is the case, we can say that I am not responsible for my past action and not concerned about the future consequences. But this criticism would not be correct, since the self which is extended through time emerges from the discussion of personal identity in Book Two of the *Treatise*.[38] According to Hume's view in Book One, *my* past includes those perceptions, thoughts, and actions related to my present self, in terms of causation and resemblance. Now, in Book Two of the *Treatise,* Hume attempts to explain why those past perceptions, thoughts, and actions affect my present *feelings*, and therefore why I am interested in

[37] T. 216.

[38] J. L. McIntyre focuses on Hume's view on the self in Book Two to make sense of the self which is extended through time. In my analysis of the diachronic outlook of the self, I am indebted to her discussion. See McIntyre (1989), pp. 545-557.

them and they are important to me. This is the way in which the self is extended through the past and present. How then is the self extended to the future? *I* am extended to the future, according to Hume, for I act from concern for a future collection of perceptions that will have a relation for me, now, the way in which I relate to past. In this way Hume provides an account of personal identity as it regards the passions in Book Two.

Now Hume classifies two categories of passions when he discusses the past and the future regarding self-concern: indirect and direct passions. He attempts to explain our concern with the *past* in terms of so-called indirect passions - pride and humility, love and hatred. Now a question arises: how then do we have the conception of ourselves in terms of the indirect passions? It can be said that repeated experiences of pride in something would lead me to have the associated conception of myself as the owner of that which make me proud. For example, if our home-owning is something in which we have taken pride repeatedly, then our pride in our valuable things might turn our attention toward ourselves as homeowners. In this respect, we can say that pride functions to convert the belief that I own a house, for example, into the belief that I am a houseowner.[39] If this is true, we can say that Hume wants to distinguish between explanatory beliefs as belonging to someone, and identificatory ones as defining him. On this interpretation, the latter, which I call 'identificatory belief,' engages with the "concern we take in ourselves"[40], whereas the former, which I call 'an explanatory

[39] Donald Ainslie takes a similar view to mine in interpreting Hume's view of the genesis of the idea of the self in terms of pride. See Ainslie (1999), p. 483. But many commentators take a different view in reading Hume's view of the generation of the idea of self. For example, according to J.L. McIntyre, we have the idea of the self in terms of pride as follows: my beautiful or ugly house, for example, which belongs to me, is a source of pleasure or pain. This results in a feelings of pride of humility. These passions have the self for their object, so a second relationship to the self is set up. The pleasure caused by something related to me is associated with the pleasurable feeling of pride; the cause itself is related to the self. These two associations reinforce each other, generating the feeling of pride, and turning my attention towards myself. See McIntyre (1989), p. 551.

[40] T. 253.

belief,' does not.[41] Now although we have the conception of the self in terms of pride in Book Two, the problem is, it seems to me that this view cannot explain the unity of a person's life, unless the self is extended over time. Hence another conception of the self is needed.

6.2.2. The Self in Diachronicity

Now let us try to find the conception of the self in diachronicity. When Hume tries to describe the origin of the indirect passions - pride, humility, love and hatred,- we can see why my past actions continue to concern me, even though they are not the actions of the self which is strictly identical through time. My past action will affect me through my feelings. My shame, for example, in a past wrongdoing, is the result of two conditions: that the past act is related to me, and that elicits feelings of pain or displeasure. Hence it can be said that when we feel pain due to our projects which are frustrated, for example, 'I behaved idiotically', this can be given as a reason for my feeling terrible today.

We can see that pride and humility are related to my present self in the social realm. Hume claims that a feeling of pride can be persistent when it is "seconded by the sentiments and opinions of others"[42] through sympathy. If others can take pleasure in it, it allows us to maintain our pride. Accordingly, it can be said that whether our actions are pleasing or displeasing to others is sympathetically communicated to us.[43] As a result of this communication, we have a good or bad reputation, which is itself a source

[41] In interpreting Hume's view of 'pride', G. Taylor also makes this classification. I owe this classification to Taylor. See Taylor (1985), p. 27 and p. 34.
[42] T. 316.
[43] Ibid.

of pride or humility. If this is so, we can say that reputation refers to a person's past. Furthermore, we respect another's love and approval of our actions and character, since it can be the objective evaluation of our frail self-esteem.[44] If this is so we can say that the reason why we are concerned with reputation is because it has the more public aspect of concern with one's past.[45]

When we gloss Hume's analysis of pride and humility in connection with our concern with the past, we can appreciate personal identity with regard to our past and present. Personal identity with regard to our past and present also appears in Book One. When Hume discusses personal identity in Book One, he focuses on the relationship of the present self to its past. In his discussion of associating a series of *past* perceptions in Book One, he appeals to resemblance and causation which he calls the principle of association of our perceptions. In doing so, *memory* plays a key role. If so, we can say that memory helps to explain our past, but it pushes the self back, not forward. However, Hume says in Book Two that the self is extended into the future through the passions. Thus we can say that passions can help to provide a kind of causal link to the future.[46] In Book Two he discusses the relationship of the self to the future such that the idea of the future can affect the will and intention.[47] Now the way in which our intention and will is affected by the idea of future is in terms of passions: our concern for the future. According to Hume, there are many passions that are concerned with the future, for example, hope and fear.[48]

In unifying the self diachronically, Book Two, like Book One, takes causation to be the

[44] T. 316.
[45] As interpreted by McIntyre (1989), p. 552.
[46] T. 261.
[47] T. 430.
[48] T. 442-5.

primary relation between the perceptions, thoughts, and actions that construct the self through time. However, in Book Two Hume remarks on another causal relation in explaining the self—the relation between intention and action. Here we should take notice of his conception of causality. Having established 'necessity' as 'the constant conjunction, or the inference of the mind from one event to the other,' Hume claims that 'necessity, in both these senses, has universally...been allow'd to belong to the will of man, and no one has ever pretended to deny, that we can draw inferences concerning human actions, and that those inferences are founded on the experience'd union of like actions with like motives and circumstances.'[49] Thus understood, we can say that for Hume my present self will persist in terms of the causal connection between something central to my present self –that is, my intentions, motives and character- *and* the actions and circumstances of a future person.

Now why for Hume is a person (the self) extended into the future, and why are the passions affected by our ideas of the future? The reason according to Hume, is because we are concerned for our future and want to regulate our present action for the sake of our future interest.[50] Now the way in which our present action is regulated by a future interest is that the *prospect* of pleasure or pain, reward or punishment, influences our choice of action.[51] If this is true, it can be said that our concern for the future is explained as being based on our *desires* - pursuing pleasure, or reward and avoiding pain or punishment - that we have *now*.[52] Then we can say that we see ourselves in the future in the actions that follow from my present desires - intentions, motives and

[49] T. 409.
[50] T. 534-36.
[51] T. 414.
[52] T. 410.

character. Hence when I think of *myself* in the future I think of the actions that follow from my motives, intentions and character.

Now in order for Hume to say that a self is extended through time, we need to look at Hume's conception of the past, present, and the future. In Book Two Hume says:

> We always follow the succession of time in placing our ideas... When the object is past, the progression of the thought in passing to it from the present is contrary to nature, as proceeding from one point of time to that which is preceding... On the other hand, when we turn our thought to a future, our fancy flows along the stream of time, and arrives at the object by an order, which seems most natural, passing always from one point of time to that which is immediately posterior to it.[53]

This is because, according to Hume, our thought always follows the succession of time, and therefore moves forward more easily than backward.[54] Hence he says:

> ...we conceive the future as flowing every moment nearer us, and the past as retiring. An equal distance, therefore, in the past and in the future, has not the same effect on the imagination; and that because we consider the one as continually encreasing, and the other as continually diminishing[55] .

In the next section, I shall show that Hume's conception of time is defective in that it misses the role of memory's recreation of one's past. Then I shall present an alternative way to solve Hume's difficulty by considering the role of 'experiential memory.' But meanwhile, I will focus on our issue: how passions can help to exhibit the self extended into future.

Hume argues that passions can help the self forward by descending with greater facility

[53] T. 430-1.
[54] T. 430.
[55] T. 432.

than they ascend,[56] whereas the imagination is 'necessitated every moment to reflection on the present.'[57] This is why for Hume passions can help to explain why we have so much concern for the future. But Hume does not mean that we do always act due to concern with future interest. What he is getting at is that we identify a future interest as our own due to adopting a moral (impartial) point of view. Although we tend to ignore long-terms interest in favour of immediate advantages, if we take an impartial point of view through operating sympathy, then we can overcome our tendency in favour of short-term self-interest and understand why we identify a future interest as our own, even though we are not strictly identical through time.

However, as I have raised a question in the previous Chapter (Chapter 5), how in Hume's view could we take an impartial point of view? If this were possible, one might argue that it could be so-called 'prudence.' Thomas Nagel argues that we have our conception of prudence in terms of the following ideas that we have of ourselves: "a belief in the reality of the future...a conception of oneself as temporally extended,"[58] a belief that "other times are equally real."[59] But if this were true, how then do we know about the reality of the past, and the future? On the interpretation of Hume established so far, it is in terms of our evaluative attitude (emotions) toward past pleasure and pain, - for example, pursuing good reputation and avoiding punishment,- and our anticipation of future pain. Furthermore, our differing attitudes can be comprehensible only if we conceive of events flowing towards persons (or self), and of the person passing in and out of periods of pleasant and unpleasant experiences. Later in this Chapter I shall

[56] T. 342.
[57] T. 428.
[58] Nagel (1970), p. 71.
[59] Ibid., p. 88.

develop this idea by illuminating some distinctive temporal characteristics of memory concerning emotion and a temporal orientation that is forward-looking rather than exclusively retrospective in orientation.

We have seen so far that in Book Two of Hume's *Treatise* the passions play a central role in creating the self which is unified through time. Now if this is the true story of the self, why then does Hume confesses to his labyrinth difficulty in the Appendix?

6.2.3. Hume's Real Difficulty

I have argued that by identity Hume means identity through time and he claims that the identity of individuals through time is not perfect. According to Hume, a thing at one time and the very same thing at a different time is not perfectly identical. They are wholly distinct temporal parts. Hence we can say that for Hume identity through time is the relation which binds these temporal parts into a whole. I have then argued that the problem for Hume is by what connection the idea of a self (a person) is bound into a whole. I have shown that in order to connect a self as a whole Hume appeals to memory and causation. Furthermore, we have seen that causation is essential in the production of the identity of a temporally extended self, since for Hume causation is the only relation "that can be trac'd beyond our senses, and informs us of existences and objects, which we do not [now] see or feel,"[60] the only thing that can yield a "conclusion beyond the [present] impressions of our senses."[61] However, if we accept this view, we are committed to circularity problem.

According to Hume's view of causation, causality is not a relation we perceive between

[60] Hume, T. 74.
[61] Ibid.

objects; rather we regard a pair of objects as related as cause and effect. When we have observed a constant conjunction of similar pairs of contiguous objects, our experience shows us that they have been connected in the past. But it cannot tell us that they will continue to be connected in the future. As result of this observed constant conjunction we are disposed to expect the second member of the pair on perceiving the first. Hence, Hume concludes that the necessary connection that we attribute to causally connected objects is a projection onto them due to our mind's disposition to connect its ideas.[62] Now in order for us to have the belief that my perceptions, as the bundled elements which constitute an episodic self, are causally linked (Book One), they would have to exhibit a multitude of long-standing constant conjunctions. But in order to see a multitude of constant conjunctions, there must be a self who observes that conjunction. Hence Hume's argument in Book One seems to be circular: it presupposes the existence of the enduring self which it aims to establish.

Now in Book Two, like in Book One, Hume also appeals to causation in order to explain the connection between the present and future self, namely, in terms of intention and action. To say that I am extended through the present and future is to say that there must be 'the experienced union of like actions with like motives and circumstances' and that we can draw inferences in the mind from the one to the other.[63] However, in order to draw the inference, there must be the self who observes it. In Book Two Hume also appeals to his conception of causation in order to explain the connection between the past and present self when he discusses the indirect passions. According to Hume, if most people we have seen take pride in certain things (in Hume's term, causes), for

[62] T. 165-7.
[63] T. 409.

example, some valuable possessions, then we can expect to find any person possessing riches to be proud of them. It seems to me, however, that there must be more to say this. For not only do we expect the person who takes pride in a certain thing - for example, their riches - to feel pride, there must also be a self who observes the expectation of the repeatedly experienced regularities. Hence once more, like in Book One, the argument in Book Two seems to be circular.

The circularity problem can be avoided if we consider Hume's remarks in the Appendix. Hume says: "[d]id our perceptions inhere in something simple and individual, or did the mind perceive some real connextion among them, there wou'd be no difficulty in the case."[64] In other words, Hume sees two ways of providing an account of a unified self, that is, a bundle of perceptions which is unified at one point in time, and unified across time. One way of doing this would be if the individual perceptions were themselves the intentional objects of a 'meta-self', extrinsic to the bundle of perceptions. This would be a transcendental account of the self, similar to that of Kant's when he claims that the 'I think' accompanies all our judgments. The alternative is to find some intrinsic, as opposed to extrinsic, connection between the perceptions. As Hume's metaphysics cannot allow for either possibilities he is left with the insoluble problem. Hence Hume says:

> I must plead the privilege of a sceptic, and confess that this difficulty is too hard for my understanding.[65]

In what follows, I suggest an alternative way to solve Hume's problem, which goes beyond his system. Hume rejects our belief in the simplicity of the soul in Book One,

[64] T. 636.
[65] T. 636.

since the doctrine of the mind's simplicity is the doctrine that it has no parts, but according to Hume's view of mind as a bundle of perceptions, the self is a collection of parts.[66] In order to challenge the doctrine of the mind's simplicity, he appeals to our belief in the self's diachronic identity. His theory of passion in Book Two supports the diachronicity of the self, for the passions which construct the self as a bundle extend through time. Although he accepts the diachronic conception of the self in Book Two, the deep problem, it seems to me, is his recognition of a problem about how the perceptions or passions in Book Two - as long as passions are perceptions – to which we attribute identity are held in consciousness in order for us to be able to do this to them. In other words, the real issue for Hume is to explain the synchronic unity of those perceptions that make up the process of judging all the past perceptions to form a unit.[67] In order to solve Hume's difficulty, in what follows, I suggest that Hume must take notice of 'the experiential content of memory' which is embedded in certain *feelings*. The fact that experiential memory plays a central role in explaining personal identity has been neglected by many philosophers in the mental connectedness tradition. In order to understand memory as experiential, we need to understand the affective element attached to some memories. I argue that memory affects not only my past thought but also my past emotions and those emotions deriving from the past stay on to affect my whole being and my future. Thus, for example, the shame or regret that I experience when I remember my past wrongdoing represents the whole of me up to now as well as at that time. In this way, I shall argue, we bring the person under the influence of his past, such that the past is reconstructed and then influences the future. Hence, a

[66] T. 251-3.
[67] A similar point has been made by T. Penelhum and D. Ainslie. See Penelhum (2000), p. 116 and Ainslie (2001), pp. 557-78.

life- the diachronic expansion of a person- is unified.

6.3. Three theses that the traditional psychological connectedness criterion overlooked.

The criteria of personal identity in much recent work have divided into two domains: bodily continuity and psychological continuity theses respectively. The one have sought the criterion in the domain of the physical, identifying personal identity with sameness of body; the other have sought it in the domain of the mental, equating personal identity with continuity of memory. The latter view has been represented in recent literature by psychological continuity theories of personal identity that take Locke's original insight and develop it into a more detailed identity criterion. These memory-based theories of personal identity take memory as providing a retrospective form from present experiences to past experiences when they explain the unity of a person's life.

In what follows, I deny the psychological continuity theory described above. Instead, I will defend a psychological continuity theory which memory has as much as to do with a person's future as with person's past concerning emotions. In order to do this job, the view which follows will elaborate 'experiential memory' suggested by Wollheim.

6.3.1. Thesis A: Memory is a faculty

It can be said that if we do not see memory as a faculty, we have no idea of how to communicate or play the piano, or dance. But the traditional psychological connectedness criterion such as Locke and Hume present, overlooks this idea by conceiving of memory as a kind of store. The assumption of the view which I shall defend is that it is not the occurrent psychological states alone that are part of one's

psychological make-up, but underlying traits and dispositions as well.

Another assumption which I shall defend is that there is more to mental activity than that of which we are aware. We cannot doubt the existence of unconscious mental activity, even if we are not ourselves committed to Freud. Suppose when a forgotten name comes to mind after I have given up looking for it from my memory storage which I suppose that the name was stored in my mind all the time. Furthermore suppose that the name is the correct name and the correct name were to come into my mind for no reason at all. Moreover, suppose that the name which comes to my mind is unconnected with the fact that I learned in the past that the person had that name. If this is true, according to Hume's and Locke's view, it would be magical. In order to explain such phenomena I suggest we have to posit the existence of both the unconsciously stored name and the unconscious activity of looking for it.

6.3.2. Thesis B: Memory is a disposition

Memory is broadly understood as a mental phenomenon. Memory as a kind of mental phenomenon has been divided into two categories by a number of philosophers: memory, on the one hand, is a mental state, on the other hand, a mental disposition. What is the difference between a mental disposition and a mental state? A mental disposition differs from a mental state in that mental states are events, in the sense that they can be ascribed a discrete temporal location, whereas mental dispositions cannot. Mental states are "episodic or transient phenomena," whereas mental dispositions are persisting and they "manifest themselves intermittently"[68] . A disposition explains

[68] Wollheim (1984), pp. 34-5.

events that occur in a person's life whilst underlying mental states. I shall try to illustrate later, by using memory, the way in which a mental disposition and a mental state interact, that is, how a mental disposition relates to a mental state.

In order for my current mental state and the earlier experience to be connected some kind of causal connection is needed. But what kind of connection?: not just any causal connection, but the current mental state and the past event have to be connected in the right way. This idea can be illuminated in terms of what Wollheim calls 'dependency relation.' What then is the 'dependency relation'?

The way in which dispositions manifest themselves is in terms of their specific causal influence in a person's life. Now if we see memory as a disposition, it can be said that it has causal influence in the way that it bears on an earlier event. Wollheim calls this the dependency relation. There is a three term relation to be found in the orbit of experiential memory. First, "the original event causes the disposition to be." Second, "the disposition causes the mental state to occur."[69] Third, the mediating disposition has the role of keeping an event "causally alive"[70] (Ibid). Here we can say that "causation occurs not just once but twice in the orbit of experiential memory, or that the memory onwardly transmits the causal influence of the remembered event."[71]

Now once we view memory as a disposition, it can be said that it exerts a causal influence over the manifestations of a person's later life. Hence we can explain a person's future life in terms of memory as disposition. Later I shall show how such an account might be developed.

---

[69] Wollheim (1984), p. 98.
[70] Ibid.
[71] Wollheim (1980), p. 308.

6.3.3. Thesis C: Memory is embedded in certain *feelings.*

William James defines the concept of memory in terms of "the knowledge of an event, or fact with the additional consciousness that we have thought or experienced it before."[72] He presents two sense of experience as follows:

> A succession of feelings, in and of itself, is not a feeling of succession. And since, to our successive feelings, a feeling of their succession is added, that must be treated as an additional fact requiring its own elucidation...[73]

Here James seems to suggest that the feeling of succession cannot be much like sensations or emotions. He suggests that it is one thing to be presented by a succession of feelings, it is another to appreciate it as a succession. On this view, memory plays an important role for taking notice of succession, since without the 'taking notice' (the act of mind), the mere occurrence of successive perceptions would not provide us with self-conception. But 'taking notice' is not itself a feeling. Hence in order to appreciate this idea, we need to take a close look at the thesis C - Memory is embedded in certain feelings. I shall try to show later what the thesis C means.

The theses A, B, and C described so far have been neglected by the traditional psychological connectedness theorists. Hence in order to lend support to the three theses, in what follows, I shall propose four supporting claims, following Wollheim.

a) an event is one that I have lived through:

b) an event is remembered *as* it was experienced and not just as it occurred:

---

[72] James, W. (1890/1950), p. 648.
[73] Ibid., pp. 628-9.

c) when we remember an event experientially an event is remembered from a point of view and this point of view is represented within the memory.[74]

d) an affective tendency is attached to some memory.[75]

Experiential memories take the form of 'I remember doing such-and such,' or 'I remember your, or X's, doing such-and-such', rather than 'I remember that I *X*-ed'[76]. Now in order to see the distinction between 'I remember *X*-ing' and 'I remember that I *X*-ed', let us take an example: the distinction between a person remembering herself watching a horror movie and remembering *that* she watched a certain horror movie. We can say that the former is an experiential memory that reminds her of what it was *like* to watch such a horror movie. In experientially remembering watching a horror movie, a person thinks and *feels* as she experienced when she watched it. When I experientially remember watching a horror movie those years ago, I tend to remember also the thoughts I had, the feelings I experienced, while watching the movie, and when I do so, the thoughts and feelings tend to come back over me. I once again many sense my muscles tensed, my pulse quickening, and adrenaline flowing, and those sensations deriving from the past stay on to affect my whole being and my future. The view described here does not mean that whenever I relive past events I should feel exactly as I did then. What Wollheim is intending to describe is the fact that when I remember something in my past i) I draw my attention to *what it was like to be* experiencing something from my point of view. ii) The affection that a memory preserves reconnects me with a broader affective dimension, which is that of a specific time in my life. Hence

[74] Wollheim (1984), p. 104.
[75] Wollheim (1980), pp.309-10.
[76] Wollheim (1979), p. 195.

for example, when I remember my failing some examination, the despair, shamefulness, and regret that I experience involves the sum of my experience from that point on, considered as a unified whole. My past experiences affect my present condition - my sense of shamefulness - and this feeling expresses my awareness of the past. Then my concern for my future - a desire for a successful life- causes painful feelings and emotions, for example, regret, in my present condition. This is possible in terms of memory's recreation of past and creation of future. Furthermore we can say that this kind of unity of consciousness over time requires unity of consciousness at one time. In what follows I shall show how such an account might go in more detail.

6.4. Identity and Change

6.4.1. Transitivity implies change.

Now one might wonder how the view described here might be put to use in accounting for a person's emotional change and their identity. I can recognize that my emotions or feelings have changed if I look back on past experience. For example, when I was 5 years old, I was scared of a monster which appeared in a horror movie, but now I am not. Change can also be seen in, for example, recovery from shock or recovery from grief. How then can Wollheim explain our emotional change- for example, the discordance between being afraid and having recovered- by appealing to experiential memory?

If we apply the above claim b) – an event is remembered *as* it was experienced and not just as it occurred – to explain one's identity we would be left with the following: what it is to be a psychologically identical being is to be a creature who can be, from a first-person, phenomenological perspective, in the same state. I shall argue that this is the problem of synchronic unity. The synchronic unity question is: in virtue of what do different experiences or mental states occurring at the same time count as belonging to one and the same person? Now even if the view which describes the above can explain the synchronic unity, how can the view explain personal identity through time? Questions about identity over time can be said to be questions about the diachronic unity of consciousness of some kind, e.g., persons or tables. In order to understand personal identity over time, we should consider the above claims a), c), and d) in conjunction. Wollheim would argue that it is not occurrent psychological states alone that are part of one's psychological make-up, but underlying traits and dispositions as

well - this concerns the above claim a), namely, an event is one that I have lived through. Furthermore, if we accept the view that memory has as much as to do with a person's future as with a person's past with reference to emotions, we can understand a person's identity over time. This concerns the above claim c), namely, when we remember an event experientially an event is remembered from a point of view and this point of view is represented within the memory. In what follows I shall argue that experiential memory plays a vital role in explaining this problem.

Now some people say that personal identity is constituted by sameness of the psychological subject, but it is not obvious what is meant by this. Locke, for example, claims that personal identity consists, not in sameness of substance, but in 'sameness of consciousness' and holds that continuation of the consciousness itself, not that of body or soul in which that consciousness resides, is necessary for identity. Butler claims that consciousness is 'successive' and, therefore, "cannot be the same in any two moments, nor consequently the personality constituted by it."[77] Reid argues that "consciousness, and every kind of thought are transient and momentary, and have no continued existence..."[78] Consciousness is, at least according to these theorists, a momentary phenomenon and does not remain numerically the same over time. If this is the case, it is not clear what it means to say that the same consciousness has continued to exist.

In order to answer the question of what it means to say that the same consciouness has continued to exist, Locke claims that if you cannot currently remember the past thoughts and action of some person, then you simply are not the same person who had

[77] Joseph Butler, "Of the Personal Identity", in John Perry, (1975), p. 102.
[78] Thomas Reid, "Of Mr. Locke's Account of Our Personal Identity," in John Perry, (1975), p. 116.

and did them.[79] Here what Locke makes necessary for identity with a 'past self' is, as Shoemaker notes, "not that one remember the actions and experiences for that past self but that one have 'memory continuity' with that past self."[80] Reid tries to show this with his 'brave officer' example.[81] At a certain time a boy steals apples. Years later the same person, now a young officer, performs an act of bravery in battle, remembering still his boyhood's apples. Many years later our man is an elderly general, who remembers the act of bravery in battle, but no longer remembers the incident of robbing apples. Reid charges that on Locke's theory the general both is and is not the same person as the small boy; he is the same because he is identical to the young officer who is identical to the small boy (because identity is transitive), but the general is not the same person as the small bay because he has no memory of the boyhood incident.[82]

Locke himself acknowledges that 'the same Consciousness' is not the same 'individual Action' but 'a present representation of a past Action.'[83] Here we should allow, as Shoemaker points out, that "one's current person-stage contains a memory of something even if one has temporarily forgotten that thing, as long as one has the potentiality of remembering it." "In such a case the stage will retain a 'memory trace' that is the basis of that potentiality."[84] I shall argue that the Lockean memory criterion of personal identity which is related to the issue of *transitivity* implies 'change,' and show that

[79] For Locke, 'person' is a forensic term, which is related to our practices of attributing responsibility and distributing rewards and punishments. According to him, a person should not be held responsible and punished for his actions which he cannot remember of his deed.
[80] Shoemaker (1984), p.81.
[81] John Perry (1975), pp. 114-15.
[82] Locke might reply to Reid as follows: the general is not the same *person* as the boy, while pointing out that the general is nonetheless the same *human being* or *man* as the boy. This is pointed by E. J. Lowe. See Lowe (1995), p. 113.
[83] Locke, Essay Concerning Human Understanding, p. 337.
[84] Shoemaker Ibid.

change is constitutive of, rather than a threat of personal identity.[85]

In this respect, we should recall Locke's remarks that present consciousness includes in itself temporally distant actions and experiences in precisely the same way it does present ones. He says that a present self "extends it *self* beyond present Existence to what is past, only by consciousness, whereby it becomes concerned and accountable, owns and imputes to it *self* past Actions, just upon the same ground, and for the same reason, that it does the present"[86]

6.4.2. Recognition of Emotional Change: From the Past to the Present

With understanding of the psychological sameness described in the above paragraphs, we can explain a person's emotional change and his or her identity. According to Wollheim, emotional change is possible because when my affective store from which the affective tone comes was reinforced by an experiential memory, the impact of the memory was mediated by my current *beliefs, desires* and *feelings*. Hence, it can be said that I should not be scared of the monster in the horror movie due to my current *beliefs* that it is a fictional character, even though I was scared of it in my childhood. Furthermore, I am not going through grief over the death of my mother due to my present *recognition* that the perceptual image of my mother is weakened; the *belief* that I can no longer see her; prospect of my future - *desire* for my well-being in her absent. Hence we can say that I am psychologically identical being, from a first - person phenomenological perspective even though my emotions or experiences have been changed.

---

[85] Some people read Locke's transitivity of consciousness as one what William James (1890, ch. 10) later calls 'stream of consciousness.' I share with this view. Cf. See E. J. Lowe(1995), p. 114

[86] Locke, Ibid., p. 346.

Wollheim claims that change is constitutive of, rather than a threat to, a person's identity. Then we would be left to clarify how change is constitutive of personal identity. The change must occur in order that a person not be for an indefinite period struck down, for example, by grief or disappointment. In this respect, Wollheim claims that "it is in the nature of persons that they change," since time is not merely a static but a transiant thing and a development of the new out of the past and the present. Hence "as they live, they necessarily change, and therefore it would be a merit in a criterion of personal identity." Thus Wollheim argues that the mental connectedness criterion of personal identity must permit change. "It must exhibit how change is compatible with identity"[87]. Let us consider how such an account might be developed.

It might be said that emotional change is a kind of recovering. But once a person notices herself to have recovered, for example, from deep grief or shock, she could say that she is not the same as the person who once grieved or shocked. That is, she has changed. But the recognition that her feelings have been changed with passage of time is not itself a feeling. Then what is it? One might say that recognition of emotional change is disaffection with oneself. When someone, for example, finds herself not only changed but also finds that she has become an alien to herself many times over, she realizes that she is in a disaffective state. Martha C. Nussbaum sees this disaffective state as various cognitive features, such as a personal point of view, cognitive denseness, and cognitive freshness.[88] Now if this view is right, it can be said that the way in which a person has recovered from grief, for example, is when she is in the state of 'a detached, unemotional, or intellectual deliberation.'[89] Although I agree with Nussbaum in the

[87] Wollheim (1980), p. 313.
[88] I have elaborated these various cognitive features in Chapter 2 of this thesis.
[89] Nussbaum (2001), p. 24.

sense that the recognition of emotional change depends on cognitive features, I do not accept her view in the sense that cognitive elements are both necessary and sufficient for the requisite of such recognition, since it seems to me that we can recognize our emotional change in terms of noncognitive features, such as bodily feelings or affects. Now in order to see the difference between Nussbaum's and my view, let us look at Nussbaum's example of the role of emotional change in explaining a person's identity in her discussion of a cognitive account of emotional responses. When someone hears the news of her mother's death, Nussbaum describes the person's feelings of grief as follows[90]:

> <...> my feeling, of terrible tumultuousness, of being at the mercy of currents that swept over me without my consent or complete understanding; the feeling of being buffeted between hope and fear, as if between two warring winds; the feeling that very powerful forces were pulling the self apart, or tearing it limb from limb <...>
> [91]

Nussbaum in the above passage intends that non-cognitive or bodily feelings are not necessary or sufficient for elements of an emotion. Rather she takes evaluative recognition to be the constituent parts of an emotion, since she identifies emotion with judgment. In what follows I shall show that this account fails to render plausible of the role of non-cognitive element in our emotional experience. Consider the fact that the emotion of grief fades while the judgment remains unharmed. This shows us that there is something more to grieving than judgment. But Nussbaum ignores anything other than judgment by saying that the change which takes place when an emotion fades is a

[90] I have also used this quotation in Chapter 2 in order to challenge to strong cognitivism. I use this again for the sake of my argument.
[91] Nussbaum (2001) p.26

cognitive change. Nussbaum calls this features of emotion "cognitive freshness."[92] Nussbaum describes emotional changes in a several stages. First, cognitive freshness makes the given emotion a background emotion rather than a situational one. When grief progresses, for example, we find it becomes a background emotion rather than a situational emotion. What then is the difference between a background and a situational emotion? According to Nussbaum, the first is one that persists throughout situations of numerous kinds, whereas the other is one that is elicited in the context of some particular situation.[93] Now the second stage in the progress of mourning, Nussbaum writes, is a cognitive reconstruction. For example, she looks forward to celebrating her mother's birthday, then remembers that this is not to be, thus she has to amend her prospects and plans for the future. The recognition of her own future goals or projects shift the judgments of the importance of her mother into the past tense, hence her grief is gradually diminished by this judgmental change. Third, a fading of the imagination takes place as a result of the lost person's absence. The perceptual image of the object is weakened, she no longer sees her mother before her, leading the content of the emotion to shift. Although Nussbaum's account of the phenomenon of fading grief in terms of cognitive reconstruction seems to be plausible in explaining the fact that grief fades as time passes by, her view is defective in that she ignores the aspects of *felt urgency* and the *affective* side of the emotional experience. In order to appreciate cognition and affect in the recognition of our emotional change, in what follows I shall show that the affect – bodily feeling- also reveals our emotional change.

In order to lend support to this idea, I will use an example from *Remembrance of Things*

[92] Nussbaum (2001), pp.79-85.
[93] Nussbaum (2001), p.69.

Past by Proust.[94] When Marcel hears of the death of Albertine, he suffers great anguish. Marcel, who believed confidently just a moment before that he did not love her any longer, realizes with certainty that he loves Albertine. Proust tells us that Marcel deceives himself about love. If we say that recognition of our condition can best be attained by a detached, unemotional, and disaffective cognition, we cannot explain Marcel's shock of loss and the attendant welling up of pain. The news of Albertine's death brings him immediate anguish, and this removes his intellectual judgment that he did not love her any longer. He has been a victim of self-deception. Then he says:

> I had been mistaken in thinking that I could see clearly into my own heart. But this knowledge, which the shrewdest perceptions of the mind would not have given me, had now been brought to me, hard, glittering, strange, like a crystallised salt, by the abrupt reaction of pain.[95]

Marcel's feeling -anguish-is said to have the power, just in terms of its own *felt* quality, 'like a thunderbolt[96]', and forces his assent to his state of love. His state is revealed by his feelings, not his intellect. Proust tells us that our feelings are systematically buried by the workings of habit, the primary manner in which self-deception operates. The truth of Marcel's condition is brought about by his feelings. His painful anguish demonstrates his prior state of self-deception. If this is indeed a typical or common phenomenon, then it shows that a full account of our condition cannot be given simply by appeal to a cognitive-evaluative state. If so, the attempt by Nussbaum to reduce the role of cognition and affect in our recognition of emotional change to a single cognitive

[94] I have also used this example in Chapter 2. Here I use this in defense of my argument. In her Ph D thesis, Pillippha Byers also discusses a person's emotional changes and their identity in order to explain the connection between memory and personal biography, drawing on Proust's *In Searching of Lost Time.* See Byers (2005), esp. pp. 175-86.
[95] Proust (1981), Vol. III, p. 426.
[96] Ibid., Vol. III, p. 431.

element is implausible. If we accept this, we must concede that both cognition and affect play an important role in our mental lives.

A qualified cognitivist, for example, Peter Goldie claims that Nussbaum's cognitivism fails to recognize the fact that "how profoundly and systematically our emotional feelings can mislead us – of how the emotions can distort perception and reason."[97] In order to solve this difficulty, Goldie suggests that we should take notice of the aspect of 'feeling- towards' which is "unreflective engagement with the world beyond the body; it is not a consciousness of oneself, either of one's bodily condition or of oneself as experiencing an emotion." [98] According to him, feeling-toward shows someone undergoing an emotion without being reflectively aware of having that emotion. If we adopt the idea of feeling-toward, we can explain Marcel's case: Marcel has been experiencing love for Albertine without being reflectively aware of having that emotion until Francoise tells him of the death of Albertine; and then Marcel notices, or becomes aware of his feeling of love for her with the attendant painful feeling - anguish. Before hearing the news, Marcel had feelings of love toward Albertine but was not aware that this was so. If my argument so far is right, the claim that the recognition of one's condition is possible by referring only to a cognitive evaluative state is likely to fail. Hence, we can say that Nussbaum's attempt to reduce the cognitive and affective aspects of our recognition of emotional change to a single element, such as cognition is implausible.

6.4.3. Recognition of Emotional Change: From the Past-Present to the Future

[97] Goldie (2004), p.91.
[98] Ibid., p.96

Now I shall discuss how experiential memory can explain a person's future. Having established that memories, as dispositions, have a backward and a forward looking dimension, Wollheim argues that a memory criterion of personal identity is not merely an 'indicative' criterion: "it is creative in that, when it holds, it brings about or helps to bring about that for which it is a criterion." [99]

On this view, a person's memories are not merely instances of the spatio-temporal connections of psychological states, in the way that the traditional mental connectedness proponents, including Locke and Hume, propose. Rather, memory and the forward-looking dimension in the temporal unity of a person's life are interwoven.[100] The most distinctive feature for this is the emotional tone and richness of each person's life, since as Byers notes, for Proust, "memory is the emotionally textured expanse or register of a person's life, as without it the past would be empty, rather than emotionally nuanced, textured, and on occasion, devastating."[101] This idea suggests that in a person's life memory is deeply implicated in the forward-looking orientation or impetus of a person's life. Now in order to understand this idea, the implication that memory has perspective needs to be spelled out. This concerns my claim c: when we remember an event experientially an event is remembered from a point of view and this point of view is represented within the memory.

---

[99] Wollheim (1984), p. 109.

[100] I am indebted for this point to Philippa Byers (2005). Byers discusses the relation between memory and time drawing on Proust's *In Searing of Lost Time.* On Byers's analysis of Proust, the time that memory recalls is time that memory recreates. According to Byers's Proust, although the stuff of memory is the past, what is produced from the past, as the result of the labour of remembering, is new. Memory calls forth something new, rather than something repeated.(p. 170) This view implies that memory not only recreates time but also has a future. In this respect, Byers criticizes the traditional memory-based theories of personal identity, such as Locke's, where memory is characterized as presenting a retrospective form of unity. See Byers (2005), especially, Chapter 5, Section 5.5.

[101] See Byers(2005), p. 175

Now the discussion which follows will show the role of perspective or point of view in our emotional change. This idea will show us how experiential memory creates one's future. The reason why the emotional change is possible is that memory embeded in feelings also has a perspective, or a forward-looking dimension. In order to appreciate this point, let us look at the following example.

> Suppose that Kate wants to gain revenge on her ex-husband who had an affair with her friend, Lucy and left her. Kate is also jealous of Lucy, and wants to gain revenge on her as well. Kate knows that her ex-husband depended on her financially and also Lucy is impoverished. Kate desires to see him financially ruined so that he will regret leaving her. But one day Kate falls in love with a man whom she dreams of marrying. Then she realizes that she has become a person who is no longer obsessed with her ex-husband, and no longer jealous of Lucy.

Now if emotions do not also have perspective, then we cannot understand how Kate who has been in the grip of jealousy is eventually indifferent concerning the object of jealousy. But from Kate's self-concern in her future, this emotional change is understandable. Her self-concern and self-love, which possess perspective, have forward-looking dimensions. If we adopt my above claim a) – an event is something that I have lived through – then we can say that the person has been and will be persisting, since a person is only able to experience something on the basis of his or her past action or experiences and concern for his or her future life. Now the reason why Kate is psychologically continuous over time is because she is aware of the fact that her past action or experiences, and her concern for her future life, influence her present condition. Hence we can say that the recovery from the obsession or jealousy due to her awareness of a future with the other man shows the fact that Kate is a succession of an embodied subject who not only thinks but also feels and who has entered into past states and will enter into future states. If this is right, we can say that change does not imply

the loss of identity, rather change reveals how one and the same person is obsessed and miserable and then become no longer obsessed and miserable. This view implies that the unity of a person's life include changes, for our condition and experiences change.

In explaining change there are elements of both discovering and creation. Kate's recognition of her emotional change involves discovering something new in herself. Recognition of something new is awareness of both the fact that the past is differentiated from the present, and the fact that the future is the new creation of the past and present. In this respect, we can say that experiential memories of Kate's living with her ex-husband recreate a dimension of her past life and create her future life. She tends to remember the thought, and the feelings that she had while living with him with regret: "what a shame I lived with manipulated man!; I've wasted years of my life living with a man who wasn't my type." And her remembering drives her life towards the future. This shows us that experiential memory is not just confined to recalling events or experiences that the subject has experienced, but concerns the narrative structure of her life as whole.

One might raise an objection that although this kind of psychological criterion can explain the problem of unity of consciousness over time, it cannot lend support to the unity of consciousness at a time, since in order for me to stitch consciously together the events of my life into a 'whole self' I have to already know which events those are. So some notion of personal identity is presupposed. In other words, one might say that although this view can explain the diachronic identity of a person, since diachronic identity can best be understood in terms of a *memory relation,* but the view which I have described so far cannot account for the synchronic identity, that is, what makes for

personal identity *at* a time.

In order to reply to this objection, let us consider the following example proposed by Marya Schechtman:

> Consider what it is for different psychological components to be part of the same consciousness at a time. ...the fact that the various beliefs, desires, images, and so on that are part of the same consciousness at a given time interact in such a way that each provides a context in which to interpret the others. The difference between one person thinking a six-word sentence and six people each thinking one word of that sentence is, at least in part, that in the former case each of the words provides an interpretive context that constrains the understanding of the others so that what is experienced is not the same as six separate words side by side, but a single six-word unit. The meanings of the other five words are, in a sense, already present in the consciousness of each word when one is conscious of such a sentence.[102]

On the basis of the unity of consciousness at one time described here, we can extend unity of consciousness over time. Since the latter involves "[weaving] temporally distant actions and experiences into the network of one's conception of oneself."[103] Hence, extending the unity of consciousness over time is possible in terms of the fact that our past action or experiences, and our concern for the future, influence our present conditions and experiences. Thus, we can say that in a synchronic case, a person is 'jointly' conscious of all the things of which he is currently conscious,[104] while in a diachronic case, the person at a later time has a first-person memory of some conscious thought, experience or action which occurred to them at some earlier time. Hence she makes the unity of consciousness, since it is not the case that "each time a person has an

[102] Schechtman (1994) p.208
[103] Ibid.
[104] In this respect, E. J. Lowe presents a useful example: if I am conscious of a pain in my toe, and simultaneously conscious of thinking about a philosophical problem, then I am conscious of thinking of the problem while having the pain. See Lowe (1995), p. 109.

experience she consciously considers the past and future and decides how to understand the present on that basis, any more than in understanding a six word sentence one need always consciously attend to the interactions of the words."[105]

[105] Schechtman (1994), pp.208-9.

6.5. Difficulty

Finally, I want to consider another objection to the view which I have supported. One might say that this kind of psychological criterion of personal identity can be faced with the so-called problem of personal fission, namely, splitting one person, P1, into two distinct persons, P2 and P3."[106] If we adopt the psychological connectedness criterion, we can say that P 2 and P3 are identical with P1 in that both P2 and P3 are related psychologically to P1. Then it is logically possible that someone else will exist tomorrow who is related to me psychologically in just the same way in which I am related to my self of yesterday. Now if we apply the problem of fission to memory, the difficulty seems to be apparent. Suppose that both persons, P2 and P3, into which the original person P1 divides, seem to remember doing what P1 did. How would the proponent of the psychological connectedness criterion deal with this matter? Now in order to explain this phenomenon, some people appeal to the concept of 'quasi-memory.' Derek Parfit suggests the following definition of the concept of 'quasi-memory.'

I have an accurate qusai-memory of a past experience if

a) I seem to remember having an experience

b) someone did have this experience

and

c) my apparent memory is causally dependent, in the right kind of way, on that past

[106] Many people discuss cases of fission, for example, Shoemaker (1963); David Wiggins (1967); Parfit, in *Philosophical Review,* LXXX; John Perry (1972); Williams (1973); Lowe (2000).

experience.

On this definition, ordinary memories are a sub-class of quasi-memories. They are quasi-memories of our own past experiences.[107] Now if this definition is right, it can be said that quasi-memory is a broader sense than that of experiential memory. My experiential memory of doing something can only be justified if it was indeed done by myself, but I can be justified my quasi-memory of doing something that was done by someone else can be justified, if I am related psychologically to that person in the same way in which I am characteristically related to my earlier self, as a case of personal fission might show. But many people challenge to the notion of qusai-memory[108], since if personal fission is possible, then I might be a product of such fission, in which case many of my experiential 'memories' are in fact only quasi-memories.

In this respect, some philosophers defend the notion of quasi-memory in order to avoid the circularity objection which is attributed to any psychological criterion of personal identity which appeals to personal memory. In brief, the circularity problem can be summarized as follows: If we say that a person, P2 existing at one time, is identical with a person, P1, existing at an earlier time, only if P2 remembers doing some of the things done by P1, then we are committed to a circularity. Since as Lowe remarks, "a necessary condition of P2's remembering doing something is that P2 should actually have done that thing." "One cannot establish, then, that P2 does indeed remember doing something done by P1, until one has established that P2 is indeed identical with P1, which is what the criterion was supposed to enable us to determine."[109] One way that the

[107] Parfit (1987), p. 220.
[108] For the references for criticizing quasi-memory, see Evans (1982), pp. 242-8, Wollheim (1984), pp. 117-27, and Andy Hamilton (1995), pp. 332-49.
[109] Lowe (2000), p. 281-2.

circularity objection can be avoided is by replacing reference to memory by reference to quasi-memory.[110] But if we demonstrate that the concept of quasi-memory is incoherent, as Lowe notes, "the implication would seem to be that no non-circular psychological criterion of personal identity is available."[111]

Hence in what follows I present several reasons for demonstrating that concept of quasi-memory is incoherent in terms of the problem of fission.[112] One way to show the impossibility of fission concerns my claim b) - phenomenology: to show the fact that what it is like for one rememberer to remember experientially an event does not match what it is like for the other remember experientially to remember it. Since "the beliefs, desires and feelings of the rememberer, by which the impact of the memory is mediated, will differ qualitatively from those of the other potential rememberer."[113] In a similar vein, Lowe suggests a way to meet this kind of objection. According to him, if a person "remember his own thoughts, experiences, feelings and actions in a special way – that is, so- called a 'first person' way and if the person remembers them, as it were, 'from the inside', as episodes undergone *by him,"* this is different from the fact that the person "remember events in the life of another person only in an 'external' or 'third-person'

---

[110] Andy Hamilton attempts to undermine the circularity objection in terms of the IEM thesis , according to which a sincere memory-claim in the continuous verb form implies that either I o-ed or else no one o-ed. Following Noonan, he locates the force of the circularity objection in its affirmation of the epistemological priority of personal identity over memory. Then he formulates the circularity objection as follows: concept X is being used to explain concept Y, when in fact concept Y is presupposed by concept X, and so the order of explanation must be incorrect. Having understood the circularity objection in this way, he criticizes it on the ground that personal identity can appear conceptually prior only from a third-person viewpoint. "In the first-person case," he continues to writes, " in contrast, IEM says that when I judge on the basis of memory that I o-ed, it makes no sense for me to attempt to verify that it was myself as opposed to someone else who o-ed. There is no question, according to Hamilton, in this case of having to establish personal identity. (Andy Hamilton, *Memory and Body: A Study of Self-Consciousness*; forthcoming)

[111] Lowe (2000),. p. 282.

[112] In presenting reasons for rejecting the possibility of fission my argument depends on Wollheim's; see Wollheim (1979) and (1984). But some people present a different argument from Wollheim's against the genuine possibility of what is entertained in the fission thought experiment. E.g., see Wiggins (1976).

[113] Wollheim (1979), p. 230.

way, as episodes undergone by *somebody else.*"[114]

The second way to show the impossibility of fission concerns my claim d): *an affective tendency is attached to some memory.* Wollheim says that "the affective tone that accrues to the memory – and the impact of which is just what the beliefs, desires, and feelings –will differ qualitatively from that which would have accrued to it had the other entertained the memory."[115] And this is so because "when the rememberer's affective store from which the affective tone comes was reinforced by an experiential memory, the impact of the memory was mediated by his current beliefs, desires and feelings, which become increasingly peculiar to him."[116] However, this seems to make trouble with Wollheim rather than helping him. For, as Lowe notes, "if one *can* have first-person memories of episodes in the lives of other people, then, clearly, such memory does *not* , after all, provide a satisfactory criterion of personal identity across time."[117]

Hence, we need the third way to show the impossibility of fission. This concerns my claim a): that is, that if we take a memory to be disposition, a memory disposition which an ordinary rememberer has and a Q-memory disposition which a Q-rememberer has would manifest themselves differently, even though both depend on the same initiating event. If so, it is difficult to say that these respective dispositions would manifest identical experiential memory states in two different people. For these reasons, it seems to me that the notion of personal fission in terms of memory replication and memory transfer is implausible.

In order to do justice to this, Wollheim proposes a more specific argument against Q-

[114] Lowe (1995), p.109.
[115] Wollheim, Ibid., p. 230.
[116] Wollheim, Ibid.
[117] Lowe, Ibid., p.110.

memory theory. He discusses the reasons why the causal influence of memory in a person's life is inconsistent with the view of Q-memory theorists that a person could remember an event experienced by someone else. His argument against Q-memory shows us that memory is far more than the conveyor of information and keeping track of past experience.

Now let us look at Wollheim's arguments against Q-memory in more detail. He proposes a triad of claims against Q-memory as follows:[118]

a) Every centred (experiential) event-memory is a memory of an event in the life of the person who has the memory;

b) the event is remembered as the person in whose life it occurred experienced it;

c) and the person who remembers the event not only centrally (experientially) remembers it, he remembers it from his point of view.[119]

Q-memory theorists may agree on the above claim b) - phenomenology. But they overlook the above claims a) - memory as disposition - and c) a person has a point of view - which are central to what it is to lead the life of a person, and therefore they think that memory replication and memory transfer are possible. In what follows I shall show why this view is implausible.

If we reject the claim a), then we have to think that memory dispositions could originate in one person and be transferred to another. Suppose that I were to Q-remember my

[118] These claims are the same one that I have made about experiential memory before, where I have made a), b), c). In addition to these three claims, I have made claim d) to emphasize the role of feeling in experiential memory. In fact, Wollheim stresses the claim d) in his paper (1980).
[119] Wollheim (1984) p. 112.

mother as a girl travelling in Greece. It would have to be the case that a memory disposition, established in her by her living through the event that I then come to remember, migrates at some moment in its history from her to me. In principle this is possible. From the moment of this migration I would have the disposition, and the disposition would then be free to manifest itself in some appropriate mental state of mine.

But the trouble with this supposition is, firstly, that it is impossible that "mental dispositions could shift from one person to another and remain intact."[120] Since "dispositions form...a web or network: they are ancillary to one another, and there seems no method for determining what in the way of other dispositions, what in the way of beliefs, emotions, desires, fears, and other memories, would have to be transferred along with it if the original memory is to run across persons."[121] If I am to Q-remember my mother's childhood traveling, I should also have to have a familiarity with the thoughts she would be inclined to have when she looked around the city, or the taste of coffee she drunk, the feeling she had from the weather...and so on. In the Q-memory, such dispositions would be backgrounded, if not foregrounded. If this is so, would these dispositions disappear for ever, or for the duration of the Q-memory? If this is in turn true, how would this be engaged? Once the transfer of the memory disposition had been effected, by whatever means, the disposition would then, on suitable occasions, manifest itself in appropriate mental states. But what would a mental state that was appropriate for the manifestation of Q-memory be like? In order do answer this question, the Q-memory theorist might appeal to the claim c) that the person remembers the event

[120] Wollheim (1984) p. 113.
[121] Ibid.

from his point of view. But a Q-memory state is a first-person state and this state does not have to have the point of view of the person who has it.

Now finally I shall show a more serious difficulty that the Q-memory theorist might be encountered. The final objection to Q-memory theory concerns memory as the repertoire. If I were to experientially remember someone else, then I should have to assume that person's repertoire. But if this is so, "it would be incompatible with the character of memory that, before some kind of memory state could occur, I should first have to assume a repertoire."[122] But "it seems impossible that before a repertoire becomes available to memory, it should have to be assumed."[123]

For the reasons so far I have presented, I argue that the idea of memory replication and memory transfer are both implausible. If the argument I have so far presented is right, then the identity-neutral thesis of Q-memory theory and the view that ordinary memory is a subclass of Q-memory are both implausible.

[122] Wollheim (1984), p. 116.
[123] Ibid.

6.6. Conclusion

In this Chapter I have defended a psychological continuity theory of personal identity which claims that memory has as much as to do with a person's future as with a person's past with reference to emotions. Specifically, on the view which I have defended, memory does not simply invoke past events, but creates dispositions for future actions, thoughts and emotions. In this respect, we can say that memory is both past and future-regarding. In defence of this view, I have challenged the view of memory as retrospective only. This may cause problems for a linear conception of time because the time one remembers is the turbulent dynamics of the stream of consciousness and the past and future-regarding dimension are interwoven in the temporal unity of a person's life.

In order to illustrate this, I have elaborated on the conception of 'experiential memory' suggested by Wollheim. The characteristics of experiential memory are: first, it is a memory of an event lived through; second, it involves remembering an event *as* I experienced it, intrinsic to which are affective elements, as opposed to the mere recall of facts; third, it is a memory of an event from a person's own point of view. Hence, when we experientially remember a past event with feeling and perspective, our remembering not only *recreates* a dimension of our past life but also drives our future life in a certain direction. If my argument is right, we can say that experiential memory is not just confined to the recalling of events or experiences that the subject has experienced, but concerns the narrative structure of person's life as a whole.

I have argued that if we adopt the notion of experiential memory to identify mental

connectedness with personal identity, we bring the person under the influence of his past, such that the past is recreated and then influences the future. Furthermore, if a person consciously considers the past and future from the present perspective, the person creates the unity of consciousness at one time. Hence, a life - the diachronic expansion of a person - is unified. If this is right, I argue, the problem of personal identity could be resolved in terms of 'experiential memory.'

One might say that the view which I have discussed implies a narrative conception of the self. What then is the narrative conception of the self? According to Galen Strawson, "the notion of form finding is essential to being narrative and that goes beyond be diachronic."[124] Although form finding is an important element of something's being narrative, this does not make the difference between Narrative and Diachronic conception of the self, since being diachronic also puts a certain structure on one's life: one was in the past and will be in the future. The crucial difference between the Diachronic and Narrative outlook is the manner in which one's life is construed.[125] We can see one clear statement of the narrative thesis in Marya Schechtman's work. Schechtman claims that "a person creates his identity [only] by forming an autobiographical narrative – a story of his life." One must be in possession of a full and "explicit narrative [of one's life] to develop fully as a person."[126]

The proponent of a narrative thesis claims that autobiographical memory is an essentially constructive and reconstructive phenomenon rather than a merely reproductive one. There are cases in which this claim is proved by experimental

[124] Strawson (2004), p. 442.
[125] Ibid.
[126] Schechtman (1997), pp. 93, 119.

psychology.[127] In reconstructing our past life by autobiographical memory, *emotions*, especially moral emotions-for example, pride, self-love, conceit, shame, regret, remorse and guilt-are interconnected. This proves that in reconstructing our past life by experiential or autobiographical memory, social and moral aspects of emotion play an important role. As we have seen in this Chapter, my feeling of pride and humility makes ineliminable reference to other agents in my social group. Hume claims that a feeling of pride can be persistent when it is "seconded by the sentiments and opinions of others"[128] through sympathy. If others can take pleasure in it, it allows us to maintain our pride. Accordingly, it can be said that whether our actions are pleasing or displeasing to others is sympathetically communicated to us.[129] As a result of this communication, we have a good or bad reputation, which is itself a source of pride or humility. We can also say that Kate's guilt at putting her partner in the clinic is, as we have seen in Chapter 5, inextricably bound up with her social group. We can see what kind of person she is in terms of whether her action is pleasing or displeasing to others, and in terms of whether she has a good or bad reputation.

Alasdair McIntyre claims that "the unity of an individual life is the unity of a narrative embodied in a single life." He maintains that "the good life for man is the life spent in seeking for the good life for man."[130] G. Strawson calls this view 'the ethical narrative thesis.'[131] The ethical narrative thesis holds that in seeking the good life we need a narrative perspective. Although whether narrativity is necessary for a good life is controversial, I raise the ethical narrative thesis in order to resolve Hume's difficulty

[127] See Brewer (1988); see also, McCanley (1988).
[128] T. 316.
[129] Ibid.
[130] MacIntyre (1981), pp. 203-4.
[131] See Strawson (2004).

and suggest that Hume can be read as proposing an ethical narrative view in Books Two and Three. Since, when Hume's theory of passions is properly understood, we can say that identity through time is central to what we care about in our lives. "One thing", as John Campbell notes is, "I care about is what I have made of my life."[132] It seems to me that when I reflect on what I have made of my life I realize that I have been influenced by other people who have had more influence in making me what I am than I had on myself. To that extent it can be said that the education and upbringing of children play a very important role for a human being. The irony, it seems to me, is that we may have more chance to make others as we want them to be than we have to make ourselves what we want to be. If this is right, I would say that it would be more correct to describe us as "selves made by others" rather than "self-made selves."[133] To that extent I think Hume's view in Book Two leaves room to be interpreted as an ethical narrative thesis. However, the difference between Taylor and McIntyre, on the one hand, and Hume, on the other it seems to me, is that the former take a subjective point of view, whereas Hume takes an external point of view in terms of sympathy.

Finally I want to point out the difficulty of a psychological narrative thesis. The problem with autobiographical memory is that it involves the possibility of revision when a person reconstructs his or her past life. However, if revision is necessary to being a narrative, a person's life which is reconstructed by revision is not a genuinely identical one. If this is true, I do not want to commit myself to the narrative theory. Whatever narrativity means by the unity of a life, it seems to me that it is not the same

[132] Campbell (1994), p. 190.
[133] I attribute the latter expression, which I have taken from Dennett (1983), Ch. 4, to the psychological narrative thesis. In regard to the distinction between 'psychological' and 'ethical' narrative thesis, see Strawson (2004).

as personal identity but seems to presuppose it, since in order for us to stitch together the event of our life into a whole self we should already know which events those are. So some notion of personal identity is presupposed. Hence, in order to do justice to the narrativity thesis, narrative theorists need to respond to some facts which seem to be problematic for their account: for example, a person can remember and relive certain parts of their lives but not most of what they did, particularly those things that happened to them when they were young children. I have attempted to meet this kind of objection by taking memory to be a disposition and a capacity. However, if we adopt the narrative thesis, then we cannot avoid the following objection: there is an important difference between what one *actually* did at that a particular time and a *story* that is made up to explain those events that they do remember. I leave this kind of difficulty open in this Chapter. Instead I want to suggest that as a philosopher, we'd suggest we go for an historical narrative for a past self while trying to be creative about a future self, leaving fiction for the novelists, since I am not convinced that I have been constructed. Autobiographies are not sufficiently reliable for philosophical account of identity.[134]

[134] This Chapter was presented in the conference of *Emotion, Others, and the Self* in *Self* at Abo Akademi University, Turku, in August 2005. I am very grateful to the audience at the conference, especially, Lilii Alanen, Kathy Behrendt, and Michael Lacewing for their useful comments. I also thank Roger Squires, who suggested, in personal correspondence, some important improvements. I also extend my thanks to Louise Richardson and Francis O'Sullivan.

Final Remarks

In this project I have demonstrated that emotions do not form a natural kind distinguishable, from motive, moods, and attitudes. Emotions as mental events, for example, fear, anger, jealousy, and pride, are phenomenologically distinguishable from propositional attitudes like astonishment, and amusement. Those occurrent emotions are further distinguishable from many desires (for example, hunger and lust) and moods (for example, depression and anxiety). Yet, emotions, like anger and fear, similarly to moods such as, depression, or nostalgia pervade the psychological domain, affecting a person's perceptions and construals. Indeed, for other emotions, for example, pride, love, benevolence, aesthetic delight, and moral indignation, it is difficult to say whether they are feelings, virtues, passions, or sentiments.

In explaining emotion, I have addressed three kinds of reductionist views: the feeling-centred theory in Chapter 1, judgmentalism in Chapter 2, and a Humean functionalism in Chapter 4. Next, I have demonstrated the difficulties confronting these reductionist views. In order to avoid such redcutionism, I have suggested two alternative approaches to emotion: a qualified cognitivism and an 'embodied appraisal' theory (Chapter 3). I have demonstrated that this latter view faces difficulty in explaining moral emotion. In order to meet this challenge, I have provided an alternative which focuses on the socio-cultural nature of the emotions (Chapters, 5 and 6).

I have argued that specific emotions are individuated and identified, firstly, with reference to other mental states, notably, perceptions, beliefs, and desires (Chapters 1, 2,

3, and 4). For example, every tremble of fear, stab of jealousy, sense of indignation, ache of sadness occurs in a set of other supportive and opposed perceptions, beliefs, desires, imaginings.[1] Secondly, I have shown that emotions are individuated and identified within the context of personal narrative and are culturally determined (Chapter 4). I have shown that the reason why emotions are individuated in a person's narrative is because emotions are often changeable, unstable, and are sometimes ambivalent (Chapter 2).

The instability and ambivalence of the emotions is a feature of many literatures. As Rorty points out, "loyalty turns to ambition and envy (Iago, Macbeth): love turns to hate (Medea, Saul); fear turns to love (Jane Eyre); friendship and brotherhood turn to rivalry and sometimes to treachery (Absalom and Amnon, Pandu and Dhritarastra)."[2] The changeability of the emotions can be seen in our everyday life. I have explored this characteristic of emotion through Kate's life (Chapter 6). Kate and her husband Paul adored each other. They seemed to be a perfect couple. Her husband comes to depend on her support and admiration. She was ready to transform her life for his sake. When she became absorbed in her newborn child, he lost his bearings and was easily propelled into leaving Kate for an affair with Lucy, Kate's friend. Kate wanted to exact revenge on her ex-husband. Kate was also jealous of Lucy, and wanted to revenge on her as well. However, one day Kate's emotions changed radically when she met and fell in love with another man. Through this example, I am simply drawing attention to how one's emotional life can change over the course of a lifetime.

---

[1] For example, fear can be analysed in terms of perception and belief, jealousy in terms of belief and desire, sadness in terms of imagining. See Chapters, 1, 2, 3 and 4 of this thesis.
[2] Quoted from Rorty (2004), p. 274

However, despite their instability, emotions are sometimes long-standing (Chapter 3). Some emotions, are sustained, they increase and grow when other people reflect or echo them, like instruments of an orchestra joining in and strengthening the melody played by the soloist. The proud man sustains his pride by the recognition of his worth which he finds in others. His self-esteem is reinforced by other people, when it is mirrored by them. The proud person, hence, can sustain his pride only when others let him have it – either in terms of their approval – e.g., admiration and recognition – or in terms of their disapproval – e.g., envy. Some emotions are, on the other hand, related to morality (Chapter 5). To go back to our example in Chapter 5, let us imagine that Kate's new partner was an alcoholic and abuses his family psychologically and physically. As a result, she made a decision to put him into a clinic, even though she knew she would guilt for doing so. As has already been asked, however, can Kate's guilt be warranted? Can her partner's anger be warranted? I have argued that the appropriateness of Kate's guilt in this case neither depends on Kate's blaming herself, nor on her partner's anger in being put into the clinic, but depends on whether anger on the part of others is warranted.

I have argued that emotions also have perspective (Chapter 6). I have shown that if emotions do not have perspective, then we cannot understand how Kate, who has been in the grip of jealousy, is eventually indifferent concerning the object of jealousy. Her self-concern and self-love, which possess perspective, recreate a dimension of her past life and create her future life. Hence, she could say the following: "What a shame I live with manipulated and selfish man! I've wasted years of my life living with a man who wasn't my type." I have shown that these characteristics of emotions – stability, long-standing duration, and possession of perspective – helps us one of the most enduring

philosophical problems, that is, the problem of personal identity (Chapter 6).

I have argued that the instability and complexity of the emotions shows that they are irreducible simply to one state. Hence, I have argued that any plausible theory of emotion should depend on its integration within a comprehensive framework constituted by philosophy of mind, philosophical psychology, biology, economy, politics and socio-cultural context. In this project I have mainly focused on philosophy of mind, philosophical psychology and socio-cultural context, leaving, biology, economy and political context for a future project. A person's or a society's characteristic emotional repertoire is framed and reinforced by political and economic arrangements. As we have seen in Chapter 5, while anger itself is a normally regarded as something to be condemned, society or culture uses the emotion to encourage the development of what is regarded as an appropriate reinforcement of morally impermissible actions.

Emotions on the other hand play a significant role in the interaction between the customer and supplier. 'Market-based consumer-oriented economic systems' reinforce the range of emotion-based motives. They are designed to present images of satisfied desire. Although I have taken a close look at the interpersonal relationship, stressing the fact that all encounters are human encounters, these relationships cannot be taken out of their social context. Many institutions in a society, for example, schools, religious institutions, courts, workplaces, museums, armies, mass media, banks, etc,- create a feeling of security with reference to those things which they have already desiquoted as fearful. Schools promote respect, a regard for orderly behaviour, ambition and an ambivalent attitude to rivalry. Businesses tend to promote competitiveness and greed.

Norms of familial relations are promoted by the media and religious groups. The presence of such norms and valorised emotions means that 'emotional nonconformists' are, like Camus's Mersault, seen as odd or suspicious. Such cultural influences on emotion are elaborated on in more detail by A. Rorty. She puts the point as follows:

> Social institutions provide the models for the feelings of responsibility and accountability; they set norms for the tenor of social interactions, finely attuned for status and power, formality or intimacy, empathically tactful or aggressively confrontational. They form the patterns and the habits of aggression and cooperation that are exercised in generating and resolving ordinary conflicts."[3]

Our emotional repertoire can also be extended by political institutions. I leave the issue of the interdependence between emotion and politics for a future project.[4] Instead, in this conclusion, I want to raise an important point for those who want to do political philosophy without paying serious attention to the role of emotion in our political life. They might say that in order for a political principle to be realizable, we should not focus on 'what a community feels, but on it believes.' If we agree with them, we accept judgmentalism which we have discussed in Chapter 2. However, as we have examined, the existence of recalcitrant emotions gets judgmentalism into trouble. Now let's apply this idea to political philosophy. For example, as Susan James puts it, a man as a member of the American Republican party may fear Islamic fundamentalism and holding the unchanging belief that fundamentalists are dangerous, despite evidence to the contrary. He may avoid confronting the possibility that his fear is ungrounded.[5] In this project, I have only focused on recalcitrant emotions which are experienced by individuals as opposed to groups. However, as Susan James nicely points out, "political

[3] Rorty, Ibid., p. 277
[4] An interesting account of the relationship between emotion and politics, see James, S. (2003), pp. 221-234.
[5] James, S., Ibid., p. 228

philosophers may feel able to ignore the recalcitrant emotions of individuals, perhaps on the grounds that they will cancel each other out, but the recalcitrant emotions of groups are harder to dismiss."[6] For example, as a group, American Republicans may feel fear and anxiety regarding Islamic fundamentalists, which has been cultivated by their shared experiences and their own understanding of history.[7] Cases like this show that the emotions of groups are often not responsive to the rational arguments of political philosophers. I leave further investigation of this issue for a future project.

Finally, I want to propose another direction for my future research with regard to emotion and personal identity. In addressing the problem of personal identity with regard to emotion, I have remained non-committal on whether we conceive of time as liner or non-linear. The clarification of this issue in a future project will contribute to clarifying the relationship of selfhood or consciousness to time and memory as manifested in narrative. It will also illuminate the relationship between memory, temporal sequence and narrative structure.

[6] Ibid.
[7] Ibid.

BIBLIOGRAPHY

Ainslie, D. (1999), 'Scepticism about the Persons in Book II of Hume's *Treatise'*, *Journal of the History of Philosophy,* 37, 469-78.

Ainslie, D. (2001), 'Hume's Reflections on the Identity and Simplicity of Mind', *Philosophy and Phenomenological Research*, 62, 557-78.

Alanen, L. (2003), 'What Are Emotions About?', *Philosophy and Phenomenological Research*, 67, 311-334.

Alston, W. P. (1980), 'Emotion and Feeling', in P. Edward (ed.), *The Encyclopaedia of Philosophy*, 2, New York: Macmillan.

Anscombe, G. E. M (1957), *Intention,* Oxford: Blackwell.

Annas, J. (1992), *Hellenistic Philosophy of Mind,* University of California Press.

Aristotle, (1928), *On the Soul*, in W. D. Ross (tans. & ed.), *The work of Aristotle* (12 vols.), Oxford: Oxford University Press.

Aristotle, (1985), *Nicomachean Ethics,* (trans.), T. Irwin. Indianapolis, IN: Hackett.

Ayer, A. J. (1952), *Language, Truth and Logic,* New York: Dover.

Baier, A. (1980), 'Master Passions', in A. Rorty (ed.), *Explaining Emotions,* University of California Press.

Baier, A. (1986), 'The Ambiguous Limits of Desire', in J. Marks (ed.), *The Ways of Desire.* Precedent Publishing.

Baier, A. (1990), 'What Emotions are About', *Philosophical Perspectives*, 4: *Action Theory and Philosophy of Mind.*

Baier, A. (1991), *A Progress of Sentiments: Reflections on Hume's Treatise*, Harvard University Press.

Byers, P.(2005), *Personal Identity: Temporality and Continuity* (a PhD thesis), University of Sydney

Ben-Ze'ev, A. (1990), 'Envy and Jealousy', *Canadian Journal of Philosophy*, 20

Ben-Ze'ev, A. (2004), 'Emotions are not Mere Judgments', *Philosophy and Phenomenlogical Research*, 450- 457.

Bergman, F. (1978), 'Review of Solomon's the passion', *Journal of Philosophy*, 75.

Biro, J. (1976), 'Self- Identity and Memory', *The Review of Metaphysics*, 26

Blackburn, S. (1981*a*), 'Reply: Rule Following and Moral Realism', in S. Holtzman and S. Leich (ed.), *Wittgenstein: To Follow a Rule*, London: Routledge & Kegan Paul.

Blackburn, S. (1981*b*), 'Non-Cognitivism and Rule-Following,' in S. Holtzman and S. Leich (ed.), *Wittgenstein: To Follow a Rule*, London: Routledge & Kegan Paul.

Blackburn, S. (1984), *Spreading the Word: Grounding in the Philosophy of Language*, Oxford: Clarendon Press.

Blackburn, S. (1993), *Essays in Quasi-Realism*, New York: Oxford University Press.

Blackburn, S. (1998), *Ruling Passions*, New York: Oxford University Press.

Blum, L. (1980), *Friendship, Altruism, and Morality*, London: Routledge & Kegan Paul.

Blum, L. (1987), 'Particularity and Responsiveness', in J. Kagan and S. Lamb (ed.),*The Emergences of Morality in Young Children*, Chicago: University of Chicago Press.

Brett, N. (1972), 'Substance and Mental Identity in Hume's *Treatise,' The Philosophical Quarterly,* 22, *87.*

Brewer, W. F. (1988), 'Memory for randomly sampled autobiographical events', in, U. Neisser, & E. Winograd (ed.), *Remembering Reconsidered: Ecological and Traditional Approaches to the Study of Memory*, Cambridge University Press.

Bricke, J. (1996), *Mind and Morality: An Explanation of Hume's Moral Psychology*, Oxford Clarendon Press.

Brook, D. H. M. (1994), *The Unity of the Mind*, St. Martin's Press

Butler, J. (1975), 'Of the Personal Identity', in John Perry, (ed.), *Personal Identity*

Calhoun, C. (1984), 'Cognitive Emotions?', in Calhoun and R. Solomon (ed.), *What is Emotions?,* New York: Oxford Press.

Campbell, J. (1994), *Past, Space, and Self*, Cambridge, MA: MIT Press.

Damasio, A. R. (2000), *The Feelings of What Happens: Body, Emotion in the Making of Consciousness*, New York: Harcourt Brace.

Dancy J. (1986), 'Two Concepts of Moral Realism', in *The Proceedings of the Aristotelian Society,* supp., 60, 167-87.

Danto, A. C. (1999), 'Basic Actions and Basic Concepts', in *The Body/Body Problem,* Berkeley: University of California.

D'Arms, J. and Jacobson, D. (1994), 'Expressivism, Morality, and the Emotions', *Ethics,* 104, 739-763.

D'Arms, J. and Jacobson, D. (2000*a*), 'The Moralistic Fallacy on the "Appropriateness" of Emotion', *Philosophy and Phenomenological Research,* 61, 65-90.

D'Arms, J. and Jacobson, D. (2000*b*),'Sentiment and Value,' *Ethics, 110, 722-48.*

D'Arms, J. and Jacobson, D. (2003), 'The Significance of Recalcitrant Emotion', in Anthony Hatzimoysis (ed.), *The Philosophy and the Emotion,* Cambridge University Press.

Darwin, (1889/1998), *The Expression of the Emotions in Man and Animals,* with an introduction, afterward and commentary by P. Ekman, London: HarperCollins,

Dennett, D. (1983), *Elbow Room: The Varieties of Free Will Worth Wanting,* Cambridge, MA: MIT Press

Deigh, J. (1994), 'Cognitivism in the Theory of Emotions', *Ethics, 104, 824-54.*

Deigh, J. (2004), 'Primitive Emotions', in Solomon (ed.), *Thinking about Feeling,* Oxford & New York: Oxford University Press.

Deonna, J. A. (2006), 'Emotion, Perception and Perspective', *Dialectica,* 60, 29-46.

Descartes, R. (1984,1991), *The Philosophical Writings Of Descartes,* Vols. I. and II, (ed.), J. Cottingham, R. Stoothoff, and D. Murdoch, and A. Kenny, Cambridge; Vol. III Correspondence, (ed.), J. Cottingham, F. Stoothoff, D. Murdoch, and A. Kenny, Cambridge

Descartes, R. (1984), *Meditations on First Philosophy,* in Cottingham (ed.), *Philosophical Writings, et al.,* II.

Descartes, R. (1984), *The Passions of the Soul,* in Cottingham (ed.), *Philosophical Writings, et al.,* I.

De Sousa, R. (1974), 'The Good and the True', *Mind,* 83.

De Sousa, R. (1987) *The Rationality of Emotion,* Cambridge, Mass., London: MIT Press, 543-51.

De Sousa, R. (2002), 'Emotional Truth', *Proceedings of the Aristotelian Society,* Supplementary volume 76, 247-64.

De Sousa, R. (2004), 'Emotion – what I know, what I'd like to think I know, and what I'd like to think,' in Solomon (ed.), *Thinking about Feeling,* Oxford & New York: Oxford University Press.

Doering, S. (2003), 'Explaining Action by Emotion', *The Philosophical Quarterly,* 53, 214-230.

Doering, S. (2004), *Gründe and Gefühle: Rationale Motivation dürch emotionale Vernuft,* Universitat Essen.

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

Dunn, R. (2004), 'Moral Psychology and Expressivism', *European Journal of Philosophy,* 12.

Ekman, P. (1972), 'Universals and Cultural Differences in Facial Expressions of Emotion', in J. Cole (ed.), *Nebraka Symposium on Motivation,* 1971 (Vol. 19, pp. 207-283), Lincoln; University of Nebraska Press.

Ekman, P. (1980), 'Biological and Cultural Contributions to Body and Facial Movement in the Expression of Emotion', in A. Rorty (ed.), *Explaining Emotions,* Berkeley: University of California Press.

Ekman, P. (1992), 'Are there basic emotions?: A reply to Ortony and Turner', *Psychological Review,* 99, 550-3.

Ekman, P. (1999), 'Basic Emotions', in T. Dalglesish & T. Power (ed.), *The handbook of cognition and emotion,* 45-60, New York: Wiley.

Evans, G. (1982), *The Varieties of Reference*, Oxford: Clarendon Press.

Gibbard, A. (1990), *Wise Choices, Apt Feelings: A Theory of Normative Judgment,* Cambridge, Mass.: Harvard University Press.

Glathe, Alfred B. (1950), *Hume's Theory of the Passions and of Morals: A Study of Books II and III of the "Treatise",* University of California Press.

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

Goldie, P. (2002), 'Emotions, Feelings and Intentionality', *Phenomenology and the Cognitive Sciences,*1, 235-54.

Goldie, P. (2004), 'Emotion, Feeling and Knowledge of the World', in Solomon (ed.), *Thinking about Feeling,* Oxford & New York: Oxford University Press.

Grahma, G. (1990), 'Melancholic Epistemology', *Synthesis*, 82.

Greenspan, P. (1980), 'A Mixed Case of Feeling', in A. Rorty (ed.), *Explaining Emotions.*

Greenspan, P. (2003), 'Emotions, Rationality and Mind/Body,' in Anthony Hatzimoysis (ed.), *Philosophy and the Emotion*, Cambridge University Press, 113-125

Hacker, P. M. S. (2002), 'Is there Anything it is Like to be Bat?,' *Philosophy*, 77.

Hatzimoysis, A. (2003) (ed.), *Philosophy and the Emotion,* Cambridge University Press.

Hamilton, A. (1995), 'A New Look at Personal Identity', *Philosophical Quarterly,* 45.

Hamilton, A. (forthcoming), *Memory and Body: A Study of Self-Consciousness*

Hobbes, T. (1991), *Leviathan,* (ed.), R. Tuck, Cambridge.

Humberston, L. L. (1992), 'Direction of fit', *Mind*, 101

Hume, (1978), *A Treatise of Human Nature*, (ed.), L. A. Selby-Bigge, 2nd ed. Revised by P. H. Nidditch, Oxford: Clarendon Press.

Hume, (1777), *Enquiry Concerning the Principles of Morals*, 2nd ed., (ed.), L. A. Selby-Bigge Oxford: Clarendon.

Izard, C. E. (1971), *The Face of Emotion*, New York: Appleton-Century-Crofts.

IL-Whan, Im (1996), 'Understanding Emotion and Feeling', in Dae-Heoyn Jeong (ed.), *A Philosophy of Emotion*, Min Um Press, Seoul, Korea.

James, S. (1997), *Passion and Action: The Emotions in Seventeenth-Century Philosophy*, Clarendon, Press, Oxford.

James, S. (1998), 'Explaining the Passions', in Stephen Gaukroger (ed.), *The Soft Underbelly of Reason: The Passions in the Seventeenth Century*, Routledge.

James, S. (2003), 'Passion and Politics', in Hatsimoysis (ed.), *Philosophy and the Emotions,* Cambridge University Press

James, W. (1884), 'What is an Emotion?,' *Mind* 9, 188-205.

James, W. (1890/1950), *The Principles of Psychology,* New York.

James, W. (1897/1956), *The Will to Believe and Other Essays in Popular Philosophy*, New York: Dover Publications.

James, W. (1902), *The Varieties of Religious Experience: A Study in Human Nature*, New York, Bombay: Longmans, Green and Co.

James, W. (1907/1970), *Pragmatism*, Indianapolis, Cambridge, Mass: Hackett.

Johnston, M. (1997), 'Dispositional Theories of Value', *Proceedings of the Aristotelian Society* Supplementary Volume.

Kenny, A. (1963), *Action, Emotion and Will,* Routledge & Kegan Paul.

Lazarus, R. (1984), 'On the Primacy of Cognition', *American Psychologist*, 39, 124-29.

Levenson, R. W., Ekman, P., & Friesen W. V. (1990), 'Voluntary facial action generates emotion-specific autonomic nervous system activity', *Psychophysiology*, 27, 368-84.

Locke, J. (1975), *Essay Concerning Human Understanding,* (ed.), Peter H. Nidditch Oxford.

Lowe, E. J. (1991), 'Substance and Selfhood', *Philosophy*, 66, 81-99.

Lowe, E. J. (1995), *Locke: on Human Understanding,* Routledge and New York.

Lowe, E. J. (2000), 'Personal Identity and Self-Knowledge', in *An Introduction to the Philosophy of Mind,* Cambridge University Press.

Lyons, W. (1980), *Emotions*, Cambridge University Press.

MacIntyre, A. (1981), *After Virtue*, London: Duckworth.

McIntyre, J. L. (1989), 'Personal Identity and the Passions', *Journal of the History of Philosophy,* 27, 545-57.

McGinn, C. (1982), *The Character of Mind,* Oxford University Press.

McGinn, C. (1983), *The Subjective View,* Oxford.

McCanley, R. N. (1988), 'Walking in our own footsteps: Autobiographical memory and reconstruction', in *Remembering Reconsidered*, Cambridge University Press.

McDowell J. (1981) 'Non-Cognitivism and Rule-Following' in Holtzman & Leich (ed.), *Wittgenstein: to Follow a Rule,* London: Routledge.

McDowell, J. (1985), 'Values and Secondary Qualities', in T. Honderich (ed.), *Morality and Objectivity: A Tribute to J.L. Mackie*, London: Routledge.

McDowell, J. (1998), *Mind, Value and Reality*, London, Cambridge, HA: Harvard University Press.

McNaughton, D. (1988), *Moral Vision*, Basil Blackwell.

Morton, A. (2002). 'Emotional Accuracy', *Proceedings of the Aristotelian Society,* Suppl, 76, 265-75.

Nagel, T. (1970), *The Possibility of Altruism,* Princeton University Press.

Nash, N. (1989), 'Cognitive Theories of Emotions', *Nous*, 23.

Noonan, H. W. (1989), *Personal Identity*, Routledge.

Nussbaum, M. C. (1993), 'Poetry and Passions: Two Stoic Views' in J. Brunschwing and M. Nussbaum (ed.), *Passions and Perceptions*, Cambridge University Press.

Nussbaum, M. C. (1990), *Love's Knowledge: Essays on Philosophy and Literature,* New York: Oxford University Press.

Nussbaum, M. C. (1993), 'Passions and Perceptions', *Hellenistic Philosophy of Mind*, Cambridge University Press.

Nussbaum, M. C. (1994), *The Therapy of Desire: Theory and Practice in Hellenistic Ethics*, Princeton: Princeton University Press.

Nussbaum, M. C. (2001), *Upheavals of Thought,* Cambridge University Press.

Ortony, A. and Turner, W. (1990), 'What's basic about basic emotions?', *Psychological Review*, 97, 315-331.

Parfit, D. (1987), *Reason and Persons*, Oxford University Press.

Penelhum, T. (2000), *Themes in Hume: The Self, The Will, Religion*, Clarendon Press, Oxford.

Perry, J. (1972), 'Can the Self Divides?', *Journal of Philosophy*, 73, 463-88.

Perry, J. (1975), *Personal Identity*, California University Press

Plato, (1961), *The Collected Dialogues*, (ed.) Edith Hamilton and Huntington Cairns. Princeton University Press.

Platts, M. (1979), *Ways of Meaning*, London: Routledge and Kegan Paul.

Peters, R. (1967), *Hobbes*, Penguin Books.

Prinz, J. (2003), 'Emotion, Psychosemantics, and Embodied Appraisals', in A. Hatzimoysis (ed.), *Emotion and Philosophy*, Cambridge University Press.

Prinz, J. (2004a), 'Embodied Emotions', in *Thinking about Feeling*, Oxford University Press.

Prinz, J. (2004b), *Gut Reactions: A Perceptual Theory of Emotion*, Oxford University Press.

Proust, M. (1981), *Remembrance of Things Past*, Vol III. Scott Moncrieff, C. K., Mayor, A. (trans) & Kilmartin, T. (rev), London: Vintage.

Rawls, J. (1977), *A Theory of Justice*, Cambridge, Mass: Harvard University Press.

Ratcliffe, M. (2005*a*), 'William James on Emotion and Intentionality,' *International Journal of Philosophical Studies*, 13. 179-202.

Ratcliffe, M. (2005*b*), 'The Feeling of Being', *Journal of Consciousness Studies,* 12, 43-60.

Reid, T. (1975), 'Of Mr. Locke's Account of Our Personal Identity', in John Perry (ed.), *Personal Identity.*

Robert, D. (2004), 'Moral Psychology and Expressivism', *European Journal of Philosophy,* 12.

Roberts, R. (1988), 'What an Emotion Is: A Sketch', *Philosophical Review.*

Roberts, R. (1991), 'What is Wrong with Wicked Feeling?', *American Philosophical Quarterly,* 28, 13-24.

Roberts, R. (2002), *Emotions: An Essay in Aid of Moral Psychology*, Cambridge University Press.

Robinson, J. (1983), 'Emotion, Judgment and Desire,' *Journal of Philosophy*, 80, 731-41.

Rorty, A. (1976), *The Identity of Persons*, Berkeley: University of California Press.

Rorty, A. (1980), *Explaining Emotions*, (ed.), A. Rorty, University of California Press.

Rorty, A. (2004), 'Enough Already with "Theories of the Emotions"', in R. Solomon (ed.), *Thinking about Feeling,* Oxford University Press.

Rudd, A. (2005), 'Narrative, Expression and Mental Substance', *Inquiry*, 48.

Salmela, M. (2005), 'Emotional Feelings as Twofold Representations', in Å. Carlson (ed.), *Philosophical Aspects of Emotions.* Stockholm: Thales.

Salmela, M. (2006), 'True Emotions', *The Philosophical Quarterly,* 56, 382-405

Schaffer, J. (1983), 'An Assessment of Emotion', *American Philosophical Quarterly*, 20, 161-73.

Schechtman, M. (1994), 'The same and the same: Two views of Psychological continuity' *American Philosophical Quarterly*, 31. 199-212.

Schechtman, M. (1997), *The Constitution of Selves*, Itach: Cornell University Press.

Schiffer, S. (1979), 'A Paradox of Desire', *American Philosophical Quarterly*, 13.

Schueler, G. F. (1991), 'Pro-Attitudes and Direction of Fit', *Mind*, 100.

Schueler, G. F. (1995), *Desire,* The MIT, Cambridge, Massachusetts.

Searle, J. (1983), *Intentionality: an Essay in the Philosophy of Mind*, Cambridge University Press.

Searle, J. (2001), *Rationality in Action,* A Bradford Book, The MIT Press, Cambridge, Massachusetts, London, England.

Shoemaker, S. (1963), *Self-Knowledge and Self-Identity,* Ithaca, NY: Cornell University Press.

Shoemaker, S. (1984), 'Personal Identity; A Materialist Account', in Sydney Shoemaker and Richard Swinburne (ed.), *Personal Identity*, Oxford: Basil Blackwell.

Smith, (1987), 'The Humean Theory of Motivation', *Mind,* 96.

Smith, M. (1994), *The Moral Problem,* Oxford: Basil Blackwell.

Sobel, D. and Copp, D. (2001), 'Against direction of fit accounts of belief and desire', *Analysis,* 61.

Sorabji, R. (1993), *Animal Mind and Human Morals*, Duckworth.

Sorabji, R. (2001), *Emotion and Peace of Mind: Stoic Agitation to Christian Temptation,* Oxford University Press.

Solomon, R. C. (1976), *The Passions,* New York: Doubleday.

Solomon, R. C. (1980), 'Emotion and Choices', in A. Rorty (ed.), *Explaining Emotions,* Berkeley: University of California Press.

Solomon, R. C. (1993), *The Passions: Emotions and the Meaning of Life,* 2nd ed. Indianapolis, IN: Hackett.

Solomon , R. C. (2004), 'Emotions, Thoughts, and Feelings: Emotions as Engagements with the World', in R. Solomon (ed.), *Thinking about Feeling,* Oxford University Press.

Spada, M. (1998), *Space and the Self in Hume's Treatise,* Cambridge University Press.

Stampe, D. (1987), 'The Authority of Desire', *The Philosophical Review,* 96.

Stroud, B. (1977), *Hume,* Routledge & Kegan Paul.

Stroud B. (1993) '"Gilding or Staining" the world with "Sentiments" and "Phantasms"', *Hume Studies,* Vol. XIX, No.2.

Strawson, G. (2004), 'Against Narrativity', *Ratio,* 17, 428-452.

Tappolet, C. (2000), *Emotions et Valeurs.* Paris: Presses Universitaires de France.

Tappolet, C. (2005), 'Ambivalent Emotions and the Perceptual Account' of Emotions', *Analysis,* 65. 229-33.

Taylor, C. C. W. (1986), 'Emotions and Wants', in J. Marks (ed.),*The Ways of Desire,* Chicago: Precedent Publishing.

Taylor, G. (1975), 'Justifying the Emotions', *Mind,* 84.

Taylor, G. (1985), *Pride, Shame, and Guilt: Emotions of Self-Assessment*, Oxford: Oxford University Press.

Taylor, G. (1988), 'Jealousy and Envy: Emotions and Vices', *Midwest Studies in Philosophy, 13.*

Velleman, D. (1992), 'The Guise of the Good', *Nous,* 26, 3-26.

Walton, K. (1979), 'Fearing Fictions', *The Journal of Philosophy*, 75.

Wiggins, D. (1967), *Identity and Spatio-Temporal Continuity,* Oxford: Blackwell

Wiggins, D. (1987), 'A sensible Subjectivism', in *Needs, Values, Truth,* Basil Blackwell.

Williams, B. (1973), *Problem of the Self,* Cambridge University Press.

Wollheim, R. (1979), 'Memory, Experiential Memory, and Personal Identity', in G. F. McDonald (ed.), *Perception and Identity: Essays Presented to A. J. Ayer with His Replies to Them*, London: Macmillan.

Wollheim, R. (1980), 'On Persons and Their Lives', in A. Rorty, *Explaining Emotion.*

Wollheim, R. (1984), *The Thread of Life*, Cambridge University Press.

Wright C. (1988), 'Moral Values, Projection and Secondary Qualities', *The Proceedings of the Aristotelian Society.*

Zangwill, N. (1998), 'Direction of Fit and Normative Functionalism', *Philosophical Studies*, 91.