The Role of Emotion in Rational Decision-Making.

Alison Kathryn Stratton
Adelaide University
School of Humanities
Discipline of Philosophy
July 2012
# Table of Contents

Chapter One: The Pro-Emotion Consensus .......................................................................................... 7

1. The Roles of Emotions in Practical Rationality .............................................................................. 9

2. Pro-Emotion Theories ................................................................................................................. 20

2.1 Evolutionary Psychology Theories ................................................................................................ 20

2.2 Strategic Theories ....................................................................................................................... 24

2.3 Perceptual Theories ................................................................................................................... 27

2.4 Neo-Jamesian Theories ............................................................................................................. 28

Chapter Two: Historical Theories of Emotion and Emotions as the Cause of Irrational Behaviour .......................................................................................................................... 35

1. Historical Theories of the Emotions ............................................................................................... 37

1.1 Theories of Emotion in Ancient Greece ...................................................................................... 38

1.1.1 Plato ......................................................................................................................................... 38

1.1.2 Aristotle ................................................................................................................................... 40

1.1.3 The Stoics ................................................................................................................................ 42

1.2 Theories of Emotions in the 17th & 18th Century ........................................................................ 45

1.2.1 Descartes ................................................................................................................................. 46

1.2.2 Hume ...................................................................................................................................... 49

1.2.3 Kant ....................................................................................................................................... 51

2. Historical Theories and the Lack of ‘Anti-Emotion Consensus’ ...................................................... 53

3. Case Studies .................................................................................................................................. 62

3.1 Recycled Wastewater .................................................................................................................. 62

3.2 Emotional Eating ........................................................................................................................ 64

Chapter Three: Strategic & Evolutionary Psychology Theories Revisited ....................................... 66

1. Strategic Theories ......................................................................................................................... 68
1.1 The insufficiency of explanations for emotionally driven irrational behaviour .......................................................... 71
1.2 Is the strategic theory a pro-emotion view? .......................................................... 75
1.3 Emotional Training and Regulation on the Strategic Account .................................. 79

2. Evolutionary Psychology .......................................................................................... 81
   2.1 Evolutionary Psychology's explanation of emotionally-driven irrational behaviour .................................................. 82
   2.2 Why don't these explanations work? ..................................................................... 85
   2.3 Accounting for Regulation and Training of Emotions ........................................... 90

3. Prinz's Embodied Appraisal Theory ....................................................................... 93

Chapter Four: The Success of the Perceptual Accounts .............................................. 100
1. Perceptual theories revisited ...................................................................................... 101
   2.1 Damasio's Account ............................................................................................. 101
   2.2 De Sousa's Account ........................................................................................... 102
2. Explaining emotionally-driven irrational behaviour ............................................... 105
3. Application to the Case Studies ................................................................................. 108

Chapter Five: Dealing with Emotionally-Driven Irrational Behaviour ....................... 116
2. Breaking the Connection Between Emotion and Behaviour ................................... 122
3. Managing Occurrent Emotions ............................................................................... 127
4. Changing dispositions to feel emotions ................................................................. 132
5. Emotion regulation in public policy ........................................................................ 135

Bibliography ..................................................................................................................... 146
Abstract

Within philosophy of emotion, there has been the development of a pro-emotion consensus that claims emotion plays a positive and important role within decision-making. It has been suggested that the pro-emotion consensus is replacing ‘traditional’ views of emotion which claim emotion is disruptive to rationality and is often the cause of irrational behaviour; claims which appear to be supported by empirical evidence.

In this thesis I will be examining several pro-emotion theories and historical theories of emotion. I will argue that the suggestion that all historical theories of emotion are ‘anti-emotion’ theories is incorrect and fails to take into account the many differences between historical accounts. In fact, claims made by some of these historical theories are not unlike those made by some pro-emotion accounts; the view that emotion is disruptive to rationality is limited to a small number of historical theories.

Although all pro-emotion theories agree that emotion is necessary for decision-making, there are differences between accounts regarding the exact function emotion serves. I will argue that perceptual pro-emotion theories provide the best explanation of the role of emotion and of emotionally-driven irrational behaviour. I aim to show that the claims made by these perceptual theories and (some) historical theories are compatible and when we combine the insights of these accounts we have a comprehensive explanation of the role of emotions within rational decision-making, which can account for emotionally driven irrational behaviour and still maintain that emotions are necessary for decision-making.
Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library catalogue and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Signed:
Acknowledgements

This thesis would not have been possible without the support of a number of people. First and foremost, I would like to extend my deepest gratitude to my principal supervisor, Dr. Jennie Louise. Throughout the entirety of this thesis your enthusiasm, positivity and support has been unwavering. I’ve learned so much from you – and not just about philosophy or writing a thesis. Thank you so much!

Many thanks to my co-supervisor, Prof. Garrett Cullity, whose feedback and objectivity were immensely helpful in the editing and fine-tuning processes of this thesis. I am most appreciative of your insights and assistance.

Lastly, but not least, thank you to all my family and friends for the continual love and support. Special mention to my parents, Lewis and Kathryn, brother, Daniel and my new husband, Kyle; the four of you kept me going and encouraged me to succeed when I was ready to tear my hair out. Thank you!
Chapter One: The Pro-Emotion Consensus.

Introduction

In the latter half of the twentieth century, one of the major developments in theories of emotions has been the emergence of a ‘pro-emotion consensus’ (Jones, 2006) concerning the relationship between emotions and practical rationality. This consensus holds that emotions are useful, and indeed necessary, for rational decision-making; supposedly this view overturns previous assumptions that emotions are merely disruptive influences on rational deliberation and decision. As Jones (2006: 3) describes it, “within both philosophy and psychology, a new pro-emotion consensus is replacing the old dogmas that emotions disrupt practical rationality, that they are at best arational, if not outright irrational.” In this thesis I will explore the claims of the pro-emotion consensus concerning the contribution that emotions make to rational decision. In particular, I will be asking how pro-emotion theories reconcile their view of the necessity of emotions for practical rationality with a very commonsense (and, I will argue, accurate) idea emphasised by historical theories of emotion: namely, that emotions also often cause irrational behaviour, and require regulation by reason.

The focus of the thesis will therefore be on the issue of the function or role of the emotions. The members of the pro-emotion consensus will obviously have views on other issues in the philosophy of emotion – for example, whether emotions have propositional content, whether they form a natural kind, or whether there are basic emotions. However, I will not be concerned with these issues: the question of the function of emotions in practical rationality can be explored without having to address them, and it allows me to categorise members
of the pro-emotion consensus simply according to what roles they assign to emotions.

In this chapter I will describe the pro-emotion consensus and the positive functions emotions are supposed to play in rational decision-making. I will then discuss the various differences of detail that exist between members of the consensus, and categorise pro-emotion theories into three groups: strategic theories, Evolutionary Psychology theories, and perceptual theories. In Chapter Two I will examine the claim that the pro-emotion view overturns a longstanding traditional consensus according to which emotions are antithetical to practical reason. I will argue, using scholarship on ancient and early modern theories of emotion, that this overstates the case. In fact, traditional theories often held emotions to play much the same roles in practical reason that pro-emotion theorists do. Additionally, their focus on the tendency of emotions to cause irrational behaviour, and the need for emotions to be trained and regulated by reason, is something that pro-emotion theorists should accommodate within their theories.

Therefore, in Chapters three and four I consider how the different kinds of pro-emotion theory can explain emotionally-driven irrational behaviour, and what resources they offer for emotional training and regulation: in Chapter Three I argue that strategic and Evolutionary Psychology theories fail to give a consistent and satisfactory account of irrational behaviour and emotional regulation, and in Chapter Four I argue that perceptual theories can give an adequate account. I use two case studies of modern emotionally-driven irrational behaviour for illustration: opposition to wastewater recycling programmes, and ‘emotional eating’. In Chapter Five I bring the conclusions of the previous chapters together to show how ‘perceptual’ pro-emotion theories give the most satisfactory overall
picture of the role of emotions in practical rationality, and I show how the theory can be used to give practical recommendations for dealing with emotionally-driven irrational behaviour both on an individual and public policy level.

1. The Roles of Emotions in Practical Rationality

The members of the ‘pro-emotion consensus’ have their differences, which I will detail in later sections. However, what creates a consensus is the functional view of emotions: that emotions play particular roles in practical decision-making which are helpful and indeed necessary. Because of this, they are also broadly evolutionary views: they all believe that we have emotions because of their useful functions, which are adaptations to the environment faced by our ancestors. Emotions do certain things which are (or at least were) adaptive for the organisms that had them, and which helped those organisms to live and reproduce more successfully than the available alternatives. As Robert Levenson (1994:123) writes, "Emotions are short-lived psychological-physiological phenomena that represent efficient modes of adaptation to changing environmental demands." Or as Karen Jones puts it (2006:3), “emotions are clever design solutions to the problem of making fast decisions in response to significant practical problems posed by the natural and social worlds.”

We can also give a general characterisation of emotions that all members of the consensus (and probably most philosophers and psychologists studying emotion) would agree to. The following are generally accepted as being defining characteristics of emotions (Faucher & Tappolet 2008: 106-108):

1. Emotions involve some sort of perception, comprehension or ‘grasp’ of an object, even if this is not conceptual.
2. Emotions involve some sort of appraisal or evaluation, which may again be non-conceptual.

3. Emotions involve physiological changes and often characteristic facial expressions.

4. Emotions involve a ‘characteristic feeling’.

5. Emotions involve action tendencies; they ‘push’ us towards certain actions or ‘pull’ us away from others.

According to the pro-emotion consensus, these aspects of emotion have evolved because they are useful to us in deciding what to do. But there are various different ways in which states with the characteristics of emotions might be supposed to help us in rational deliberation, decision, and action. Jones (2006:5-7) distinguishes four general kinds of helpful role that emotions might play in practical rationality:

1. Improving our access to reasons.

2. Facilitating planning agency.

3. Supplementing decision-theoretic reasoning and remedying its deficiencies.

4. Facilitating appropriate action.

I will discuss these in turn, and describe how emotions are thought to play these roles.

1. Improving Access to Reasons
According to Jones, emotions improve access to reasons by leading us to “interpret considerations as reasons where we otherwise might not have” (Jones; 2006:5). In other words, emotions can direct our attention to relevant features of the environment. They carry information about how a situation has been evaluated; they signal events and features that are relevant to our concerns and interests. As Frijda (1994: 113) says, emotions “represent a process of relevance signalling…. Emotions signal the relevance of events to concerns.” In addition to calling our attention to events and features of the environment, emotions can make these immediately salient to us; they will affect not just attention but also interpretation. As Clore (1994: 105) writes,

a primary function of emotion is to provide information about how a situation has been appraised…It serves as data for judgment and decision making processes and also for reordering processing priorities…Emotions guide one’s attention to things that are relevant to goals and concerns that are implicated in the emotional situation. Such processes ensure what appears most important is attended first (1994:105).

This function is useful in practical deliberation because it allows us to very quickly identify those aspects of a situation or event which are important, and why. This means that we will be less likely to overlook important features of our situation, and also that we will not have to spend as much time finding them.

2. Facilitating Planning Agency

The second way in which emotions might contribute positively to practical rationality is by playing a role in organising our goals and sub-goals into planning hierarchies: that is, in helping us to decide which of our concerns is more
important when several are implicated in a choice situation, and also in deciding to pursue long-term goals, which requires us to sacrifice present goods for greater future ones. Although Jones does not provide an explanation of this category, she seems to have in mind the function for emotions proposed by Damasio, who has written extensively about impaired planning agency in people with compromised ability to experience emotion. Damasio discusses patients who have suffered brain trauma to the pre-frontal cortices, which have made them unable to experience normal emotions. These patients do not seem to have impaired intelligence or theoretical reasoning, but they are profoundly affected in their ability to plan and make good decisions.

Damasio’s most striking and famous example is that of Phineas Gage, the railroad worker who sustained a major trauma to his pre-frontal cortex when an iron rod penetrated his skull. While Gage recovered, his personality and temperament underwent fundamental changes. Prior to the accident he had been reliable, polite and efficient (Harlow, 1868), but afterwards he became impulsive, impatient and prone to bad language and fits of temper (Damasio, 2006: 8). He could no longer hold down a job or normal relationships. Damasio observes, from Gage and other patients with similar injuries, that these patients appear to be normal in their general intelligence and ability to reason, but that they are quite abnormal in their lack of affect. Damasio argues that the loss of capacity to feel emotions has caused an inability to order goals and to form and act in accordance with longer-term plans. Therefore, one of the functions of emotions is to enable us to do these things (Damasio, 1994: 192-3).

We can see that this function – facilitating planning agency – is in fact closely related to the signalling function discussed above; in fact, it can be seen as another type of signalling. Emotions can signal relevant aspects of our current
situation or environment, but they can also signal relevant aspects of our choices and of imagined future events. The way that emotions help us to form and act on plans is by attaching a feeling, and motivational force, to possible future events, thus allowing us to pay sufficient attention to them when making decisions. People who cannot experience emotions are not able to pick out those things that are most important about their choices, and cannot properly represent the value of future events in order to pursue those rather than present desires.

We can see how the ability to represent future goods can assist in planning agency by considering the work of Ainslie on the phenomenon of hyperbolic discounting and its relation to weakness of will. Ainslie’s interest is in the fact that the normal way of discounting future goods in choice situations leads to irrational behaviour, in particular time-inconsistent preferences that lead us to prefer smaller goods now over larger goods later, and hence undermine our longer-term plans. According to Ainslie (2001), we require strategies and methods to overcome this tendency, and emotions would be one way of doing this if they play the signalling role suggested by the pro-emotion consensus.

It is generally considered rational to have some rate of discounting for future goods: the associated uncertainty means that I should value $100 in one year’s time as being worth less than $100 now. Decision theorists think that an ‘exponential’ discount curve is appropriate: that is, that the future goods are discounted according to an exponential function. However, Ainslie’s own research has led him to conclude that most people discount future goods according to a ‘hyperbolic’ function. This function is “more bowed” (Ainslie, 2001: 31) than the exponential function, and its telltale characteristic is that it leads to time-inconsistent preferences. Suppose that I am offered a choice between (a) $50 in 6 months or (b) $100 in 12 months, and I prefer to take $100 in 12 months. If I am
time consistent in my preferences (as I would be if my discount function were exponential), I should always prefer $100 at time $t+6$ months to $50$ at time $t$, no matter when time $t$ is. However, if I have a hyperbolic discount function, then I will reverse my preferences as time $t$ gets closer to the present. Therefore, although I preferred $100$ in 12 months to $50$ in 6 months, I will tend to prefer $50$ now to $100$ in 6 months. This causes a problem for my ability to make long-term plans and goals, because I will often be undermining them due to short-term preference reversal.

According to Ainslie, there are several ways in which we can learn to manage ourselves so as to keep ourselves committed to our long-range plans and avoid undermining by short-term preference reversal. Some of these, such as making the undermining option physically impossible, or manipulating attention, he thinks are ‘unsophisticated’ and often impractical (Ainslie, 2001:74–76). And although his apparent favoured technique is to make personal rules by ‘stacking’ a series of future choices together (which increases their rewarding-ness), he believes that emotions can be more successfully used to help long-range plans ‘win’ over short-term ones (2001:79–85). And we can see that, if emotions can help us to represent the value of future goods in our present choice situation (by playing a signalling role), then this would help us to remain consistent in our preferences over time, and hence achieve our long-term plans.

3. Supplementing Decision-Theoretic Reasoning

There are several different ways in which emotions might be thought to supplement decision-theoretic reasoning and remedy its deficiencies. One way is for emotions to reduce the number of things we have to take into account, and the number of options we have to consider, so that we are not endlessly deliberating
about a potentially infinite set of factors. Another is for emotions to help with the so-called ‘paradoxes of rationality’ such as the Prisoner’s Dilemma and other problems of commitment, both in committing us to ‘rationally irrational’ courses of action, and in signalling our dispositions to others.

The first problem with ‘pure’ decision-theoretic reasoning is illustrated by Daniel Dennett and adopted by Ronald de Sousa (1987:193-4). He describes a robot being informed that a bomb is about to explode in its hangar. The robot decides to leave the hangar, but the bomb was actually attached to the robot itself. The robot knew this and numerous other features of the environment, but it failed to draw any inferences from such knowledge. So the robot’s designers make some alterations, and programme the robot (presumably after reassembling its blown-up pieces) to draw inferences from the knowledge it has. They repeat the experiment with the bomb in the hangar; but this time the robot sits lost in computation and never acts, so it is again blown up. The designers return to the drawing board, this time instructing the robot to ignore irrelevant implications,

“[T]heir next model [was] the robot-relevant-deducer…R2D1. When they subjected R2D1 to the test…they were surprised to see it sitting, Hamlet-like, outside the room containing the ticking bomb…“Do something!” they yelled. “I am”, it reported, ‘I’m busily ignoring some thousands of implications I have determined to be irrelevant…” but the bomb went off. (Dennett, 1987:42)

This scenario suggests that reasoning alone is insufficient for decision making. It also demonstrates that while it is absolutely crucial that we identify all of the relevant features of our situation when making decisions, this is in tension with the fact that we also need to make decisions without excessive delay, which
means we cannot waste time evaluating the indefinitely large number of potentially relevant features in order to locate the relevant ones. It is important that we have a ‘pre-selection device’ that will allow us to quickly focus our attention on those aspects of our situation that are most relevant. The claim is that the emotions function as such a device. They do this by providing a narrowed focus for a limited time, reducing the number of features we attend to (De Sousa 1987:195). This will also help in considering features of potential situations as well as actual situations. Damasio (1994) notes this latter possibility, saying that, of a set of options, those with outcomes that are ‘marked’ with negative emotions will be quickly removed from the choice set, and allow us to more quickly identify the truly viable options (Jones, 2006: 6).

The decision-theoretic defect which is supposedly remedied by the attention-focusing aspects of emotion is the fact that decision-theoretic reasoning alone gives us no mechanism to quickly and reliably identify those features of our choice situation which are actually relevant. However, there is another problem of pure decision-theoretic reasoning: namely, the existence of situations in which choosing according to the criterion of maximising one’s own interests can lead to outcomes in which one is in fact worse off by that very criterion. These paradoxes of rationality include the Prisoner’s Dilemma (in which there is a dominant strategy – defection will always give the greater payoff no matter what the other player does – which nevertheless leads to a suboptimal outcome), the Paradox of Hedonism (choosing what will give me the most pleasure in all individual choice situations will give me a life with less pleasure overall) and various problems of collective action, cooperation, and commitment. The challenge of these paradoxes of rationality is how individuals can commit themselves to acting in a “rationally irrational” way: that is, in a way which leads to lower payoffs in an individual
choice situation, but which leads to greater payoffs overall. For example, it is
tbetter for me to be disposed to keep my promises because of the benefits of
cooperation with others, but in any particular situation, it will be better for me to
break my promises once I have received these benefits.

The argument is that emotions can help with the paradoxes of rationality,
by disposing us to act in ‘rationally irrational’ ways. They can, in other words, act
as commitment mechanisms which lead us to act in ways which are ‘locally’
irrational (in that they do not maximise our expected payoffs) but which make us
better off overall. This role is obviously related to the role emotions might play in
overcoming hyperbolic discounting, in the sense that in both cases emotions are
supposed to cause us to forgo present gains for greater long-term gains. However,
it is important to understand that in the case of hyperbolic discounting, emotions
are supposed to help us to behave *rationally*: when we reverse our preferences in
the face of a present temptation, we are actually preferring a smaller good to a
greater one, since the (rationally) discounted value of the later good is still larger
than the value of the present good. On the other hand, in the case of paradoxes of
rationality, what emotions would be causing us to do is to act *irrationally*: it is,
after all, irrational to cooperate in the (one-off) Prisoner’s Dilemma since it leads
to a lower expected payoff no matter what the other player does. Nevertheless,
cooperation in the Prisoner’s Dilemma (and keeping one’s promises, or
contribute to a collective good) is *rationally irrational*, in that it is rational for us
to be disposed to act in irrational ways in such situations.¹

¹ This is true of the one-off Prisoners Dilemma. There is also the *iterated* Prisoners
Dilemma where players interact repeatedly, and in these it can be (rationally) rational to
cooperate since there is the possibility that defection in one round will be punished in
future rounds, and that cooperation will be rewarded in future rounds.
How are emotions supposed to help in resolving paradoxes of rationality?

There are two relevant functions: one has to do with the motivational effects of emotions, and the other has to do with the signalling effects of emotional expression and behaviour. The first, motivational, function is essentially the signalling and attention-focusing effect of emotions. If I tend to feel guilt if I betray others or break my promises, then the prospect of feeling guilt will make me averse to betrayal or promise-breaking. A similar point can be made regarding a tendency to carry out threats or avenge wrongs committed against me. If I issue a threat, the intention is that it acts as a deterrent; if I am in a position to decide whether to follow through, the threat has already failed, and since it will be costly to carry it out, it will not usually be rational to do so. A purely rational agent will therefore not ever carry out threats, and if they are known to be purely rational, they therefore won’t be able to make a credible threat in order to deter others from mistreating them. But if they are disposed to carry out on threats, and this is known, a credible threat can be made, which can deter others from harming them. A similar argument applies for a disposition to avenge wrongs. And I can be disposed in this way if I am prone to feeling anger when wronged (Elster, 2000: 46–47).

Of course, this relies on my emotional dispositions being known to others, so that they are deterred from harming me by the prospect of my anger, and are led to cooperate with me, or give me benefits for the promise of future repayment, because of the trustworthiness that results from my tendency to feel guilt. This can happen if I in fact do keep my promises, or carry out my threats, so that others learn that I am disposed to do so. But another way of conveying this information is via the characteristic expressions and behaviours. These expressions serve as a signal to others that I have the relevant dispositions so that they know that they
can trust me, or that they should not harm me, without having to find out 'the hard way' (Frank, 1988: 109-111). And this means that I can get benefits (of cooperation, deterrence, etc.) which I could not get without my emotional dispositions.

4. Facilitating Action

The final way in which emotions contribute positively to practical rationality is in facilitating appropriate action. This is because of the physiological changes involved in emotions, which usually consists of some sort of physiological arousal that primes us for action. The reason why this is an aid to practical rationality is that deliberation by itself may not produce sufficient motivation for action, or it may be too slow in doing so. Emotions can therefore aid practical reason by ensuring that when we identify an option as salient, we are also motivated and primed to carry it out. As Jones writes, “in emotionally laden deliberation, the agent’s motivation is already engaged and so she is prepared to embrace the selected action option” (Jones, 2006:7).

The claim may be that specific emotions prepare us for specific actions – that is, that a particular emotion (e.g., fear) will prepare us for the sorts of actions (e.g., flight or fight) that are called for when we are in danger. Thus, Oatley and Jenkins (1992: 176) suggest that, “emotions make ready a small suite of plans, already assembled either in evolution or individually, that can be called upon when time or other resources are scarce”. And Panksepp (1982:411) claims that emotions “organise behaviour by activating or inhibiting classes of related actions (and concurrent autonomic/hormonal changes) which have proved to be adaptive”. On the other hand, emotions may prepare us for action more generally – that is, all (or most) emotions may involve a kind of general physiological arousal that can be
used as an all-purpose motivational force. Prinz (2004) argues that emotions prime us for behaviour in a more generalised way, rather than for specific behaviours. This could mean that they prime us to do something – whatever it is – or that they prime us for an open-ended range of actions.

2. Pro-Emotion Theories

As we have seen, all pro-emotion theorists agree that emotions play some helpful role and contribute positively to practical rationality. However, any individual pro-emotion theory usually only takes emotions to play some of the roles discussed above. Pro-emotion theories therefore differ on which of the roles they think emotion actually plays in practical rationality (or which roles they focus on as being the most important). They also differ on the question of how emotions perform the relevant functions. The pro-emotion consensus can therefore be broken down into the following rough categories:

1. Evolutionary Psychology Theories
2. Strategic Theories
3. Perceptual Theories
4. Neo-Jamesian perceptual theories (Prinz and Damasio)

I will discuss each of these in turn.

2.1 Evolutionary Psychology Theories

As noted earlier, all of the pro-emotion theories are evolutionary theories in the sense that they believe that emotions have evolved to fulfil a particular function in practical rationality. However, I use the name Evolutionary Psychology theories for this category of pro-emotion theory for the fact that these
theories accept not just the general idea of emotions as evolved functions, but also the more specific claims of Evolutionary Psychology. (Throughout the thesis, I capitalise the label ‘Evolutionary Psychology theories’ to make it clear that I am referring to a specific type of evolutionary claim rather than the general claim of emotions having evolved to serve functions.) According to Jesse Prinz (2004:108) these more specific claims are: (1) that “the mind is an information-processing system that can be broadly characterized as computational”; (2) that “higher cognition can be divided into highly specialized, modular subsystems that are relatively impervious to direct influence from one another”; and (3) that “the processes in those subsystems include rules and reputations [sic] that are adaptive responses to survival challenges faced by our ancestors in the Pleistocene or other earlier environments.”

The most important characteristic of Evolutionary Psychology pro-emotion theories is their acceptance of the notion of modularity (item (2) above). This is the idea that the mind is made up of a host of domain-specific modules, each specialized to solve a different adaptive problem. Jones (2006: 15) describes modules as "special-purpose mini-computers equipped with their own proprietary database and charged with undertaking information processing regarding a single domain." There are various characteristic features of modules. As listed by Prinz (2004:232) modules are:

1. Localized: they have “dedicated neural architecture”
2. Subject to specific breakdowns: there are characteristic ways in which they go wrong
3. Mandatory: they operate automatically outside of our control
4. Fast: they perform their task quickly
5. Shallow: they “have relatively simple outputs"
6. Inaccessible: “higher levels of processing have limited access to the 
representations within a module”

7. Informationally encapsulated: the module cannot receive or make use of 
information from outside

8. Ontogenetically determined: they develop in a particular, biologically 
determined way

9. Domain specific: they “cope with a restricted class of inputs.”

According to the Evolutionary Psychology pro-emotion view, emotions are 
modules which evolved in our ancestral past (in the ‘Environment of Evolutionary 
Adaptedness’, usually supposed to be the Pleistocene) and were selected for because 
they gave our ancestors an adaptive advantage. The main proponents of this view 
are Cosmides and Tooby, who hold that emotions function "to coordinate 
independent programs (programs governing perception, memory, learning, goal 
choice, etc) to produce adaptive responses to certain classes of stimuli that are 
associated with reliable cues signalling their presence in the environment” (Faucher 

The reason why emotions might be adaptively beneficial is that 
deliberation, by itself, is likely to be too slow and inefficient. As Robert W Levenson 
(1994:124) writes,

Evolution has endowed humans with exquisite cognitive powers, which 
provide myriad possibilities of action and inaction for responding to 
environmental challenges. Emotion is the antidote to the problems caused 
by this embarrassment of riches. Hearkening back to an earlier stage of 
phylogenetic development, emotions provide us with access to a limited 
number of simpler time-tested modes of efficient, effective adaptations...In
situations where hesitation could have the most dire of consequences, emotion functions to set aside cognitive processing that is too cumbersome, too obsessive, too self-indulgent, and, ultimately, too likely to be inconclusive.

Cosmides and Tooby (2000) give the example of flight from a predator. Those early hominids that didn’t quickly flee when approached by a predator would not have lasted too long. In order to respond appropriately, it is important to have a hierarchy of concerns that is flexible to incoming environmental cues. In the case of predator threats, I need to be able to immediately turn my attention to it, to de-prioritise whatever had been my main concern beforehand, and to quickly decide on the correct course of action. Emotions are a way of doing this. Described by Cosmides and Tooby (2000:92) as ‘super-ordinate’ modules, the emotions are able to orchestrate and coordinate the sub-modules of our minds because they are able to affect many facets of the body, such as physiological/body states, feeling states, behavioural inclinations and motivations, and cognitive appraisals/mental states. Having access to all these features allows emotions to organise our concerns according to the information we receive from the environment, and then set in motion the bodily changes required to act on those emotional states.

Thus, Evolutionary Psychology holds that emotions play several of the roles described above. Firstly, they help to overcome the deficiencies of decision-theoretic reasoning, by putting in place a fast, mandatory (automatic) and informationally encapsulated mechanism which helps us to quickly make assessments of situations and choose appropriate actions. Secondly, these modules improve access to our reasons by signalling important aspects of our situation.
And thirdly, they facilitate appropriate behaviour by preparing us for the sort of action usually called for in situations eliciting the relevant emotion.

2.2 Strategic Theories

Strategic theories should perhaps be seen as a subtype of Evolutionary Psychology theories; however, they are considered on their own here, firstly because not all of those who advocate strategic theories are necessarily committed to the claims of Evolutionary Psychology, and secondly because they make a more specific and narrow claim about the role of emotions in practical rationality. Essentially, strategic theories consider only one of the roles listed above: the use of emotions in supplementing decision-theoretic rationality and remedying its defects. The main proponent of this view is Robert Frank (1988).

Frank’s argument regarding emotions is essentially that they are *commitment devices* which enable us to resolve the paradoxes of rationality and to overcome time-inconsistent preferences of the sort described by Ainslie. As Elster (2000: 45) points out, emotions have often been seen as problematic for self-control and pursuit of long term goals, but strategic pro-emotion theorists “turn the idea on its head, by arguing that *emotions are solutions rather than problems.*” Frank argues that emotional dispositions can motivate us to act in ways that we would not act if we were purely decision-theoretically rational; that is, they promote rational irrationality. For example, Frank argues that an emotional disposition can enable me to make credible threats because it will also lead me to follow through on those threats. As noted above, it will usually not be rational to follow through on a threat, since this involves costs and risks. So if I am (and am believed to be) rational, I won’t be able to make a credible threat, since no one will believe that I will follow through. On the other hand, if I am disposed to become
enraged or spiteful when wronged, then I will be prone to lash out and carry out a threat even at a cost to myself, because of the motivational force of my anger. So a disposition to become angry will solve the commitment problem and allow me to make credible threats, which make it less likely I will ever need to carry them out. A similar argument can be made regarding a tendency to be vengeful. As Frank argues,

Perfectly rational persons with perfect self-control would always seek revenge whenever the future reputational gains outweigh the current costs of taking action. The problem…is that the gains from a tough reputation come only in the future while the costs of vengeance-seeking occur now... Being predisposed to feel anger when wronged helps solve this impulse-control problem. (1988:83)

Analogous arguments can be made for other emotions, such as envy or guilt. A perfectly rational agent would always cheat if given the opportunity to do so without being caught. But this makes them less trustworthy, and it is less likely that they can persuade others to cooperate with them in mutually beneficial endeavours. On the other hand, if the agent is disposed to feel guilt and remorse when they cheat, the prospect of experiencing these aversive states will motivate them to behave honestly. If someone “is emotionally predisposed to regard cheating as an unpleasurable act in and of itself – that is, if he has a conscience – he will be better able to resist the temptation to cheat” (Frank, 1988:82).

Frank even believes that romantic love can be explained as a commitment device. If our decisions to enter into and remain in committed relationships were based solely on rational assessment of the current costs and benefits, we would be more likely to leave relationships whenever they weren’t going well. Frank
imagines, for example, that “my wife” would leave if “Tom Selleck bought the house next door and announced his availability” (1988: 196); presumably the point could also be made for those with male partners and for more plausible objects of desire. If we were prone to leave relationships on the basis of a rational cost-benefit assessment, then this would prevent us from being able to make the kind of long-term plans and commitments (mortgages, children, and so on) that marriage and other committed relationships require; even if we did want to enter into them, our partners would probably not take that risk. On the other hand, if we entered relationships and stayed in them on the basis of romantic love, we would be more likely to stay in the face of short-term temptations to leave, firstly because we want to stay, and secondly because we would feel guilty for leaving (Frank, 1988: 196-197).

Frank’s argument centres on the notion of rational irrationality: the idea that it is rational to have emotional dispositions that lead us to act in certain ways, but that it is irrational to act in those ways. According to Frank, emotions perform their function by making us ignore the rational balancing of costs and benefits in a particular choice situation, and instead to perform an act which has a lower immediate payoff. The reason that it is rational for us to have these emotional dispositions is that the overall payoff in the long term is greater. On the other hand, Greenspan has noted that emotions can alter the payoffs of our choice situation so that the emotionally-motivated action has the best immediate payoff as well as the best long-term payoff. She illustrates this phenomenon with an example of someone who has tolerated numerous infidelities from their partner, and who finally works themselves up into a rage (Greenspan, 2000: 471-473). In this enraged state, they can issue a credible ultimatum: that unless the infidelities cease, they will leave the relationship. Greenspan’s point is not just that the
ultimatum is credible because the person’s anger will lead them to (irrationally) carry out their threat. Rather, she notes that having become angry enough to issue the ultimatum has shifted the payoffs of the various outcomes: if the partner ignores the ultimatum and continues to be unfaithful, staying would now constitute an unendurable humiliation and would signal to the partner that their feelings can be safely ignored.

2.3 Perceptual Theories

Perceptual pro-emotion theories are so-called because they hold that emotions either are a kind of perception, or that they are analogous to perception. In particular, emotions are supposed to be perceptions of values or evaluative properties. Ronald de Sousa, for example, believes that emotions are closely analogous to perceptual faculties and mimic their behaviour (de Sousa, 1987:154-155, 303). De Sousa believes that emotions, because they function like perceptions, help us to solve what he calls the ‘philosophers’ frame problem’, (192-193) which is the problem experienced by Dennett’s robot described earlier: how to pick out just the information we need from all of the information we have.

According to de Sousa, emotions can solve the philosophers’ frame problem by mimicking the way in which perceptual faculties operate; in particular, by temporarily inducing a kind of informational encapsulation that is typical of perception. Although he does not believe that emotions are modules (or that they are themselves informationally encapsulated), he thinks that emotions can temporarily induce this characteristic of modules:

The role of emotion is to supply the insufficiency of reason by imitating the encapsulation of perceptual modes. For a variable but always limited time, an emotion limits the range of information that the organism will take into
account, the inferences actually drawn from a potential infinity, and the set of live options among which it will choose. (de Sousa, 1987:195)

What this seems to mean is not that emotions are encapsulated, but rather that they operate by encapsulating our decision-making, or turning it into a temporary module. Encapsulation essentially means that the module will not incorporate ‘outside’ information: for example, the stick in water continues to look bent even when we know it to be straight. So emotions pick out a subset of available information and then allow only that information to be incorporated into our deliberation. Thus, de Sousa’s theory is that “emotions are species of determinate patterns of salience among objects of attention, lines of inquiry, and inferential strategies” (de Sousa, 1987: 196). On this view, then, emotions therefore play mainly the first and second roles mentioned in Section 1: they improve access to our reasons, and remedy the deficiencies of decision-theoretic rationality.

2.4 Neo-Jamesian Theories

The final category includes two theorists – Prinz and Damasio - who have a lot in common with the other types of pro-emotion theory above, but also have some important differences. Prinz’s theory is in part an Evolutionary Psychology theory (since he believes that emotions are modules) and partly a perceptual theory (since he thinks that emotions are perceptions). Similarly, Damasio’s theory has a lot in common with the perceptual theorists. However, they are considered in their own category because of their belief that emotions are a particular kind of perception: namely, a perception of bodily states and changes. As such they are referred to as ‘neo-Jamesians’.
According to Prinz, emotions are “embodied appraisals,” that is, “mental states that represent core relational themes by registering bodily changes that co-occur with those themes” (2004:142). Emotions have both ‘nominal contents’ – the bodily changes – and ‘real contents’ – their core relational themes. The idea behind core relational themes (C.R.T.) is derived from Lazarus (1991:122) and includes such things as “a demeaning offense against me and mine” as a C.R.T. for anger, “having transgressed a moral imperative” for guilt, “facing an immediate, concrete, and overwhelming physical danger” for fear, and “making reasonable progress toward the realization of a goal” for happiness (2004:16). These are things that are important for survival and flourishing, and they reliably co-occur with certain bodily changes, because they have evolved to do so (2004:68). For example, facing immediate physical danger reliably co-occurs with a racing heart, perspiration, and shaking. Emotions perceive these bodily changes and as a result represent the core relational themes that co-occur with them (2004:69).

Building on ideas developed by Dretske (1981, 1986), Prinz suggests that as representations emotions are ‘set up to be set off’, which is to say that “a state represents that which it has the function of reliably detecting” (Prinz, 2004:54). This may seem on the one hand tautologous, and on the other hand rather mysterious, but Prinz apparently means that an emotion (say, fear) functions to reliably detect danger as instantiated by particular situations, in that we are set up to undergo particular bodily changes whenever a situation is dangerous; these bodily changes characteristic of fear then represent danger, in that our perception of them amounts to a perception that we are in danger. Thus, on Prinz’s view, emotions improve our access to our reasons, and facilitate appropriate action.

Damasio also believes that emotions are perceptions of bodily changes and that they represent core relational themes by reliably tracking them (Jones,
2006:19). As noted above, he puts forward this view as a result of his experiences and research involving patients with frontal lobe damage, who appear to be cognitively normal except in their (lack of) emotion experience, and who display debilitating deficits in planning and decision-making. Damasio therefore argues that emotions facilitate practical reason by bringing some options to the forefront of our attention and allowing others to be rapidly discarded. They do this via what Damasio calls ‘somatic markers’. These are positive or negative valences which are attached to aspects of our environment (e.g., people, potential choices). The perception of these aspects causes mental representations which in turn cause bodily changes. An emotion is the experience of these changes, which then ‘marks’ the images (and the things they represent) as negative or positive in particular ways (1994:134-138). This allows us to quickly ascertain what is relevant, and just as quickly discard ‘bad’ options.

Without emotion, Damasio argues, we would not be able to do this. He gives the example of a business owner trying to decide whether to meet and make a deal with her best friend’s arch-enemy. If we were to reason unemotionally, as Damasio’s frontal-lobe-damaged patients do, we would take the different scenarios apart and…perform a cost/benefit analysis of each of them. Keeping in mind ‘subjective expected utility’, which is the thing you want to maximize, you infer logically what is good and what is bad….You are, in fact, faced with a complex calculation, set at diverse imaginary epochs, and burdened with the need to compare results of a different nature which somehow must be translated into a common currency for the comparison to make any sense at all. (1994:158)
He continues: “if this strategy is the only one you have available, rationality, as described above, is not going to work. At best, your decision will take an inordinately long time....At worst, you may not even end up with a decision at all because you will get lost in the byways of your calculation” (1994:158). This is obviously a similar point to that made by Dennett’s robot and de Sousa’s ‘philosophers’ frame problem.

On the other hand, a ‘normal’ person equipped with emotions can use somatic markers as a short cut: thinking about a potential bad outcome causes “an unpleasant ‘gut feeling’, i.e., a negative somatic marker, which ‘forces attention on the negative outcome to which a given action may lead, and functions as an automated alarm signal....The signal may lead you to reject, immediately, the negative course of action and thus make you choose among other alternatives” (Damasio, 1994:159-60). Therefore, somatic markers “assist the deliberation by highlighting some options (either dangerous or favourable) and eliminating them rapidly from subsequent consideration” (Damasio, 1994:160).

The somatic marker hypothesis also, according to Damasio, explains the deficits in planning agency in patients with damaged frontal lobes. Damasio thinks that our ability to prefer future rewards rather than present ones is due to our ability to ‘mark’ future rewards with emotions that increase their salience in our deliberations. In patients with damaged frontal lobes, however,

even if the images of future consequences were stable, damage in the ventromedial prefrontal cortices would preclude the evocation of pertinent somatic state signals...and consequently the relevant future scenarios would no longer be marked. Their significance would not be apparent, and their impact on the decision-making process would be voided, or easily overcome by the significance of immediate prospects. (Damasio, 1994:193)
The neo-Jamesians therefore hold that emotions play much the same roles as in the perceptual theories. They improve our access to our reasons, and remedy the defects of decision-theoretic reasoning, along with facilitating appropriate action. According to Damasio, they also facilitate planning agency, since the way in which emotions aid reasoning is one which allows us to ‘mark’ future consequences in ways that make them more salient in present deliberation.

**Conclusion**

In this chapter I have considered the so-called pro-emotion consensus from two different angles. Firstly, I have outlined the various *roles* that emotions might play in facilitating practical rationality. Secondly, I have given a brief overview of the major subtypes of pro-emotion theory and the particular functions that are given to emotions in each of them. The overall idea is that emotions help us in decision-making because they are quicker and more ‘automatic’ than ‘pure reason’, which cannot properly identify appropriate options, pick out the most relevant information, or motivate us to act in appropriate ways. According to the pro-emotion consensus, we would not, and could not, be practically rational without the help of emotions.

In general, though, the pro-emotion consensus needs to do more than just give an account of the facilitative functions of emotions in practical rationality. Specifically, there is the commonsense observation, confirmed by centuries of experience and observation, that emotions do quite often cause us to act *irrationally* (and not in the sense of ‘rationally irrational’), and that extensive emotional education and training (particularly involving the ability to *regulate* emotions) is required in order for emotions to be helpful and not harmful in
rational decision-making. Indeed, this point is made quite forcefully by most historical theories of the emotions prior to the modern pro-emotion consensus.

Of course, the modern pro-emotion consensus is often framed by its advocates (and by some of its critics) as doing away with an old-fashioned view of emotions as purely disruptive influences on practical rationality. However, I will show in the next chapter that this is a misrepresentation of the reality of historical theories of emotion. In fact, historical theories are quite diverse in their views of how emotions relate to practical rationality, and many of them hold that emotions have the very same kind of helpful functions as those discussed in this chapter. They do tend to claim that emotions will lead to irrational behaviour if not properly trained and regulated by reason. But I will argue that this is not only a plausible insight, but one which should be accommodated by pro-emotion theories.

Fortunately (I will argue) the historical and commonsense insight is not incompatible with the pro-emotion position. But noting general compatibility is one thing, and actually explaining how to accommodate the insight within given pro-emotion theories is another. In chapters three and four I will explore the various kinds of pro-emotion theory in greater detail, with a view to understanding how they can accommodate the phenomenon of emotionally-driven irrational behaviour and the possibility of emotional training and regulation by reason. I will give reasons for thinking that strategic and Evolutionary Psychology theories are inadequate in these respects, in the former case because the theory is not a complete theory of the functions of emotions in practical rationality, and in the latter case because the commitment to Evolutionary Psychology creates a tension in the explanation of irrational behaviour and emotional regulation. On the other hand, perceptual theories are able to give a coherent and satisfactory account, and in the final chapter I show how a complete
perceptual theory can be used to offer practical advice for managing emotionally driven irrational behaviour on both a personal and a public policy level.
Chapter Two: Historical Theories of Emotion and Emotions as the Cause of Irrational Behaviour

In the literature on the pro-emotion consensus, both pro-emotion theorists and their critics can be seen to make the claim that the pro-emotion consensus is a novel development in studies of emotion. The pro-emotion consensus is supposedly set in opposition to a contrary historical consensus, according to which emotions are entirely anti-theitical to rational decision-making. According to Jon Elster (1999: 284) “the traditional view is that emotions interfere with rational choice. They are, as it were, sand in the machinery of action.” As previously noted, Karen Jones (2006: 5) describes the pro-emotion consensus as a novel view which is “replacing old dogmas that emotions disrupt practical rationality.” Similarly, Evans and Cruse (2004: xii) make the claim that “for thousands of years, it was almost universally assumed by Western thinkers that emotions were, at best, harmless luxuries, and at worst outright obstacles to intelligent action.” Mameli (2004: 160) claims that the view of emotions as interfering with our ability to make the right choices “has been the received view in psychology and philosophy since Plato.” Gross and John (2002: 297) think that the evolutionary-functional view of emotion “has provided a much-needed corrective to the long-standing view of emotions as dysfunctional and disruptive.” De Sousa (1987: 3-4), while acknowledging that “emotions have fared unevenly in the history of Western thought”, nevertheless says that “a long tradition views all emotions as threats to rationality.” And Damasio gives his book the title *Descartes’ Error*, precisely because he takes Descartes’ dualism to hold that ‘pure’ reason exists separately from bodily processes such as emotion.
In this chapter I am going to show that this idea of a 'historical consensus' opposed to the pro-emotion consensus is inaccurate and misleading about historical views of emotion. In fact, most major historical theories of emotion have much in common with the pro-emotion view. Far from holding that emotions are entirely detrimental to practical rationality, many historical philosophers took emotions to play roles or functions in rational decision that are very similar to those suggested by the pro-emotion theorists. That is, most historical theorists thought that emotions played a helpful and perhaps indispensable role in practical rationality. Of course, they also thought that emotions tended to cause irrational behaviour unless they were under the control of reason, but (as I will try to show) this does not put them at odds with the pro-emotion consensus. This leads to the question as to where, precisely, the disagreement lies between the pro-emotion consensus and the sorts of things commonly asserted about emotions in historical theories.

In the next section I will consider some of the major historical theories of emotion: firstly, from Ancient Greece (Plato, Aristotle and the Stoics), and secondly, from early modern philosophy (Descartes, Hume, and Kant). The aim is not to present an original or novel interpretation of these philosophers' ideas. Rather, it is to outline their theories of the emotions as they are commonly interpreted by scholars in the history of philosophy. I will show that while some of these historical figures tended to be hostile to emotions, more of them (and the more influential among them) tended to see emotions as playing a helpful role in practical rationality so long as they were properly trained and regulated by reason.

Following this section, I will draw out what I consider to be the main lessons of examining historical theories of emotion, and discuss what, if any,
differences exist between those historical theories that are not entirely hostile to emotion, and the modern pro-emotion consensus. There are two possible differences: firstly, that the historical theories divide reason and emotion into separate faculties, whereas the pro-emotion consensus thinks that emotion is part of the reasoning faculty; and secondly, that the historical theories seem to at least allow the possibility of ‘pure’ unemotional reason (that is, that it might be possible to reason without emotion), whereas the pro-emotion consensus takes emotion to be necessary to reason. I will argue that these differences are not in fact very dramatic, and that the major claim of the historical theories – that emotion will lead to irrational behaviour unless properly trained and regulated by reason – is not incompatible with the view that emotions are helpful or even indispensable to practical rationality.

Finally, I will discuss the implications of all of this for the pro-emotion consensus. As I have said, pro-emotion theorists should not deny that emotions can, and often do, lead to irrational behaviour. They also should not deny that emotional training, and regulation of emotions by reason, is both possible and desirable in order to prevent emotionally-driven irrational behaviour. But in this case, a satisfactory pro-emotion theory will have to explain not only how emotions contribute to practical rationality, but also how they sometimes derail it, and how it might be possible to regulate them to prevent this. At the end of the chapter, I therefore present two case studies of emotionally-driven irrational behaviour to serve as illustrations and test cases in the subsequent chapters.

1. Historical Theories of the Emotions
In this section I consider some major historical theories of emotion, selected because they are influential on subsequent thought. It should be noted, however, that there are some terminological complications in discussing 'historical theories of emotion', since philosophers have used several different terms, and meant different things by them (Schmitter, 2010). For example, the terms 'passions', 'affects', and 'sentiments' were all used in various ways, historically, to mark different distinctions. However, I will mostly gloss over these problems in the following discussion, and refer to theories of emotion. This would not be a good idea generally, but for my purposes – that is, in discussing historical theories of the function of emotions rather than their nature – I do not think it will cause major misinterpretation.

1.1 Theories of Emotion in Ancient Greece

1.1.1 Plato

Plato’s attitude towards the emotions is widely held to be hostile, and there is a great deal of justification for this view. Plato held that the emotions were the cause of akrasia, or weakness of will (Knuutila & Sihvola, 1998: 5); Ainslie also claims that he held them to be passive (hence, ‘passions’) - external forces that push and pull us, rather than part of who we are (Ainslie, 2001: 4). His famous metaphor for the tripartite soul - the charioteer with two horses, one (rational and moral impulses) obedient, the other (appetites and passions) unruly – suggests that emotions are constantly threatening to break from our control and pull us in the wrong direction (Goldie, 2000:113), if they are not always kept under the tight

---

2 An examiner has pointed out, however, that this may be a misunderstanding of Plato. The particular claim that emotions are external forces is not central to the arguments I will be making in this chapter.
control of reason. Because of this, Plato holds that the passions are dangerous, both for individual rationality, and for a well-ordered society (Nussbaum, 1986: 105). In several of his writings Plato seems to argue against being guided by emotions, and is suspicious of art forms, such as poetry, which excite the emotions of the audience (Nussbaum, 1986: 201). Rather, we should aim to be guided only by reason.

Quite a few philosophers have pointed out, however, that Plato’s hostility to emotion was not complete: he did not seem to think that we would be better off without having emotions altogether. Consider the metaphor of the charioteer again. Plato recommends that the charioteer keep firm control of the unruly horse (emotion) in order to avoid disaster, but does not tell us that we should cut the horse loose and proceed with only one. Rather, Plato thinks that all three (the charioteer and two horses) should work together harmoniously. When the emotions are properly under the control of reason, they can be helpful (Brown, 2006: 32). The metaphor also suggests that reason requires emotions in order to supply motivational force: a charioteer without horses cannot go anywhere, and a charioteer with only one horse will be much slower and less efficient. Nussbaum (1986: 214) makes this point when she writes that “since the charioteer is clearly the planning, calculating logistikon, we are invited by the image to consider that intellect alone is a relatively impotent moving force…we require the cooperative engagement of our non-intellectual elements in order to get where our intellect wants us to go.”

Plato disapproves of those who are ruled by their emotions, but does not (at least in the later works) think that having emotions is bad. In fact, properly controlled by reason, they are crucial to the good life. This is because when they are properly controlled and guided by reason, they can help us to perceive the
good and beautiful (Nussbaum, 1986: 214-5), and thereby to achieve understanding and insight. As Knuutila and Sihvola (1998: 5) write, “even though the unruly horse requires continuous control by reason, it has intrinsic value of its own and an important motivating role in the pursuit of the good. So the intellectual element is not sufficient for making good and correct choices; important information is also provided by the passions.”

1.1.2 Aristotle

In contrast to Plato, the Aristotelian view of emotions is a far more sympathetic one. It is true that, like Plato, Aristotle takes emotion to be the cause of *akrasia* (Ainslie, 2001: 4; Schmitter, 2010a). However, this was not because emotion as such was an external disruptive influence, but rather because the akratic was feeling the *wrong* emotion, or the wrong intensity of emotion: in particular, emotions have become so strong that they ‘crowded out’ other considerations (Elster, 2000: 9). The remedy is not to get rid of emotions, but to ensure that we are feeling appropriate emotions.

Aristotle’s view was that of *metriopatheia*, namely that moderate emotions are good and useful (Sorabji, 2000: 22). It is possible that he conceived of ‘ideal’ rationality as emotionless, but he nevertheless saw that actual humans are not capable of this ideal, and that as we are, we require emotions to make good decisions and to be happy (Sorabji, 2000: 190-191). In fact, emotions are essential to virtue, which involves feeling the right emotion, in the right intensity, at the right time (Goldie, 2000: 113). This does not mean that we should never feel strong emotion, but rather that our emotions should be appropriate to the situation (Sorabji, 2000: 195). As Sorabji (2000: 22) writes, this means that we should avoid both too much and *too little* emotion, and “emotions may need
intensifying as often as quieting, if the mean is to be hit.” This ability to respond appropriately (i.e., with the correct kind and strength of emotion) is necessary for practical wisdom (Nussbaum, 1986: 307) and for choosing and acting virtuously; it is therefore an indispensable part of the good life (Brown, 2006: 32). In fact, “without the right ‘passion’, the very same choice and action will cease to be virtuous” (Nussbaum, 1986: 383).

The reason why emotions are necessary for practical wisdom, for Aristotle, is that they give us information about things that are of concern to us, and they provide us with motivation for appropriate action. Knuutila and Sihvola (1998: 11) claim that the analysis of emotions in the *Rhetoric* involves three elements:

1. The emotional evaluation states that something positive or negative is happening to the subject or to someone else in a way that is relevant to the subject.
2. A pleasant or unpleasant feeling aroused is an awareness of one’s or someone else’s changed position in a typical situation.
3. The behavioural suggestion is a spontaneous impulse towards action. (1998:11)

Wonder, for example, motivates us to learn, and is therefore the “beginning of philosophy” (Brown, 2006: 142).

We can see that these elements describe some of the same functions for emotions proposed by pro-emotion theorists: emotions give us access to our reasons, and motivate us to act in response to them. In particular, this view of emotions is quite close to the ‘perceptual’ pro-emotion theories. Nussbaum (1986: 308) writes that, for Aristotle, “frequently the perception of the salient features will be achieved in a way that relies centrally upon the discriminating power of passion….Aristotle’s accounts of the so-called ‘practical syllogism’…ascribe to the desires a sorting or discriminatory power: out of the many things presented to the
agent by thought and perception, desire will single out some and not others to be foundations of action.” The similarity to de Sousa’s idea of emotions as ‘patterns of salience’, or Damasio’s ‘somatic markers’, is obvious.

Of course, Aristotle believes that emotions need to be properly trained and educated in order to make sure that they are giving us accurate information, and motivating us in the right way (Goldie, 2000: 113), otherwise they will not perform this function (or perform it badly). Aristotle thinks that emotions are “responsive to training, and therefore able to play a constructive role in moral motivation, impelling the person towards more appropriate objects in keeping with his or her evolving conception of the appropriate” (Nussbaum, 1986: 307). Thus, Aristotle gives advice not only about how we should train children and improve our own emotional dispositions (by habituation, practice, and even ‘faking’ until the real emotion follows), but also about how to arouse or quell emotions in others.

1.1.3 The Stoics

The Stoic theory of emotions is probably the historical theory that comes closest to being genuinely opposed to the pro-emotion consensus. As is well known, the Stoics (or at least some of their leaders) advocated eliminating all passions in order to reason well and attain wisdom. Sorabji (2000: 7) writes that “Stoicism as formulated by Chrysippus was opposed to nearly all emotion.” The passions were thought to be ‘disobedient to reason’ (Baltzly, 2010), and the sage – someone who had attained wisdom – would not experience them.

The reason why the Stoics disapproved of emotions is that they considered them to be false evaluative judgements. Emotions, according to the Stoics, in fact consist of two evaluative judgements: firstly, that something is good or bad, and
secondly, that some kind of reaction or action is appropriate (Sorabji, 2000: 29). But both of these judgements are false, since (on the Stoic view) the only things that are genuinely good or bad are things that are under our control (e.g., virtue and rational decision). Everything else is ‘indifferent’ – neither good nor bad – and therefore it’s inappropriate to react to it (Sorabji, 2000: 7). Because emotions are mistaken judgements about what is good or bad, they are directly harmful to rational decision in and of themselves. But they are additionally harmful in also motivating us to act: “an emotion is not an idle and innocuous false belief, but a false belief taking effect in the agent’s behaviour” (Brennan, 1998: 31). In order to attain practical wisdom, then, we ought to train ourselves to ignore the ‘appearances’ presented to us by our emotions, and not to let them influence our actions.

The apparent Stoic notion of totally emotionless reasoning as the idea of rationality suggests that this may be what some pro-emotion theorists have in mind when they describe a ‘traditional’ view of emotions as antithetical to reason. And it is true that the Stoic view was highly influential on later thinkers, in particular Galen and medieval Christian philosophers. However, even the Stoic theory is not completely at odds with the modern pro-emotion view, and it would be mistaken to think that the Stoic hostility to emotions is sufficient to vindicate the pro-emotion assertion of a ‘traditional consensus’ holding that emotions are antithetical to rationality.

Firstly, it should be noted that the Stoic view of what the emotions do is quite similar to modern pro-emotion theories of what they do. The Stoics thought that emotions were evaluative judgements, that is, that they had cognitive or propositional content, whereas the modern pro-emotion consensus is not committed to a cognitive thesis about emotions. Nevertheless, the more general
idea that emotions are appraisals (whether these are beliefs or not) is a common feature between the Stoic and the modern pro-emotion view. Both the Stoics and the modern pro-emotion consensus hold that emotions (a) tell us that something in the environment is good or bad in a particular way, and (b) motivate us to react accordingly. The difference between the Stoics and the pro-emotion theorists is just that the Stoics believe that these appraisals are uniformly false: they tell us that things are good or bad when they actually aren’t. There is nothing truly good or bad, only so from our perspective.

In fact, the Stoics do allow that the sage would experience certain affective states. Sorabji (2000: 47) writes that the Stoics approve of eupatheiai, or good states of feeling, of which there are three kinds: joy, will, and caution. These are acceptable because they are appraisals of things that are under our own control, and therefore accurate indications of value. “Instead of over-valuations of and reactions to indifferent, alien features of the external world, eupatheia are cognitively appropriate (and active) judgments directed at the things that are truly important to the good life, particularly at other rational beings” (Schmitter, 2010a). The Stoics did not think that these states were passions (Baltzly, 2010). However, we would probably classify them as emotions (Sorabji, 2000: 47). So while the Stoic view is closest to being the kind of ‘anti-emotion’ view that could be viewed as opposed to the pro-emotion consensus, even this view allows a limited place for the emotions in rational decision.

Moreover, while the Stoic theory in general was highly influential to later philosophy of emotion, the particular claim that passions were entirely bad and should be eliminated was not widely accepted, either by later philosophers or in Ancient Greece. Sorabji notes that there was a major debate in Ancient Greek
philosophy between the Stoic anti-emotion position and Aristotelian *metriopatheia*; he says (2000: 196) that “the majority view among post-Aristotelian philosophers favours Aristotle’s ideal of *metriopatheia*, rather than the Stoic *apatheia*, although the picture is not clear-cut.” Similarly Knuutila and Sihvola (1998: 17) tell us that “the Stoic *apatheia* was much criticized in ancient philosophy. It was regarded as both impossible and inhuman.” And this also holds true of theories of emotion after the Ancient Greeks. As Brown (2006: 34) notes, “most philosophers after the Stoics were keen to distance themselves from this extreme view, believing that the passions were all by nature good.” And Schmitter (2010b) writes that “even those thinkers who seem to owe the most to Stoicism (i.e., Descartes and Spinoza) explicitly criticized certain of its doctrines, including the view that the passions are erroneous judgements.”

### 1.2 Theories of Emotions in the 17th & 18th Century

Views of the emotions in pre-modern and early modern philosophy were quite divergent. Many philosophers emphasised the value of pleasurable emotions, or of emotional tranquility (Schmitter, 2010). Others did follow the Stoics in holding that emotions are irrational because they arise from false judgments: as Nussbaum (2001: 369-70) writes, such philosophers took emotions to be “irrational not in the way that hunger is irrational, but in the way that a belief in the flatness of the earth is irrational: false, based on inadequate evidence, cultural prejudice, false premises, and bad argument.” The aim should therefore be to have the false judgements “set right by true premises and good arguments;” in other words to have reason overcome emotion. However, many other philosophers abandoned some or all aspects of the Stoic analysis, and took emotions to be generally good and useful.
1.2.1 Descartes

As I noted in the introduction, at least some pro-emotion theorists take Descartes to be a paradigmatic example of the supposed anti-emotion ‘traditional consensus’. Damasio explains that his title, *Descartes’ Error*, refers to Descartes’ abyssal separation between body and mind, between the sizable, dimensioned, mechanically operated, infinitely divisible body stuff, on the one hand, and the unsizable, undimensioned, unpush-pullable, nondivisible mind stuff; the suggestion that reasoning, and moral judgment, and the suffering that comes from physical pain or emotional upheaval might exist separately from the body. (Damasio, 1994: 218)

According to Damasio, the separation of rational thought (located in the soul) and emotion (located in the body) commits Descartes to the idea that emotion plays no part in rational deliberation or decision-making. And he thinks that this error has been perpetuated by other philosophers of emotion: “for many, Descartes’ views are regarded as self-evident and in no need of reexamination” (Damasio, 1994: 218).

However, as Brown (2006: 7) points out, Descartes does not think that we (as embodied creatures) can reason without the emotions. In fact, he gives a special role to the passions: they “are the lynchpins of mind-body unity” and ensure the integrity of the mind-body union. In particular, the emotions supply motivational force: Descartes says in the *Passions of the Soul* that “the main effect of every human passion is to arouse the soul and make it will the body to move in the way the passion prepares the body for” (Descartes, 1985:343). In particular, their function is to alert the soul to things that are potentially harmful or beneficial, and
to motivate the soul to perform actions that will aid in self-preservation: “their natural use is to incite the soul to consent and contribute to actions that may serve to preserve or in some way improve the body” (Descartes, 1985:376). So the passions detect features of a situation that are relevant to our welfare or concerns, and bring these features to our attention, along with a dose of motivation to perform certain actions.

Descartes does believe that emotions can often be harmful even when performing this function, though, because they are prone to error and distortion, many things that are harmful to the body cause no sadness initially (or even produce joy), while other things that are useful to the body are disagreeable at first. Furthermore, the passions nearly always exaggerate the size and importance of the goods and evils they represent, inciting us to pursue the goods and flee from the evils with more ardour and zeal than is appropriate. (Descartes, 1985:376-377)

For Descartes, passions are forms of perception: they represent the nature of things. But, like perceptions, they can be inaccurate and lead us to form false impressions of how things really are. This is because “practical efficiency comes at the cost of some fallibility. Under non-standard conditions (e.g., when we suffer from illness, or body parts go missing), our signposts may lead us astray” (Schmitter, 2010c). It is therefore necessary, for good decision-making, for us to use our reason and experience to draw on our knowledge of what is really good or bad, rather than being guided solely by the emotions.

Nevertheless, Descartes believed that the passions were “all intrinsically good” (Descartes, 1985:403-404), in two ways. Firstly, they are the source of “the sweetest pleasures in this life.” But secondly, they perform a valuable function in
getting us to act efficiently and appropriately. What was to be avoided was not
emotions per se, but only “their misuse or their excess” (Descartes, 1985:403-404).
We should make sure that the passions are not allowed to directly cause action,
without the intervention of reason (Brown, 2006: 21). Training, and the
achievement of virtue, will allow us to “alter our dispositions to feel the passions,
by way of a kind of internal bodily training generating new and improved
‘habits’”(Schmitter, 2010c). This training does not eliminate the disposition to feel
the passions, but rather ensures that they will be felt only under “the appropriate,
rationally endorsed circumstances” (Schmitter, 2010c).

The similarity of Descartes’ view to certain of the pro-emotion theories –
in particular, perceptual pro-emotion theories – should be obvious. Brown, in fact,
explicitly draws attention to the similarity between Descartes’ view and de Sousa’s
claim that emotions allow us to solve the ‘frame problem’ (Brown, 2006: 73-74).
The main difference is that Descartes seems to allow, at least in theory, for there
to be a ‘pure reason’ which is divorced from the body and therefore from emotion.
Damasio, as we have seen, takes issue with this idea. But the disagreement may be
less strong than it first seemed. Descartes may allow for the theoretical possibility
of disembodied reason without emotion, but he also believes that emotions are
necessary for the rational deliberation of embodied creatures such as ourselves.

Brown agrees that Descartes does not rule out the possibility of
“disembodied thought,” but notes that he also acknowledges “that much of our
thought is occupied with matters that require the co-operation of the body,
including the practice of experimental science….Far from being impediments to
reason, passions and sensations play an indispensable role in the investigation and
navigation of the natural world….Although the body has the tendency to swamp
the mind with its concerns, it is possible, according to Descartes, to reason well ‘in
this life’, and precisely so because it is possible to use the sensory resources provided by the body as supplements to reason.” (Brown, 2006: 59).

1.2.2 Hume

Hume does not fit well into the implied dichotomy between ‘pro-emotion’ and ‘anti-emotion’ views. He does not seem to think that emotion is necessary for rational deliberation, but neither does he think that the passions are purely disruptive external influences on the operation of ‘pure reason’ that defines us as ourselves. He acknowledges, and rejects, the idea that “every rational creature…is oblig’d to regulate his actions by reason, and if any other motive or principle challenge the direction of his conduct, he ought to oppose it” (Hume: 2008: 296). Instead, he famously asserts that “reason is, and ought only to be, a slave to the passions and can never pretend to any other office than to serve and obey them” (Hume, 2008: 297).

However, this does not make Hume a member of the ‘pro-emotion consensus’ either. Emotions, on Hume’s view, guide us in what we should do, but this is not guidance in the sense of improving access to our reasons (i.e., helping us to perceive reasons that are independently there). Rather, the emotions give us reasons that we would not have had, and more importantly could not have acted on, without their presence. Hume’s focus is on what motivates us to act, and also what motivates us to undertake deliberation. He believes that reason has no motivational force of its own. In fact, all reason does is find relationships (causal, logical, etc.) between ideas, and discovers truth or falsehood (Brown, 2011: 221-2). It cannot motivate us to act; and because it cannot produce motivation, it also cannot oppose any existing motivation, since this would require reason to have a contrary motivating force. When reason appears to motivate us, this is only
because we happen to have a pre-existing desire to which the reason is appropriately connected.

So it is not possible, according to Hume, for the passions to be controlled or ruled by reason. Instead, reason has an instrumental use in helping us to achieve the ends given to us by the passions. Hume says that “abstract or demonstrative reasoning never influences any of our actions except by directing our judgment concerning causes and effects” (2003: 294). The operation of reason can motivate us to act indirectly, by creating the conditions in which a particular emotion will occur (Schmitter, 2010b) – for example by deducing a particular conclusion, or connecting a cause and effect.

On this view, emotions are not themselves either rational or irrational; they are “original facts and realities, complete in themselves” (Brown, 2011: 227). This might lead us to think that Hume did not see a place for emotional training or regulation. Since he did not think it was possible for reason to get any control over the passions, this would seem to rule out the regulation of emotion by reason. And if our emotions are just facts about the way things are – a reality that we are given to work with – then there is not much room, or point, to emotional training.

Nevertheless, Hume does seem to allow some room for us to improve our emotions. Indeed, as a leading proponent of the sentimentalist school, he believes that moral sentiments are the basis of our moral judgements (Denis, 2011). According to Hume, these moral sentiments are a product of sympathy, which is the capacity not just to form beliefs about what others are feeling, but to feel what they are feeling (Brown, 2011: 232-233). Hume acknowledged that sympathy operates more naturally on those close to us (and those who are like us). However, he advocated regulation of sympathy by the ‘general point of view’ – for example, we ought to think about the good of society rather than just our own interests. In this
way, the operation of reason, by leading us to the conclusion that (for example) other people are much the same as I am, and that their interests are equally deserving of consideration, leads us to experience appropriate moral sentiments.

1.2.3 Kant

Apart from the Stoics, Kant is probably the best candidate for a philosopher to fill the ‘anti-emotion’ position supposedly overthrown by the pro-emotion consensus. He does not give an explicit theory of emotion (Nussbaum, 2001: 381). However, he is often interpreted as being particularly hostile to emotions, seeing them as external disruptive forces that take away our ability to deliberate rationally and which lead us to act irrationally.

Kant is particularly concerned with the motives from which we act, and with what we can be held responsible for. Moral agency, for Kant, is based in our ability to will according to the universal moral law. And this involves being motivated to act by the motive of duty, which is seen “as uniquely expressing an agent’s commitment to morality and thus as conveying a special moral worth to actions” (Denis, 2011). We discover the moral law through rational deliberation, not through feelings or natural inclinations.

In fact, according to many scholars, Kant thinks that while an action (e.g., helping someone in need) is morally worthy if performed from the motive of duty, the very same action would not be morally worthy if motivated by sympathy (Borges, 2004: 141). This is because we are responsible for our will, but not for our emotions, so only actions motivated by the former can have moral worth. Emotions are beyond the will (Borges, 2004: 141) and motivate us to act without the assent of reason, so they are undesirable influences (Baron, 1995: 194). So
virtue consists in the will asserting mastery over the ‘inclinations’ (emotions and other non-rational influences) (Nussbaum, 2001: 381).

However, there is considerable controversy over whether this is an appropriate interpretation of Kant’s views on emotions. Many Kant scholars argue that Kant had a more complex view, and allowed that emotions, properly shaped and regulated, play a crucial part in moral deliberation and motivation. For example, Denis notes that while Kant explicitly distinguished ‘true virtue’ from ‘assisting drives’ such as sympathy or benevolence, this seems to be at odds with what he says elsewhere about the importance of these ‘assisting drives’ in getting us to act morally, because of the “weakness of human nature” and the insufficient motivational force of the moral law (Denis, 2011). In fact, “Kant is explicit that feelings – especially pleasure (or satisfaction) and pain (or displeasure), both actual and anticipated – are essential to human moral motivation” (Denis, 2011). In other words, we have to be able to feel pleasure at doing our duty, or displeasure (aversion, or guilt) at failing to do our duty, in order to be properly motivated to act morally.

Moreover, Kant also recommends that certain of our emotions should be shaped and cultivated in order for us to achieve virtue. Baron (1995: 195-6) notes that his work “is full of suggestions of how we can shape our own affective responses,” including reading literature to “refine our sentiments.” We should, for example, ‘visit places where people suffer in order to awake our natural sympathy” (Borges, 2004: 153). This is because certain emotions are the foundation of “receptiveness to the concept of duty” (Denis, 2011), as well as helping us to act morally when the moral law is insufficient motivation.

It does appear, then, that Kant thought that at least some emotions, properly cultivated, improved our access to our reasons. However, he did not
think that they were reliable guides to what is morally right. Therefore, we should not allow them to determine action directly, but rather use our reason to decide whether it really is right to act: “sympathy could inform us about the existence of someone in need, but only reason can tell us if it morally correct to help this person in that situation” (Borges, 2004: 152). As Baron (1995: 200) notes, Kant does not think that emotions are all intrinsically bad, or that they ought to be eradicated: he writes that “natural inclinations, considered in themselves, are good, that is, not a matter of reproach, and it is not only futile to want to extirpate them, but to do so would also be harmful and blameworthy.” However, it would be bad to take our emotions as a sufficient guide to action, rather than “taming” them and acting only after rational consideration and thought about what is in accordance with the moral law (Borges, 2004: 153).

2. Historical Theories and the Lack of ‘Anti-Emotion Consensus’

The previous section considered only very brief overviews of some of the more important and influential historical theories of emotion. Nevertheless, even from this we can see that there was no ‘anti-emotion consensus’ in historical thought: historical theories did not all hold that emotions were entirely antithetical to rational deliberation and decision. Rather, there was a wide variety of views, some more hostile to the emotions than others. What is interesting is that many historical theories take the emotions to play important and even indispensable functions in practical rationality, and the particular functions they describe are substantially similar to those proposed by pro-emotion theorists.
Of the historical theories considered in the previous section, those that come closest to being ‘anti-emotion’ views are those of Kant and the Stoics. Yet the Stoics allowed that emotions are evaluative judgements of things that we are concerned about (even though we should not really be concerned about them); moreover, they thought that the sage would experience affective states (that we would call emotions) in relation to things that are genuinely good or bad. And Kant believed that certain emotions were useful and perhaps essential in moral deliberation and motivation, in ways which recall the pro-emotion view that emotions improve our access to our reasons: emotions help us to see where our duty lies, and also help us to be motivated to act on it.

Other major historical theorists had a much more sympathetic view of emotions, and some of the things they say about emotions could be attributed to the modern pro-emotion consensus. In particular, there is the idea (in Plato, Aristotle and Descartes) that emotions perceive the evaluative features of the world, especially those that relate to our concerns and our wellbeing. For these philosophers, emotions play the role of determining the salience of features of our choice situations, and discriminate between relevant and irrelevant considerations in deliberation. Additionally, there is the idea that the emotions supply the motivational force necessary for us to act on our reasons. Descartes, in particular, pre-empts the modern consensus’ view that emotions are necessary to allow us to act quickly and efficiently when there is insufficient time for rational deliberation. In the previous chapter I listed the various roles that emotions were said to play in practical rationality, according to the pro-emotion consensus. We can now see that historical theorists took emotions to play these very same roles.

So are there any common features of historical theories of emotion which place them at odds with the modern pro-emotion consensus? There are several
potential candidates. Firstly, most historical theories thought that emotion and reason were two different faculties. Secondly, they focused on the tendency of emotions to cause irrational behaviour if allowed to lead directly to action. Thirdly, and as a consequence, they believed that it was necessary for the emotions to be regulated by reason (with some notable exceptions, e.g. Hume). I do not think that any of these aspects of historical theories opposes them to the modern pro-emotion consensus, and that they are either trivial differences or not in fact points that the pro-emotion consensus must or should deny.

The first possible point of disagreement between (at least some) historical theories and the modern pro-emotion consensus is the idea that reason is one kind of mental faculty while emotion is another. Descartes, for example, thought that reason resided in the soul while emotions resided in the body; and Plato’s model of the tripartite soul took emotions to be a different part of the soul (the unruly horse) to reason (the charioteer). And in fact, others have interpreted the ‘traditional view’ as saying “that emotions are quite distinct from the processes of rational thinking and decision-making” (Sripada and Stich, 2004: 133), and have made the same assumption as Damasio, namely, that this separation is inconsistent with the idea that emotions are fundamental to rational decision-making. As we saw in the discussion of Descartes, however, this separation of reason and emotion into separate faculties does not entail (as Damasio thought) a view according to which emotions are nothing more than disruptive influences on rationality. It is possible to think of the emotions as playing a helpful, and even essential, role in practical rationality while also thinking that they are not part of the same faculty.

---

3 Although Descartes rejects talk of faculties and maintains that the soul is unified. Thanks to the examiner who brought this to my attention.
as reason. This is because the material that reason has to work with, especially as it concerns empirical matters or those affecting our wellbeing, is delivered via other faculties, including sensory perception and (on this view) the emotions.

Therefore, if the separate-faculties view in fact differentiated historical theories of emotion from the modern pro-emotion consensus, it would not be a very significant difference. The question would be whether emotions play the roles attributed to them, and whether they do so from a separate faculty, versus doing so within a common faculty, would not have any major implications. However, it’s not clear that the pro-emotion consensus is uniformly opposed to the separate-faculties view. Indeed, according to the Evolutionary Psychology view, emotions are a module (or modules), which are therefore separate from whatever constitutes the general faculty of reason. And the perceptual theorists mostly do not appear to make an assertion regarding this question. Damasio believes that emotions are located in particular structures of the brain, which are separate to the structures used in cognition. So we cannot use this as a point of contention between historical and modern theories.

We might instead try to locate a disagreement between the majority of historical theories and the modern pro-emotion consensus in the fact that most historical theories hold that emotions, left to their own devices, will tend to cause irrational behaviour and decisions. Those historical theorists whose views overlap with the pro-emotion consensus tend to say that while emotions can be helpful to decision-making, they very often – even usually – are not, and that it is a very bad thing to be unreflectively guided by one’s emotions in deciding what to do. Conversely, the modern pro-emotion consensus might be interpreted as denying that emotions commonly cause irrational behaviour, and/or that it is a good idea to be unreflectively guided by emotions in practical questions.
I do not think, though, that this is a desirable or necessary position for pro-emotion theorists to take. Firstly, it would be highly counterintuitive for pro-emotion theorists to assert that emotions do not commonly cause irrational behaviour, given all of the commonsense and experiential evidence to the contrary. That idea that emotions quite often do lead us to act irrationally seems obvious and hardly in need of defence. In fact, if this were the position of the pro-emotion consensus, we might agree with Sripada and Stich (2004: 133) that “the pendulum has swung too far” away from the traditional view. The burden of proof would seem to be on the pro-emotion consensus to explain why decisions driven by emotions are not actually irrational even when they appear to be.

However, I do not think that pro-emotion theorists are committed to the claim that emotionally-driven irrational behaviour does not occur (or even that it occurs only very rarely). None of them make this claim and many (as we will see in subsequent chapters) quite explicitly agree that emotions are a major contributor to irrational behaviour. Nor is it necessary for pro-emotion theories to make this claim. The pro-emotion view is quite compatible with the notion that emotions sometimes, even often, cause irrational behaviour.

Recall the roles which emotions are supposed to play in practical rationality according to the pro-emotion consensus: emotions improve our access to reasons, facilitate planning agency, remedy the deficiencies of decision-theoretic reasoning, and facilitate action. Roughly, emotions do this by signalling evaluative features and providing accompanying motivational force. The idea is that without them we would not be able to pick out those aspects of our situation that are relevant, we would not be able to construct a manageable set of options or decide efficiently between them, we would not be able to represent the importance of
future goods in current choices, and we would not be motivated to act on our decisions.

But what all of this suggests is not so much that emotions help us to make *correct* decisions, but rather that they help us to make *decisions*. They are fundamental to the *capacity* for rationality rather than the *fact* of rationality. What the pro-emotion consensus says is that without emotions we would be severely impaired in our ability to construct sets of options and make a choice from these sets, to be motivated to pursue a future good at the expense of a present one, or to act on a decision. If we are impaired in these ways then we certainly will be severely compromised in our ability to make rational decisions, but this is (mostly) because we will be severely compromised in our ability to make any decisions. The fact that emotions play all of these roles is quite compatible with the idea that most of the time, they help us to construct the *wrong* sets of options, lead us to make the *wrong* choice from among them, lead us to *inappropriately* pursue future goods at the expense of present ones, and motivate us to *act irrationally*.

Nevertheless, it does not sit well with the pro-emotion view to suppose that emotions *always* or *overwhelmingly* cause irrational decisions and behaviour. As noted in Chapter One, all of the pro-emotion views are evolutionary views in the sense that they claim that we have evolved to have emotions because they served a useful function. If emotions were inevitably harmful, and always led us to act wrongly, then it becomes less plausible to suppose that they are evolved adaptations to common problems. Of course, it is not *impossible* that emotions are evolved features nevertheless, because it could be that all of the alternatives actually realised were worse. But a pro-emotion theorist probably does have to be committed to the view, not just that emotions are needed for us to make decisions, but also that they are often enough a reliable guide to what we ought to do. But
then the disagreement between modern and traditional theories would, at best, come down to the question of what proportion of the time emotions are helpful rather than harmful, or whether they are helpful in a more important way than they are harmful. And this would seem to be an empirical question.

Finally, we might think that historical theories can be placed in opposition to the modern pro-emotion consensus because they tended to believe that emotions needed to be regulated or controlled by reason. If this is the point of disagreement between the two ‘camps’, then the pro-emotion consensus would have to deny that regulation by reason is desirable; alternatively, they might deny that regulation of emotions by reason is possible. Again, though, this would place pro-emotion theories in the position of having to defend a rather counterintuitive claim, since it appears to be common sense not only that we can regulate our emotions, but that we ought to do so. Indeed, many psychologists and cognitive scientists take the capacity for emotional regulation to be crucial in socialisation. As Gross and Thompson (2006: 3) write, “civilization is defined by coordinated social interchanges that require us to regulate how emotions are experienced and expressed…. Clearly, a great deal hinges on our ability to successfully regulate emotions.”

But again, the pro-emotion consensus need not (and should not) deny that regulation of our emotions by reason is desirable and necessary to ensure rational behaviour. Pro-emotion theorists tend to contrast their models of rational decision to the notion of ‘pure, unemotional’ decision-making, and they argue that creatures with emotions have an advantage over hypothetical unemotional creatures. However, as Scherer (1994: 128) notes, emotions also offer an advantage over a more automatic and primitive stimulus-response model according to which a given stimulus in the environment leads directly to action. Scherer argues that
“emotion decouples stimulus and response,” introducing a gap between appraisal of a stimulus and a reaction, while still preserving the automatic signalling of the stimulus mechanism and preparation of the organism for action. If emotions were to automatically lead from appraisal (stimulus) to action (response) they might still have an advantage over unemotional decision-making. But they would not have an advantage over a system in which an appraisal was made, but in which the organism then had the opportunity to intervene to prevent a response. And emotions are this latter kind of system. As Damasio (1994: 129) notes, there would otherwise be little point in our feeling emotions.

The pro-emotion view allows that this space exists between the operation of emotions in deliberative processes, and our making a decision. Emotions can signal important facts to us, and can strongly push us in the direction of certain actions, but there is also the possibility of our intervening in order to prevent the action, or even to force attention to other aspects of the situation. So at the very least, we can regulate emotions in the sense of not being guided by them in our actions. There is also the possibility of more direct regulation of emotions themselves: that is, in trying to modify occurrent emotions or train our emotional dispositions. And if emotions do cause irrational behaviour then it will often be desirable for us to use this capacity.

What we can see, then, is that the modern pro-emotion consensus is not an entirely new development which overturns a previously existing traditional ‘anti-emotion’ consensus. There was a great deal of variety in historical theories of emotions, including views on their helpfulness or harmfulness. Many historical theories have a lot in common with the modern pro-emotion consensus. In fact, there does not seem to be a way to place these theories in opposition to the pro-emotion view. There may be a difference of opinion on whether emotions are a
distinct faculty to that of reason or on the balance of helpful versus harmful functions of emotions, but the first is a rather vague disagreement and the second is a matter of degree.

We can also see that pro-emotion theorists should take seriously the historical insight that emotions can cause irrational behaviour if not regulated by reason. This insight is not incompatible with the pro-emotion view. So if the pro-emotion consensus is going to give us a satisfactory theory of the role of emotions in practical rationality, it will have to explain how and why emotions lead to irrational behaviour, in a way which is compatible with the explanation of the role of emotions in rational decision-making as well as the empirical evidence of emotionally driven irrational behaviour. It will also have to give an account of emotional regulation and how it can be used to avoid emotionally driven irrational behaviour.

In the rest of the thesis I explore whether various pro-emotion theories can provide the needed explanations, and what they can tell us about both emotionally driven irrational behaviour, and the regulation of emotions by reason. Before doing this, however, I wish to introduce two case studies of emotionally-driven irrational behaviour which can serve both to illustrate the theoretical concepts being discussed and to test the claims made by pro-emotion theories. I have chosen these case studies, firstly, because they are everyday examples of irrational behaviour which are clearly driven by emotion, and secondly, because they illustrate different ways in which emotionally-driven behaviour can be irrational. The first case study, opposition to recycled wastewater programmes, is a case of an emotion being inappropriate to its object in a way which leads to acting against our own (and society’s) best interests. The second, the phenomenon of emotional eating, is a case where the emotion itself need not be inappropriate, but the
resulting behaviour is. As we will see in subsequent chapters, these case studies will also help to bring out potential weaknesses in some of the pro-emotion theories.

3. Case Studies

3.1 Recycled Wastewater

We are living in a world in which the availability of fresh water is fast decreasing and in which supply problems are likely to become pressing in the not-too-distant future. One way to deal with this issue is to utilise recycled wastewater, including water runoff as well as treated sewage. Callaghan et al (2011) note that wastewater recycling is the “most cost-effective, environmentally sound and sustainable solution” of all the available alternatives. However, it is a solution which is not very widely used; in fact, “recycling makes only a small contribution to water supplies compared to its acknowledged potential” (Russell et al, 2009: 57). The reason for this is that the public is generally strongly opposed to recycled wastewater schemes. And the opposition is due to what is often called the ‘yuck factor’ – people are disgusted by the idea of consuming recycled wastewater, and will not accept its use (Menegaki et al, 2009: 286; Russell et al, 2009: 59).

There is a range of ways recycled wastewater could be used, but until negative public attitudes are overcome, these benefits will not be realised. As Russell et al (2009:58) point out, “public acceptance is crucial to the future of reuse”.

A good example of the opposition to recycled wastewater schemes, and the irrational nature of this opposition, is the failure of a referendum on the issue in Toowoomba, Queensland, in July 2006. At this time, dams were at 20% capacity,
and the scheme was proposed as a way of rescuing the town from a dire situation (Hurlimann, & Dolnicar, 2010:291-293). The Toowoomba council provided information about the unsustainable nature of current water usage, and explained that recycling wastewater would guarantee continued access to clean water. However, the referendum failed. Two years later, in July 2008, the dams had dropped to 11% capacity and severe water restrictions had to be implemented. Toowoomba residents had clearly been led by their disgust to oppose a course of action that would have averted a very dangerous situation.

The opposition to recycled wastewater can fairly be described as irrational, as it is acting in a way which makes one worse off. More importantly, however, the action is irrational because the emotional reaction is irrational. Although disgust at the notion of recycled wastewater is obviously intelligible, it is in fact not appropriate to its object – recycled wastewater is not actually disgusting. The water is clean – in fact, cleaner than ‘normal’ tap water – and has no ‘objective’ disgusting properties. Additionally, people are not consistent in their disgust reactions. In particular they are not disgusted by standard tap water which, as Menegaki et al (2009: 286) point out, is if anything more contaminated: “secondary-treated wastewater is commonly dumped into streams, rivers and lakes from which potable tap water is processed; if any contact with human excrement ‘contaminates’ water forever, most municipal tap water is already ‘contaminated’.

In fact, disgust at the notion of recycled wastewater has produced some obviously ridiculous results; for example:

In Fountain Valley, a group dubbed the Revolting Grandmas led opposition to the Orange County Groundwater Replenishment System, which is the largest wastewater reclamation plant in the world. Responding to opponents’ demands, engineers now pump highly treated
wastewater leaving the plant into an underground basin, where it filters through layers of sand and gravel before being piped to the homes and businesses that use it. Ironically, the water coming out of the basin isn’t as clean as the treated water going into it...during its trip through the natural filters it picks up trace elements and contaminants that must later be removed by the water utility. (Schmidt, 2008:525)

Nevertheless, the disgust reaction is not universal. Several cities have successful recycled wastewater schemes, including in England and the US. Singapore has a recycled wastewater scheme, NeWater, which was started in 2003 and which is now being expanded. Recycled wastewater is currently responsible for “up to 30% of Singapore’s water needs,” and the scheme is being expanded with the goal of eventually supplying 50% of water demand (Gallezo, 2011).

3.2 Emotional Eating

Emotional eating is a term used to refer to “the phenomenon of eating in order to cope with emotions” (Spoor et al, 2007:368). More specifically, it refers to the tendency of some individuals to eat more in response to negative emotions. Canetti et al state that “emotional eating is most often precipitated by negative emotions such as anger, depression, boredom, anxiety and loneliness and often bears an episodic relationship to stressful periods of life” (Canetti et al, 2002:160). Whilst common amongst ‘normal’ individuals, it is certainly not universal. In fact emotional eating is more common amongst obese individuals, people with eating disorders, and those who are ‘restrained eaters’ (i.e., those who are trying to restrict their food intake) (Spoor et al, 2007:368).
While mild emotional eating may be seen as a harmless way to cope with negative emotion, the association with obesity and eating disorders shows that the consequences for a person’s health and well-being may be seriously detrimental (Macht & Simons, 2011:291). Foods chosen by emotional eaters tend to be more calorie-dense and unhealthier. Thus, emotional eating can be seen as an irrational behaviour, one which is not in the person’s best interests. In fact, we can go further than this and say that this sort of behaviour is also irrational in the sense that it goes against the person’s all-things-considered judgements. Most emotional eaters are not engaging in this behaviour because they do not care about the consequences for their health and wellbeing; rather, they themselves see the behaviour as negative and harmful.

Emotional eating is in fact a behaviour which is *maladaptive* rather than adaptive. While the negative emotions that cause the eating need not be inappropriate, eating is not an appropriate *response* to negative emotions. It does not even make much sense on evolutionary grounds: Heatherton et al point out that “on strictly physiological grounds, emotional distress ought to suppress eating, because the physiological reaction to stress mimics the internal sensations associated with feeding-induced satiety” (Heatherton et al, 1991: 138). Thus, emotional eating has been generally viewed as an “inapt” reaction to emotional distress (ibid) and a “maladaptive strategy” (Macht & Simons, 2011:291). It appears to be associated with a more generally maladaptive approach to emotional regulation: for example, some studies have found that individuals who engage in emotional eating also have a more general tendency to use “emotion-oriented coping and avoidance distraction” (that is, trying to ease the negative emotions associated with a problem, or distracting oneself, rather than addressing the problem causing the negative emotions) (Spoor et al, 2007:373).
Chapter Three: Strategic & Evolutionary Psychology

Theories Revisited.

In the first chapter, I listed the various ways in which emotions might play helpful roles in practical rationality, and divided members of the ‘pro-emotion consensus’ into different categories – strategic theories, Evolutionary Psychology theories, perceptual theories, and neo-Jamesian theories – according to which of the roles they took emotions to play, and the ways in which they took emotions to play those roles. In the second chapter, I argued that there is no historical ‘anti-emotion consensus’ which the current pro-emotion consensus overturns; rather, many historical theorists had views on the helpful function of emotions in practical rationality which are to all intents and purposes the same as the pro-emotion views. Moreover, the major insight which was widely held historically – namely, the view that emotions are prone to cause irrational behaviour unless trained and regulated by reason – is one which is compatible with the fundamental commitments of pro-emotion theories, and rather clearly correct.

But just because the ‘historical insight’ is compatible with the general pro-emotion thesis – namely, that emotions are necessary for practical rationality – this does not mean that all of the specific pro-emotion theories are able to accommodate it. The historical insight may be incompatible either with the particular roles included in a theory, or the mechanism by which emotions play those roles. Additionally, there may be other, independent, reasons why a particular pro-emotion theory is problematic or unsatisfactory, including a failure to cohere with empirical evidence, or not giving a broad enough account of the role of emotions in practical rationality. A satisfactory pro-emotion theory, on the
other hand, would be not only interesting in its own right but would lead to potentially useful claims and predictions about how we might manage both our own practical rationality and that of others.

For this, however, we would need a complete theory – that is, one which can in theory account for all of the roles emotions play in practical rationality. This theory would also have to be able to give an explanation of the apparent tendency of emotions to cause irrational behaviour, and allow for the possibility of emotional training and regulation of a kind that is generally recognised to exist. Furthermore, the account of irrational behaviour and emotional regulation must be compatible with the claims made about the particular functions of emotions on that theory, and the mechanisms underlying those functions. While a satisfactory account of this kind obviously does not guarantee the correctness of the theory, it would have a great deal of intuitive plausibility, and would be able to generate testable predictions about when emotions will be helpful or harmful, and what will and won’t work in terms of managing them. Thus, it is important to further explore the different kinds of pro-emotion theory discussed in Chapter One to see if any are able to deliver a satisfactory account of this kind.

In this chapter I consider whether strategic theories or Evolutionary Psychology theories can meet these challenges. I argue that neither is satisfactory. Strategic theories are incomplete and do not give a full account of the role of emotions in practical rationality, as well as having serious internal problems. Evolutionary Psychology theories can give an account of emotionally-driven irrational behaviour; however, this account is problematic, since it is not entirely consistent with the observed evidence, and moreover causes tensions with the underlying commitment to the central pro-emotion thesis. Additionally, the
potential for emotional training and regulation is overly limited by a commitment
to the claims of Evolutionary Psychology.

I also consider Prinz’s neo-Jamesian perceptual theory in this chapter, while leaving discussion of Damasio’s theory until Chapter Four. As noted in Chapter One, both Prinz and Damasio are, broadly speaking, perceptual theorists, but they differ from other perceptual theorists in their commitment to neo-Jamesianism. The reason for considering Prinz in this chapter is that the central objection to Evolutionary Psychology theories – namely, the commitment to the modularity thesis concerning emotions – will also be an objection to Prinz’s neo-Jamesian theory, since he too accepts the modularity thesis. Damasio, on the other hand, is not committed to modularity, and is more properly discussed along with other perceptual theories.

1. Strategic Theories

As discussed in Chapter One, strategic theories hold that emotions are essential to practical rationality because they serve as strategic commitment devices. They are able to remedy the deficiencies of practical rationality by motivating us to act in ways which are ‘rationally irrational’, in the sense that the actions are not utility-maximising in that particular choice situation, but the disposition to act in this way is utility-maximising overall. This allows us to cooperate in Prisoner’s Dilemma situations, make credible promises or threats, and commit to relationships, both because of the behavioural dispositions themselves and because our emotional dispositions are signalled to others and allow them to predict our actions. It also allows us to represent future goods in present choice situations, overcoming the time-inconsistent preference reversal
characteristic of hyperbolic discounting, and therefore aids in constructing goal hierarchies and achieving longer-term goals.

When asking whether strategic theories can properly allow for the observation that emotions often cause irrational behaviour, it is important to be aware that this condition is *trivially* met by the fact that the supposed role of emotions is to cause ‘rational irrationality’. According to decision-theoretic accounts of rationality, it is *rational* for us to defect on agreements, fail to carry out threats, or walk out on relationships, whenever there is more utility/benefit in doing so. According to the strategic view, the function of emotions is not to cause us to act rationally in these situations, but rather to act *irrationally* – i.e., not to choose the utility-maximising option. What is rational is the *disposition* to act in these ways, since the resulting overall long-term utility is greater than choosing rationally in each individual choice situation. So, in a sense, the strategic theory not only *allows* the possibility that emotions will often cause irrational behaviour, it actually *relies* on it.

Nevertheless (and disputes about the adequacy of the decision-theoretic notion of rationality aside) this is obviously not an accommodation of the historical and commonsense insight that emotions often lead to irrational behaviour. The claim is, rather, that emotions frequently cause us to act in ways that are irrational *simply*. These sorts of actions are not ‘rationally irrational’: they do not serve our long-term interests or allow us to gain the benefits of cooperation with others. The case studies introduced in the previous chapter are such examples: it is not in our long-term interests to vote against recycled wastewater schemes, and nor is it rational to be disposed to engage in emotional eating. We must therefore ask whether strategic theories can explain this kind of emotionally-driven irrational behaviour.
In fact, strategic theorists have acknowledged that emotional dispositions might cause us to act in ways that are irrational in this broader sense. This can happen if the disposition causes us to act in ‘locally irrational’ ways (that do not maximise utility in the individual choice situation) and this is not compensated for by longer-term benefits. For example, it may be generally rational for people have a vengeful disposition, since this serves as a deterrent to others. But in some contexts such a disposition can be disastrous in its long-term consequences as well as suboptimal in individual strategic decisions. Frank (1988: 1-4) gives the example of the Hatfield-McCoy feud, a long-running dispute between two clans in West Virginia and Kentucky in the late 19th century. Following a complicated history of disputes, the Hatfields were attacked by the McCoys, and although they lost many family members, they retaliated with an attack of their own. The McCoys struck back, and the cycle continued for more than ten years, destroying the lives of many people and achieving nothing of value. Clearly it would have been better if at least one of the parties had been less strongly disposed towards vengeance.

Clearly, strategic theorists do acknowledge that emotions can cause behaviour that is simply irrational. Their explanation is that a helpful disposition, like a good policy or rule, cannot be fine-tuned to every possible circumstance, but rather has to be designed for the sort of situations we are most likely to encounter. This entails that the disposition may have undesirable or unexpected results in uncommon situations. As an analogy, consider the desirability of wearing a seatbelt. Most of the time, wearing a seatbelt will prevent severe injuries and save lives. However, there are rare kinds of accidents in which wearing a seatbelt might make us worse off. Nevertheless, we ought to have a policy of wearing a seatbelt, because the expected benefits vastly outweigh the risks. In the same way, the
expected benefits of emotional dispositions are greater than the expected tendency to cause harm (Elster, 2000: 46). Although we might end up with long-running feuds that ruin the lives of generations of family members, most of the time we will be able to successfully deter other people from harming us and we can all get on with our lives in peace.

So we can see that strategic theorists explain the phenomenon of emotionally-driven irrational behaviour in a way which is consistent with their claims about the function of emotions in practical rationality. However, there are several problems facing strategic theories of emotion, if it is taken to be a complete account of emotion. Firstly, they do not account for all instances, or even the majority of instances, of emotionally driven irrational behaviour. Secondly, in admitting that emotions can cause irrational behaviour in (supposedly) ‘non-standard’ circumstances, the idea that emotional dispositions are essential for rationality, or even mostly helpful, is weakened. Thirdly, the resources this account has to explain the possibility of emotional training and regulation are somewhat limited.

1.1 The insufficiency of explanations for emotionally driven irrational behaviour

We saw above that strategic theorists acknowledge the possibility that emotional dispositions will lead to (irrationally) irrational behaviour in circumstances that the dispositions were not ‘designed’ to handle. This is fine, as far as it goes. But this is only one kind of emotionally-driven irrational behaviour: namely, only the kind of irrationally-irrational behaviour that can occur in
There are many other ways in which emotions can cause irrational behaviour – and in fact, there are many other ways in which emotions function in practical rationality apart from their strategic role. Thus, strategic theories are incomplete: they do not give us a full account of the role of emotions in practical rationality.

Even if we accept Frank’s account of supposedly strategic emotions such as vengeance/anger, guilt, or romantic love (and it is debatable whether we should do so), there are still plenty of emotions which do not have a plausible strategic explanation of this kind. The recycled-wastewater case study is one such example: it is hard to see how disgust at recycled wastewater (which causes opposition to wastewater-recycling schemes) is at all strategic. It does not function as a commitment device to ‘protect’ us from a paradox of rationality. Nor does it help me to avoid short-term maximising to achieve longer-term benefits. We can imagine a sort of strategic argument that goes: yes, it would be ‘locally’ rational (i.e., utility-maximising) to consume ‘contaminated’ water when thirsty, but if disgust forces me to look elsewhere I will be better off in the long term because I will avoid parasites and illness. But this isn’t really very plausible. In an evolutionary environment of unpredictable water supply, those who didn’t drink when they were thirsty, or who rejected clean water supplies wholesale, would be more likely to die of thirst; those who used rational means of deciding whether water was safe to drink would be much more likely to survive and thrive.

Similarly, the phenomenon of ‘emotional eating’ does not seem like something that can be given a plausible strategic explanation. As noted in introducing the case study in Chapter Two, the phenomenon is generally seen as

---

*Strategic theorists are dealing with one kind of rationality, namely whether an emotion is instrumentally/strategically rational – not whether an emotion is appropriate to its object. For further discussion, refer to pages 16–19; 24–27.*
maladaptive and at odds with the physiological effects of negative emotions. Additionally, the reason why emotional eating is seen as an irrational behaviour is that it is harmful in the long term, being associated with obesity, overweight and eating disorders. An emotional disposition which causes us to eat in response to negative emotions cannot be plausibly seen as a commitment device which helps us to behave in a rationally irrational way. The disposition to eat in response to negative emotion is an irrational disposition to have, and it is arguably not rational in the individual choice situation either, to the extent that it represents a maladaptive coping strategy of trying to deal with the symptom (the negative emotion) rather than the problem.

The phenomenon of emotional eating also alerts us to an additional gap in the strategic account. In the case of emotional eating, the function of the emotional disposition, if anything, is to inappropriately prioritise short-term goals (alleviating emotional distress) over long-term plans (staying healthy, maintaining a good diet). It is therefore the opposite of a commitment device: here, emotions are serving to undermine planning agency rather than facilitating it. As we saw, it was widely held by historical theorists that emotion is the primary cause of weakness of will, while strategic theorists “turn the idea on its head, by arguing that emotions are solutions rather than problems” (Elster, 2000: 45). But in doing so, they have ignored the fact that, whatever solutions they present, emotions are also a problem: even if they do often facilitate planning agency, they will also often undermine it. In fact, critics such as Elster argue that many common ‘self-management’ practices, such as self-binding (constraining ourselves to make a particular option impossible to choose), arise because of the need to protect plans made with a ‘cooler head’ from being derailed by the temporary influence of emotions (Elster, 2000:7).
Strategic theorists are therefore not giving a complete account of the role of emotions in practical rationality. I have mentioned two of the gaps: firstly, there are emotions that are non-strategic in that they cannot plausibly be interpreted as any kind of commitment devices; nevertheless, they have a role in both rational and irrational decisions. Secondly, there is the fact that emotions undermine planning agency and commitment as well as facilitating it. Now, the incompleteness of the strategic theory is not by itself an objection to the central claim that emotions are commitment devices. But it does mean that strategic theories are of limited usefulness in understanding the role of emotions in rational decision-making, since they only tell us, at best, part of the story.

Additionally, there are some strong objections to the idea that emotions, even when functioning as commitment devices, are generally helpful to practical rationality. As we saw above, strategic theorists acknowledge that emotional dispositions may be harmful (irrationally irrational) in some circumstances, but they would argue that these circumstances must be 'nonstandard', that is, uncommon or trivial compared to the kinds of situations the dispositions evolved to cope with. Strategic theorists (along with Evolutionary Psychologists) make the point that if emotions were entirely or even mostly detrimental, it is unlikely that we would have them. We might disagree with this assertion on the grounds that the best we can say is that our existing emotional dispositions have proved to be better than whatever alternatives happened to be realised in our evolutionary history. It may be that unemotional creatures, or creatures with vastly different emotional dispositions, have never been thrown up by the genetic lottery. Nevertheless, even if we accept the view that emotional dispositions evolved because of their strategic usefulness on balance, we do not have to think that they
are still useful; this creates a potential problem for the idea that the strategic theory is in fact a pro-emotion position.

1.2 Is the strategic theory a pro-emotion view?

As we saw above, strategic theorists attempt to reconcile their claim that emotions are strategically helpful with the fact of emotionally-driven irrational behaviour by arguing that emotional dispositions are more often helpful than harmful. The idea is that emotional dispositions evolved to handle the kind of strategic problems that were most commonly faced by our evolutionary ancestors, so that they are mostly helpful. But in any particular environment, it will therefore be a contingent matter how helpful or harmful a given emotional disposition will be. As Frijda (1994:117) writes, “the usefulness of a provision depends on the proportion of hits to misses and to false alarms, and on the costs of the alarms”. And this means that if our environment is sufficiently different from that in which our emotional dispositions evolved, the balance of helpfulness over harmfulness is likely to change. As we have seen it can be rational to be disposed to retaliate if this acts to deter people, but if we are in an environment in which everyone else is both (a) also disposed to retaliate, and (b) prone to ignore my disposition to retaliate, there is a great danger of disaster.

This does not mean that the strategic theory is false. But it does cast doubt upon the notion that it is a pro-emotion view. The best that can be said is that emotional dispositions had a useful role in facilitating commitment and planning agency in the environment in which we evolved. However, it will be entirely a contingent matter whether they still play a generally helpful role in our current environment. In fact, it would be possible that, in our current environment, our evolved emotional dispositions are almost entirely harmful and tend to undermine
commitment or planning agency much more often than they facilitate it. In this case, the strategic claim could be used as the foundations of an *anti-emotion* view: namely, that emotion is an evolutionary vestige, like the appendix, which probably used to perform a useful function but which is now entirely detrimental to rational choice to the extent that they do anything.

Another problem is that the strategic theory has been argued in relatively broad terms: for example, as saying that a disposition to become enraged when wronged (and hence motivated to carry out threats or acts of vengeance) is a rational disposition to have. However, the story must at the very least be more complicated than that. The ‘environment’ in which our emotional dispositions operate, and operated in our evolutionary history, includes other people, and their emotional dispositions; this is precisely what makes these emotional dispositions *strategic*. And whether a particular disposition is strategically useful depends, in particular, on the prevalence of this disposition in the general population. In many cases the disposition will be useful if (and *only* if) many others also have this disposition; in other cases the disposition will be useful if (and only if) others *don’t* have it; in still other cases it may be a *combination* of various dispositions which causes a problem.

This phenomenon – often referred to as frequency-dependence – is well known in the literature on strategic interaction and on evolution more generally. For example, it is widely understood that, in a population of non-co-operators, a sub-group of co-operators will, if they find each other, gain greater benefits. On the other hand, if everyone is a co-operator, this creates a target-rich environment for any non-co-operator who happens along: the non-co-operator can take advantage of a population of suckers. At the population level, what we would expect to see is therefore some sort of equilibrium between co-operators and non-
co-operators: if the level of co-operators rises, then non-co-operators move in to take advantage and drive the level down; if the level of non-co-operators rises then the payoff from being a non-co-operator is lowered (because there are fewer co-operators to exploit) and the level of co-operators will rise. Added to this is the question of co-operator-detecting mechanisms (including being able to distinguish genuine signalling of emotional dispositions from faking), and the costs of having these, which will depend on the proportion of non-co-operators.

This phenomenon of frequency-dependence creates three potential objections to the strategic theory. The first is that of explaining how the emotional dispositions arose in the first place in our evolutionary history. The second is that frequency-dependence makes it difficult to argue that a given emotional disposition is rational for anyone to have, or that it will be more often helpful than harmful in any context. The third is that when we take this phenomenon seriously, it becomes harder to see why we should have evolved emotional dispositions rather than being able to adopt policies or decision rules which are more complex and flexible.

With regard to the first point, there is a problem in understanding how anyone came to have these emotional dispositions in the first place, even if they have long-term benefits. The point of the strategic theory is that emotional dispositions cause us to act against our immediate interests in order to secure longer-term benefits. But as Elster (2000: 46) points out, “natural selection…does not always wait for the long term. Strategies of the kind ‘one step backward, two steps forward’ are not in its repertoire.” Nevertheless, it might be possible to give a satisfactory account: for example, it might be pointed out that the objection assumes that creatures with emotional dispositions arose in a population of creatures without them. Maybe, however, the emotional dispositions were there
all along, and have proved superior to any other evolutionary strategies that arose later.

The second objection is potentially stronger, but we must first be clear whether the strategic theory holds that emotional dispositions are stable features of an individual which are hardwired in by evolution, or that they are to some extent chosen or shaped through the individual’s interactions with their environment. If the former, then even in our evolutionary history, we cannot always say that a given disposition was rational for an individual to have had, since they may have been born a co-operator into a society overburdened with co-operators. But if emotional dispositions are learned or shaped, there is a different problem, since what disposition I come to have in interacting with others might not remain rational if others’ dispositions change. As Elster (2000: 48) points out, just because there is a theoretical equilibrium between (say) co-operators and non-co-operators in a population, this does not mean that a society will find this equilibrium and stay there. It could be instead that the proportion of co-operators to non-co-operators constantly cycles up and down around this equilibrium, and “the magnitude of the fluctuations or the speed of the correction mechanism” (Elster, 2000: 48) are open questions. A person whose disposition was shaped at one extreme of the cycle could find themselves badly off most of the time.

The strategic theorist might argue that this objection envisages only one way in which frequency-dependence might play out: namely, the idea that the population is ‘polymorphic’ – there is a mix of strategies instantiated in the population, but each individual has one fixed strategy. There is another possibility, however: namely, that all individuals in the population have the same ‘mixed strategy’ (Griffiths, 1997: 126). However, in the case of emotional dispositions, this ‘mixed strategy’ would be rather complex. It also may not be much of a
solution to the problem: again, if the environment changes, there is no reason why a particular disposition will continue to be on balance helpful, even if it comprises a mixed strategy. And emotional dispositions are quite difficult to change. In fact, it would appear to be far better for individuals to adopt general policies and rules (for example, Tit for Tat) rather than emotional dispositions, since these would be just as beneficial, but also easier to change if the environment does.

1.3 Emotional Training and Regulation on the Strategic Account

In the previous section I considered the question of whether strategic theories could explain irrational behaviour in a way that was compatible with the central claim to being a pro-emotion theory. In this section I look at the question of emotional training and regulation. That is, what resources to strategic theories have to explain the phenomenon of training of emotional dispositions, and our ability to regulate emotions (including emotional expression but also emotional experience)?

Strategic theories need not have difficulty with the general notion of emotional training; I noted above that strategic emotional dispositions might be formed and shaped by interactions with the environment. Thus, a person’s emotional dispositions might be influenced over time by their parents’ dispositions and training, as well as their experiences. This would be the case even if they thought that emotional dispositions were innate in some way: that we are born with certain dispositions or with strong tendencies towards certain dispositions; these could at least be modified or shaped to some extent by experiences and interactions.

There is a potential problem, though, with the phenomenon of emotional regulation on the strategic account. In order for the strategic account to work, it
has to hold that emotions (a) lead us to act, (b) in ways that are ‘locally’ irrational (i.e., not maximising in the particular choice situation), but ‘globally’ rational. This means that the way in which emotions play their role, according to the strategic theory, depends on their being relatively impervious to attempts to regulate or control them. Firstly, if we were able to stop our emotions from causing their characteristic behaviours, then emotions would not be able to play their strategic function. If I could stop the prospect of guilt from leading me to keep my promise, then there would not be much point in my being disposed to feel guilt; I would behave just like a purely rational agent. Similarly, if I could stop myself from carrying out my threats because I had rationally calculated that the costs outweigh the benefits, then a disposition to feel angry would not fulfil its purpose. So, strategic theories must hold that we have only very limited abilities to intervene in emotionally-driven behaviour.

A similar point can be made about emotional regulation which aims to change the experience of emotion, either in terms of trying to suppress or enhance an occurrent emotion, or to change our emotional dispositions. If it were possible to do this whenever we realised that our behaviour was irrational, then emotional dispositions would not be able to properly perform their function. Of course, strategic theorists can point out that it is quite hard to get rid of an emotion we are feeling, or to change our emotional dispositions. But it is not always hard; in fact, we engage in regulation of our occurrent emotions quite frequently in our day-to-day lives and often do not even notice what we are doing. In this case, we could override strategic emotional dispositions much of the time, and they would do us no good. Strategic theorists therefore have to show that emotional regulation is sufficiently difficult, as a matter of fact, to allow for strategic emotional dispositions to perform their functions.
Overall, there are several significant counts against strategic theories as offering complete accounts of the supposedly helpful roles of emotions in practical rationality. Firstly, strategic theories are incomplete, since they do not consider many of the roles emotions play in generating both rational and irrational behaviour. Moreover, strategic accounts face some difficulties in maintaining their central pro-emotion claim in the face of the fact that the desirability of emotional dispositions will be crucially dependent on the environment. It will be purely a contingent matter whether emotions are helpful or harmful, on this view, and given that the balance of benefits over costs can very easily shift, it would seem that emotions are a more risky proposition than is warranted from the point of view of practical rationality. Finally, while these theories allow for the possibility of emotional training, they do not have room for a sufficiently robust account of emotional regulation.

2. Evolutionary Psychology

As noted in Chapter One, all pro-emotion accounts are to some extent evolutionary accounts, in that they hold that emotions evolved and became fixed in our psychological makeup because of the functions they played in facilitating practical rationality. However, what I am calling Evolutionary Psychology accounts are also committed to several other claims, most importantly the modularity thesis that the mind is made up of domain-specific modules evolved to solve specific problems. On the Evolutionary Psychology account, emotions are modules: they are localized in the brain, subject to specific breakdowns, mandatory, fast, shallow, inaccessible to and informationally encapsulated from other levels of processing, ontogenetically determined and domain specific. This
allows them to play several roles in practical rationality: they signal relevant features of the environment and hence improve our access to our reasons; they enable quick and efficient decision-making and sorting of our options, thus remedying deficiencies in decision-theoretic rationality; and they facilitate appropriate action, since emotions have evolved to have behavioural outputs that are generally appropriate to the situations in which they are elicited.

2.1 Evolutionary Psychology’s explanation of emotionally-driven irrational behaviour

Evolutionary Psychology theories, unlike strategic theories, are intended as a complete account of emotions and their functions in rational decision-making. However, the explanation of how emotions cause irrational behaviour is much the same as that given by the strategic theory. Evolutionary Psychology theories, like strategic theories, hold that emotions evolved to cope with the sorts of situations most commonly encountered by our ancestors in the Environment of Evolutionary Adaptedness (EEA). However, this does not mean that an emotion will function ‘correctly’ in all possible circumstances: it may be maladaptive in different environments. Therefore, Evolutionary Psychology explains emotionally-driven irrational behaviour by claiming that while the emotion module was adaptive in the environment in which it evolved, either (a) changes in the environment have caused it to become maladaptive, or (b) the generally adaptive nature of the emotion has led to some maladaptive behaviours as a side effect. As Tooby and Cosmides point out,

if a trait is an adaptation, it follows by definition that it conferred a fitness benefit in the EEA; but nothing follows from that about the trait’s current adaptiveness... [Indeed,] the odds are massively against the whole
package of input and interaction working out in the way it would need to continue to be adaptive under conditions as different from the EEA as are our current conditions. (Jones, 2006:17)

While this is the same argument as offered by strategic theorists, the fact that the scope of the Evolutionary Psychology theory is broader allows it to give an explanation of a wider range of emotionally-driven irrational behaviour. Thus the irrational behaviour in the case studies, for example, can be explained as the function of previously adaptive modules that are inappropriate to the current environment. With regard to recycled wastewater, they can point out that the principal function of disgust is to protect us against contamination (Rozin & Fallon, 1987; Oaten, Stevenson, & Case, 2009; Curtis, De Barra, & Aunger, 2011; Schaller & Park, 2011): in our evolutionary history, it would have been useful to be disgusted by substances that were known have come into contact with human waste, even if they appeared clean (Prinz, 2004: 154). In the current situation, the water known to have once come into contact with human waste is actually perfectly clean and safe to drink, and the disgust is both inappropriate to its object and harmful to our interests. But we are still disgusted because of the operation of an emotional module that evolved to protect our welfare in very different circumstances.

An initial objection to this explanation is that the phenomenon of disgust at substances known to have been in contact with human waste is by no means universal. In the early 19th Century, the population of London drank untreated water – downstream from sewage plants (Snow, 1855). Public response to Singapore’s NeWater initiative has been largely positive, and free of the disgust response or ‘yuck’ factor (Gazello, 2011). There are also numerous other examples
of our disgust response being flexible, such as medical students and forensic pathologists who work with many ‘disgusting’ objects, yet find that their disgust response becomes desensitised with time. This might be thought to cast doubt on the idea that there is an evolved ‘disgust’ module that is wired to pick up on a history of contamination, and which continues to operate regardless of changes in the environment.

In fact, though, Evolutionary Psychology theories can account for cultural differences and the influence of desensitising learning experiences in disgust reactions. The disgust module may not have a pre-existing set of eliciting conditions, but instead may be set up to make it extremely easy for us to learn disgust reactions from others, with some disgust reactions being much easier to acquire than others: they can say that people are hardwired to learn disgust reactions taught by their culture, and that certain kinds of disgust reactions are much easier to acquire than others. Because the things which are disgusting (i.e., prone to cause illness or contagion) can vary, as Griffiths (1997:89) points out, “it might well have been advantageous for [modules] to be linked to some mechanism which can interpret the broad ecological categories of danger, novelty, and so forth, in the light of local conditions.”

In the case of ‘emotional eating’, Evolutionary Psychology theories may argue that the tendency to eat in response to negative emotion was generally beneficial in the EEA. For example, it might be suggested that emotional distress generally signals a high risk of adverse circumstances in the near future, and that the eating of as many energy-dense foods as can be found is a good way to prepare for the metabolic demands of this adversity. Alternatively, since ‘emotional distress’ is a rather broad category, they may invoke a simpler explanation: we crave energy-dense foods when upset because we are trying to maintain
homeostasis or even to increase levels of the serotonin precursor tryptophan. Macht & Simon (2011:286) note that high-carbohydrate foods raise blood levels of tryptophan which “leads to increased activity of the serotonergic brain systems.” And it is obvious that a major difference between modern industrialised society and the EEA has to do with the overabundance of energy-dense foods in the former. While we may have needed strong motivations to pursue scarce energy sources in the EEA, nowadays this same disposition causes us to act maladaptively.

2.2 Why don’t these explanations work?

Evolutionary Psychology theories can therefore explain a wide range of emotionally-driven irrational behaviour as the operation of emotional modules which evolved to cope with one set of circumstances, and are now operating maladaptively in a very different environment with a very different set of problems. Nevertheless, I do not believe that these explanations are satisfactory. One problem is the same as for strategic theories – namely, that there is a strong risk of having to give up the central pro-emotion claim that emotions are necessary or helpful to practical rationality, since this will be entirely contingent on the environment. The problem isn’t just that emotions may operate maladaptively in different environments, but that Evolutionary Psychology accounts rely on instances of emotion in order to determine their rationality. What I am suggesting, and Perceptual theorists are suggesting, is much more general: it is rational to have particular emotional dispositions (or types), even though specific instances of those emotions (tokens) may be irrational. For example, anger can alter rationality based on the circumstances in which it arises; even though
particular instances might be maladaptive it is still a rational disposition to have overall.

Apart from these points, there are two main problems with the Evolutionary Psychology account. The first is that the specific explanations (that is, for specific instances of emotions causing irrational behaviour) are not always compatible with the observed evidence. The second is that while they do allow for culture and learning to shape our emotional responses, they are committed to claiming that there are strong constraints on the operation of these factors. This also creates problems for an account of emotional training and regulation, as will be discussed in the next section.

There is a general objection to Evolutionary Psychology explanations, which also applies to the explanations of emotionally-driven irrational behaviour: namely that they might be ‘just so’ stories – plausible-seeming accounts which are nevertheless entirely inaccurate. As Griffiths notes, “inferring the problem from the solution” (i.e., inferring the evolutionary forces shaping an ‘adaptation’ from the way it happens to be) is problematic, because “an adaptive explanation can be very plausible but false.” (1997:110). To show that they are not merely offering a ‘just so’ story, Evolutionary Psychology theories need to show that their explanations of irrational behaviour are consistent with observed evidence (and to show that they can generate testable predictions that are then borne out).

And in fact, Evolutionary Psychology explanations of irrational behaviour are often at odds with the observed evidence, as can be seen by the explanations given for the behaviours in the case studies. In the case of recycled wastewater, the claim was that it would have been adaptive for us to be disgusted by (or learn to be disgusted by) anything known to have come into contact with human or animal waste products. But, as noted in Chapter One, people who are disgusted by
recycled wastewater are not disgusted by ordinary tap-water, even when they are
told that it has a similar history of contamination. This inconsistency cannot be
explained if the disgust reaction is a module that operates in the way claimed.

In the case of emotional eating, the explanation similarly appears not to be
supported by the evidence, which suggests that physiological mechanisms are not
in fact major factors. In fact, a more adaptive response to negative emotion, even
in the EEA, would apparently have been to refrain from eating. Heatherton et al
(1991:138) point out that “on strictly physiological grounds, emotional distress
ought to suppress eating, because the physiological reaction to stress mimics the
internal sensations associated with feeding-induced satiety”. Spoor et al (2007:368)
also write that “loss of appetite and decrease of food intake have been considered
natural physiological responses to negative emotions.” If we were to take an
Evolutionary Psychology viewpoint, we would have to admit that a stress
response that mimics satiety must have been adaptive – perhaps because it
prevents us from eating before we might be forced to run away or fight. If this is
the case, then eating as a response to emotional distress is not an initially adaptive
response that has become maladaptive in a different environment: rather, it has
always been maladaptive. This also fits with the evidence that emotional eating is
associated with a more general pattern of maladaptive coping which aims at
alleviating the distress rather than tackling the problem causing it (Spoor et al,
2007:373).

What about the hypothesis that emotional eating is simply an attempt to
induce homeostasis or serotonergic effects? In fact, this is not strongly supported
by the evidence either. Macht and Simon (2011:288) argue that the ‘serotonergic’
effects of high-carbohydrate foods are likely to be minimal in real life diets (as
opposed to specially constructed food in laboratory studies) and that “the
serotonin hypothesis appears to play a minor role, at least in emotional eating”. They also point out that all physiological explanations for emotional eating suffer from a significant problem:

Nutrient-dependent emotional changes need time. They can only occur after food is digested, its components are absorbed into the bloodstream and transported to the brain…. Delayed effects may be useful for coping with chronic stress, but not for responding to negative emotions arising from unpredictable stimuli as they often occur in daily life. (2011:289).

Evolutionary Psychology theorists might reply that, as pointed out above, they do allow that our emotional responses can be shaped by cultural and other environmental influences, so that they do not need to argue that our actual emotional responses or behaviours were adaptive in the EEA. Rather, they only need to argue that our propensity to learn such dispositions was adaptive. So (they might claim) it’s not a problem that inconsistent disgust reactions, or emotional eating, would not have been adaptive responses to problems in the EEA. However, I do not believe that this response is adequate: either it is insufficient to give a proper account of the actual phenomena of emotionally-driven irrational behaviour, or tensions are created with the claims about the functions played by emotions in practical rationality.

As we have seen Evolutionary Psychology theorists agree that there is some ‘plasticity’ in emotional response: the things in the environment that elicit a particular response might differ between cultures, between individuals, and may even change for one individual over time. Nevertheless, there must still be significant constraints on how much variation there can be. As Faucher and Tappyloet (2008:118-119) write,
First, they think that some appraisal mechanisms are set at birth to evaluate a small group of stimuli in a certain way...Second, the set of things one might be afraid of or disgusted by is seen as somewhat constrained. Certain things could not but be appraised in a certain way (by normal subjects). Third, not only are there things that are universal elicitors of particular emotions, but there are also mechanisms preparing us to learn to associate certain stimuli with certain emotions...Not only are those associations easier to establish, but...they are also more difficult to undo.” (2008:118-119).

Evolutionary Psychology theories are committed to this view of eliciting conditions being relatively constrained, because otherwise emotions, as modules, would not be able to perform their function in assisting practical rationality. If an emotion could in principle come to be elicited by an indefinite number of things, it would not be useful as an adaptation to a specific kind of adaptive problem. Thus, as Griffiths (1997: 130) writes, Evolutionary Psychology theorists are committed to the view that "developmental programs evolved in an environment which contained varying circumstances, but which on a larger scale of space and time constituted a single 'environment of evolutionary adaptedness'.” This means that there can only be a very limited number of conditional developmental pathways for emotions: if the environment is one way, then we develop along one pathway; if it is another way, then we develop along another pathway; but there cannot be a large number of pathways or the evolutionary advantage of the emotion is lost.

Finally, note that the account of 'plasticity' only acknowledges potential for variation in eliciting conditions of emotions; there is no real discussion of the possibility that the characteristic behaviours of emotions might vary, either from
culture to culture, from person to person, or within the individual over time. This is probably no accident. In fact, there is less scope for variation in emotional ‘outputs’ in Evolutionary Psychology theories than there is for variation in emotional ‘inputs’. The need for evolved ability to have variation in eliciting conditions of emotions is clear, since the things that are dangerous or disgusting might be different in different environments. However, the behaviours that are appropriate to dangerous or disgusting things might be presumed not to vary. Thus, if emotions are to fulfil their function of facilitating appropriate action, they will have to have as outputs a very small range of behaviours.

But this means that there is even less chance of giving a satisfactory account of phenomena like emotional eating: as noted in Chapter Two, what is irrational about emotional eating is not the ‘emotional’ part – the emotional distress may be perfectly rational – but rather the ‘eating’ part. Some individuals have come to exhibit a habitual behaviour – eating – in response to negative emotion. This is an ‘output’, not an eliciting condition, so it is hard for Evolutionary Psychology theories to explain how it might arise. Additionally, the failure to consider the ‘output’ side of emotion causes problems for an account of emotional training and regulation, as discussed below.

2.3 Accounting for Regulation and Training of Emotions

We have seen that Evolutionary Psychology theories do not give a satisfactory account of emotionally-driven irrational behaviour. In this section I will argue, following Jones (2006) that they also cannot give an adequate account of emotional training and regulation, due to their commitment to the modularity thesis.
Jones (2006) has convincingly argued that the commitment to modularity in Evolutionary Psychology theories is in tension with the idea that emotions are necessary for practical rationality, in particular because they seem to rule out the kind of flexibility in response to circumstance which is needed for them to be truly adaptive. Jones argues that pro-emotion accounts must allow that, 

emotions are capable of coming to be directed towards new objects in virtue of a cognitively modifiable range of triggering properties. Put another way, it presupposes the emotions can, with experience and regulation, become reason-tracking mechanisms that enable an agent reliably to track the way her concerns are implicated in concrete choice situations. (Jones, 2006:4; italics added).

In other words, Jones is making a similar point to the one I made in Chapter Two: that a pro-emotion account must give an important place to emotional training and regulation. If emotions are to be helpful rather than harmful to practical rationality, we have to be able to use reason and experience to modify and regulate our emotions. However, Jones argues, and I agree, that the modularity thesis is incompatible with the required degree of modifiability and flexibility in emotional responses.

I argued above that Evolutionary Psychology theories are limited in the scope they can allow for emotions to be shaped by environmental factors, because of the way in which they think that emotions function. Jones locates the problem more precisely in the commitment to modularity: in particular, with the fact that modularity entails informational encapsulation and mandatoriness. As she writes, “modular processing takes processing decisions out of the hands of conscious decision-making: the module does its job whether we will it to or not, and does it
without benefit of knowledge stored elsewhere in the system” (2006:12-13). And this doesn’t just limit the possibilities of environmental shaping of emotional dispositions; it also makes emotions generally poor adaptive solutions to many of the choice situations we would have faced even in the EEA,

Modular mechanisms would be unlikely to deliver veridical solutions to decision problems that are ‘informationally open’ and that depend, for adequate solution, on complex situation-specific cues…There is reason to believe that most practical problems are of this open-textured kind, for two reasons: first, the cues that signal a concern is implicated in a concrete choice situation are often various and recognizing them can take considerable experience and judgment. Second, choice situations are rarely single-track and typically implicate more than one of the agent’s values.” (2006:13)

This is not just saying that emotions might lead us to behave irrationally as a result of the environment no longer being relevantly similar to the EEA; it’s saying that the kind of problem emotions are supposed to help with, even in the EEA, is complex and requires input from reason, experience and judgement. However, a modular system is encapsulated – it cannot incorporate information from other levels of processing – so it is not capable of receiving this input. And this will mean that “emotions will tend to hinder rather than assist the agent in responding well to the choice situations she faces. The traditional view of the emotions as disrupters of practical rationality is more likely true than false” (Jones, 2006:14). In other words, the modularity thesis appears incompatible with the pro-emotion position. And this is because modularity does not allow for the kind of
regulation and control by reason which is necessary if emotions are to be helpful to practical rationality.

Jones' objection focuses mostly on the 'input' side of emotions: that is, on the need to learn from experience and reason when a particular evaluative property is instantiated. However, as noted above, emotional regulation, and the influence of environmental influences, also operate on the 'output' side – on the behaviours (or, at least the action tendencies) that are automatically triggered by experiencing an emotion. If modularity prevents us from having the appropriate flexibility in emotional 'inputs' – that is, from being able to modify the eliciting conditions of our emotions based on learning, experience, and reason – then it will also prevent us from having the appropriate flexibility in emotional 'outputs' – that is, being able to modify the automatic action tendencies associated with emotions.

The Evolutionary Psychology theory is therefore not a satisfactory pro-emotion theory. It cannot properly account for emotionally driven irrational behaviour, since the explanations are not consistent with the observed evidence, and although it allows for some influence of culture and environment, the modularity thesis prevents it from recognising the necessary degree of flexibility in both input and output sides of emotions. This means that it cannot account properly for emotional regulation, and that there is a real risk that it cannot maintain a commitment to the central pro-emotion claim that emotions are helpful to practical rationality.

3. Prinz's Embodied Appraisal Theory
As noted at the beginning of the chapter, Prinz is a neo-Jamesian theorist whose pro-emotion theory has elements both of Evolutionary Psychology accounts and of perceptual accounts. In particular, he is committed to the modularity thesis. Therefore, insofar as modularity is a problem for Evolutionary Psychology theories, it is likely to be a problem for Prinz’s view. Damasio is also a neo-Jamesian but is not committed to modularity; his account will therefore be discussed in the next chapter along with perceptual theories.

Recall that Prinz’s account holds that emotions are embodied appraisals representing core relational themes (C.R.T.) such as ‘irrevocable loss’ (for sadness) or ‘having transgressed a moral imperative’ (for guilt) (Prinz, 2004: 16). Prinz believes that emotions are evolutionary modules: in particular, they are “ontogenetically determined,” informationally encapsulated and mandatory. He also believes that emotions are perceptions (namely, perceptions of bodily states which in turn represent perceptions of values). In fact, Prinz argues that he, unlike other ‘perceptual’ theorists like de Sousa, is a true perceptual theorist, since he thinks that emotions are perceptions rather than just being analogous to perceptions (Prinz, 2004:221). In fact, emotions are perceptions in two ways: they are perceptions of bodily changes, and of core relational themes (2004:225).

As we saw in the previous section, a commitment to modularity poses a problem for pro-emotion theories, in that a modular account apparently cannot deliver the flexibility required for emotions to truly make a positive contribution to practical rationality. Prinz acknowledges this tension (2004:232), but believes that his account is compatible with the requisite degree of responsiveness to learning, experience and reason. His account also explains how emotions can lead to irrational behaviour.
According to Prinz, culture and other influences can affect the conditions under which we experience emotions. Firstly, culturally influenced behaviour can induce the kinds of patterned bodily changes that might lead to experiencing emotions in certain circumstances. Secondly, we can learn new elicitors of emotions by way of our ‘calibration files’. These are mental files that “contain representational states that serve as calibrating causes” which “link emotions to sets of eliciting conditions” (Prinz, 2004:148). These calibration files tell us if and when something is an instance of a core relational theme. New features and aspects can be added to calibration files if we recognise that those things relate to a core relational theme. This could happen either because cultural or social influences call attention to aspects of situations as instances of core relational themes (2004:149-150), or because we (repeatedly and reliably) recognise that a particular aspect is an instance of a core relational theme.

Consequently, there are several ways that emotions could cause irrational behaviour on Prinz’s account. Firstly, the fact that emotions are modular, and therefore informationally encapsulated, can lead to the phenomenon of “emotional recalcitrance,” where we continue to experience an emotion even though we believe it is no longer appropriate (Prinz, 2004:236). Secondly, an element may be added to a calibration file through cultural influences or incorrect beliefs even though it does not in fact represent a core relational theme; in this case, we would have emotions that were inappropriate to the situation, and therefore be motivated to act inappropriately.

This means that Prinz can explain disgust-driven resistance to recycled wastewater as an example of cultural influences creating a disgust reaction. In our culture, disgust may be a “hyper-cognized” emotion, in that our culture focuses on disgust and identifies many aspects of the external world as elicitors of it. We
learn through cultural influences (and through the media’s coverage and use of slogans like ‘toilet-to-tap’) that a core relational theme relating to disgust is instantiated in recycled wastewater.

The case of emotional eating is somewhat more complicated from Prinz’s standpoint. In this case, the negative emotions are not (or need not be) inappropriate to their object: the core relational theme relating to them might very well be instantiated. What is irrational is the behaviour that the negative emotions motivate. The most plausible explanation, given the evidence that the behaviour is (a) not universal, (b) maladaptive in any circumstance, and (c) associated with more general inappropriate coping strategies, is that this is a behaviour which is learned as a response to experiencing negative emotion. And Prinz’s account has some trouble with this, due to the way in which he tries to reconcile his commitment to modularity with his acceptance of the importance of cultural influence.

Prinz believes that emotions are both mandatory and informationally encapsulated, despite also agreeing that they can be influenced by our beliefs and judgements. He argues that these two positions can be made compatible by drawing a distinction between the “initiation pathways” of emotions, which lead from mental states to patterned bodily changes, and the “response pathways” of emotions. These response pathways are “where the actual emotions take place,” and “respond to bodily changes and core relational themes” (Prinz, 2004:234). Prinz thinks that the initiation pathways are not part of the emotion module, but the response pathways are. Therefore, emotional plasticity is possible, but only regarding the initiation pathways. The initiation pathways are where the calibration files are, so we can learn and incorporate new eliciting conditions (2004:234). Response pathways are unable to learn anything or be responsive to
judgements or beliefs or cultural influences (because of their modularity). This means that, as with the other Evolutionary Psychology theories discussed above, Prinz cannot accommodate the possibility of the shaping of the ‘output’ side of emotional dispositions. Hence, the phenomenon of emotional eating is difficult to explain on his theory.

The commitment to modularity also causes problems for Prinz in giving an account of emotional regulation and training. He does admit that we have some control over our emotions: “thinking about something in the right way can certainly influence our emotions, and calibration files can be modified through education and experience. We exert control over emotions by choosing what to think about, and by cultivating calibration files” (2004:236). However it is not clear that this is sufficient. Firstly, while Prinz explains how new things can get into calibration files, it’s not clear how inappropriate things can get out of them. Perhaps, though, he would just say that a belief that something doesn’t instantiate a core relational theme, if repeatedly brought to mind, might eventually remove that thing from the calibration file.

Secondly, though, the story of how something new gets into a calibration file isn’t very plausible, and seems overly complicated. To introduce a new element into my calibration file, I have to believe (either because of cultural influence or because of observing it repeatedly) that a core relational theme is instantiated. Prinz illustrates with the example of schadenfreude (2004:142-143): he thinks that we gain ‘someone else’s misfortune’ in our calibration file for ‘joy’ by observing someone’s misfortune together with the belief that this advances our goals. But this seems incorrect, not only as an understanding of schadenfreude, but also as an account of how we come to experience it. Schadenfreude is not simply joy at someone else’s misfortune; rather, it is often joy at someone else’s deserved
misfortune; it is also not necessary to believe that this misfortune somehow furthers my goals – it seems to me that most schadenfreude is not accompanied by this belief. In any case, though, this story of how we come to acquire schadenfreude is not intuitively plausible. It just does not ring true that we have to observe others' misfortunes and believe this furthers our goals repeatedly before acquiring the capacity for schadenfreude.

**Conclusion**

In this chapter I have argued that strategic theories, Evolutionary Psychology theories and the neo-Jamesian account presented by Prinz are not adequate as pro-emotion theories. In particular, they do not provide a plausible and comprehensive explanation of emotionally driven irrational behaviour, nor do they plausibly explain and allow for the improvement of emotional responses via regulation and training. Strategic theories, apart from being incomplete, will have trouble maintaining the central pro-emotion thesis that emotions are necessary (or, in its weaker version, helpful) for practical rationality, since the theory entails that it is a purely contingent matter whether emotions will be on balance helpful rather than harmful, and there is good reason to think they will very often be harmful. This is also a problem for Evolutionary Psychology theories, which is exacerbated by the commitment to modularity. All three theories – strategic, Evolutionary Psychology, and Prinz's neo-Jamesian view – do not allow for the required degree of flexibility in shaping and regulating emotions through learning, experience, and regulation.

In the next chapter I will consider the way in which perceptual theories (including Damasio's neo-Jamesian theory) can address these same issues. I will argue that these theories do much better than the strategic and Evolutionary
Psychology theories. In particular, they give a satisfactory account of emotionally
driven irrational behaviour which is compatible with their commitment to the
central pro-emotion thesis that emotions are necessary for practical rationality.
Additionally, they are able to account for emotional regulation, training, learning
and cultural influence. Thus, I will conclude that perceptual theories are the best
candidates for a comprehensive account of the functions of emotions in practical
rationality. In the final chapter I will see what this conclusion can tell us about
emotional management at an individual and a public policy level.
Chapter Four: The Success of the Perceptual Accounts

In the previous chapter I discussed strategic theories, Evolutionary Psychology theories and Prinz’s Neo-Jamesian theory in terms of what they had to say about emotionally driven irrational behaviour, and their potential for regulation and training of emotions by reason. I argued that these theories had several significant flaws, and therefore are not satisfactory pro-emotion theories. In this chapter I will look at perceptual theories (including Damasio’s neo-Jamesian account), and show that they give a much better account of both emotionally-driven irrational behaviour and of emotional training and regulation. They also do this in a way which maintains the commitment to the central pro-emotion claim that emotions are necessary for practical rationality. In the next and final chapter I will show how this account can be used to assist us in ensuring that emotions play a helpful and not harmful role in practical rationality, by making recommendations about emotional management at both an individual and a public policy level.

As noted in the previous chapter, a satisfactory pro-emotion theory will give a complete account of the roles played by emotions in practical rationality. It will also explain why emotions often cause irrational behaviour (and do so in a way which is compatible with the observed evidence), and explain how emotional training and regulation can occur. These explanations must be consistent with the commitment to the claim that emotions are necessary (or at least, useful) to practical rationality. In this chapter I will argue that perceptual theories can meet all of these criteria.
1. **Perceptual theories revisited**

Perceptual theories claim that emotions are analogous to perceptions: they are patterns of salience among objects of attention (de Sousa, 1987:196), or ‘somatic markers’ that attach to options. On this view the emotions play the role of signalling, and focusing our attention on, evaluative properties instantiated in a given situation. Emotions therefore facilitate practical rationality by quickly identifying just those aspects of our situation that are relevant to decision-making. They also allow us to reduce our options to a manageable set, and quickly decide from amongst them. Finally, they prepare us to quickly act on our decision.

As noted previously, the term ‘perceptual theories’ is slightly misleading because theorists such as de Sousa do not think that emotions *are* perceptions, only that they are closely analogous to perceptions. However, this is in fact an advantage. Prinz was able to claim that emotions were perceptions because he was committed to the modularity thesis; hence, on his account emotions, like other forms of perception, were informationally encapsulated. Because perceptual theorists do not think that emotions are perceptions, they are not committed to a modularity thesis, and do not have to claim that emotions are informationally encapsulated or mandatory. This means that the problems associated with the modularity thesis do not prima facie apply to the perceptual account.

2.1 Damasio’s Account

According to Damasio, the somatic markers that place a negative or positive valence on certain of our options are not set by evolution, but rather obtained through learning – they “have been connected, by learning, to predicted future outcomes of certain scenarios” (1994:159-160). There is an “internal preference system” (1994:164) – which means that we are disposed to have
emotional responses to certain general features of the environment – but the set of specific properties that instantiate those general features is a matter of our environment, and is open-ended. Moreover, we acquire new somatic markers not just in childhood but throughout our life: “the accrual of somatically marked stimuli ceases only when life ceases, and thus it is appropriate to describe that accrual as a process of continuous learning.” Damasio therefore allows a much more extensive and open-ended influence of learning and external factors on emotional development than the Evolutionary Psychology theorists. Additionally, the operation of somatic markers is not sufficient for decision-making (Damasio, 1994:159-160): they play the role of sifting through the massive number of possible outcomes and relevant considerations, to pick out the ones that are most relevant and important, but it is still up to us to decide between them using ‘unemotional’ reason.

2.2 De Sousa’s Account

According to de Sousa, emotion plays a role in remedying the defects of ‘pure’ decision-theoretic reasoning by solving the ‘philosophers’ frame problem’ of having too much irrelevant knowledge in any choice situation, and needing to retrieve only the bits that we need (1987:192-4). While he believes that evolutionary factors are very likely to be responsible for the presence of emotions in our mental makeup, he does not think that “they could yield anything like a complete theory of the emotions. One reason is the existence of conventional and social factors that are at least prima facie irreducible to biological causation” (1987:44).

De Sousa thinks that we learn emotional associations by means of ‘paradigm scenarios’, which
are drawn first from our daily life as small children and later reinforced by the stories, art, and culture to which we are exposed…Paradigm scenarios involve two aspects: first, a situation type providing the characteristic objects of the specific emotion-type…and second, a set of characteristic or ‘normal’ responses to the situation, where normality is first a biological matter and then very quickly becomes a cultural one. (1987:182)

So a paradigm scenario is a kind of foundational norm for a particular emotion: it is the standard for when an emotion is appropriate or called for. While there are some genetic predispositions for emotional response, we mostly learn paradigm scenarios by having others teach us that we are experiencing a particular emotion, and in pointing out the context in which this emotion is appropriate, so that we can learn “to feel the right emotions” (de Sousa, 1987:183). Like Damasio, de Sousa also believes that “learning these scenarios continues indefinitely… as the emotional repertoire becomes more complicated” (de Sousa, 1987:183). We learn to have emotional responses to new objects by seeing them as relevantly similar to our paradigm scenario.

Moreover de Sousa explicitly allows for ‘un-learning’ or modification of emotional responses, even to the point of deciding that our emotions are not appropriate to their original paradigm scenarios. This addresses the requirement to account for learning, regulation as a result of cultural influences, and training the output (or behaviour) that stems from emotion. He claims that the fact that a concept is learned in a given context does not mean that the concept cannot be revised and refined, that our understanding of it cannot be deepened to the point where we are able to ask without contradiction whether it is appropriate to the paradigm itself…. A paradigm can always
be challenged in the light of a wider range of considerations than are available when the case is viewed in isolation. It can also be revised in the light of competing paradigms that are also applicable to the situation at hand. (1987:187)

Thus, we can come to see (via reason or belief) that an emotion is not in fact appropriate to some context, no matter how deeply ingrained, and then unlearn it, even though this might be hard.

The perceptual account therefore allows for the kind of learning, flexibility, and potential for later modification that was a problem for the Evolutionary Psychology accounts; it explains “how emotions can come to be triggered by all sorts of things, most of which were not found in the EEA and so cannot have been the cues for situations presenting core relational themes under conditions of selection,” (Jones, 2006:20). They allow for reason to influence emotional development and regulation, because they are not committed to informational encapsulation that is characteristic of modularity.

As both Damasio and de Sousa allow for ‘higher cognitive processes’ to influence emotion acquisition and decision-making where emotions are involved, they can account for this. And because at least de Sousa holds that emotions are (analogous to) perceptions of value, this explains why “cultural views about what is valuable can be expected to affect [emotions] very directly” (Nussbaum, 2001:157).

2. Explaining emotionally-driven irrational behaviour

Both Damasio and de Sousa are quite ready to acknowledge the role of emotions in producing irrational decisions, and have explicitly allowed for this in their theories. Damasio says that “There never has been any doubt that, under certain circumstances, emotion disrupts reasoning. The evidence is abundant and constitutes the source for the sound advice with which we have been brought up” (1994:69). Moreover, for all that he titles his book *Descartes’ Error*, he writes that he *agrees* with what he takes to be Descartes’ claim that “the control of animal inclination by thought, reason and the will was what made us human” (1994:122).

Damasio’s view is that, while emotional response might have been sufficient for roughly adaptive behaviour, there has to be an evolutionary reason for our *feeling* the emotion (that is, perceiving the emotion as bodily state):

In short, feeling your emotional states, which is to say being conscious of emotions, offers you *flexibility of response based on the particular history of your interactions with the environment*. Although you need innate devices to start the ball of knowledge rolling, feelings offer you something extra. (1994:129).

In other words, we are set up the way we are so that we can strike a compromise between quick-and-dirty reaction, and endless deliberation. The emotions prime us in certain ways, but do not lead inevitably to action, and which allow for knowledge and reason to enter into the connection between emotion and decision.6

---

6 Of course, exactly how this is supposed to happen is not entirely clear. It is possible that this process is complicated, involving one emotion to be motivated by another emotion; ‘reason’ may be a matter of bringing in one emotion in order to regulate another.
So according to Damasio, we are set up to infer causal connections between some external event or person and a body state, and this can go awry. Nevertheless, the tendency to form such connections can be moderated by reason and effort: “we often make wrong connections, for instance, when we associate a person, object, or place with a bad turn of events, but some of us try to keep from making those erroneous links” (1994:150-151). When we do make emotionally-driven irrational choices, it is because we are allowing a mental representation and its somatic marker “to dominate the landscape of our reasoning and to generate a negative bias against the correct choice” (1994:174). Damasio’s wording (he does use the word ‘allowing’) suggests that he thinks there is an element of conscious control, at least potentially, which we could learn to employ in order to avoid the irrational choices.

For Damasio the potential to cause irrational decisions is inherent in the role played by emotions in practical rationality. Emotions are necessary for practical rationality because they ‘orient’ us to relevant features of our choice situation, and because they allow us to represent the goodness or badness of future outcomes in a way that facilitates planning agency. But this very role leads to the potential for irrationality. If we make bad connections when learning to attach somatic markers to certain things, then we can misidentify certain options as relevant or irrelevant. And we can also ‘allow’ the somatic marker and the possible outcome it’s attached to, to ‘dominate’ our reasoning in a way which leads us to choose badly.

But, because reasoning does enter into the pathways, both in terms of learning to ‘somatically mark’ new things, and in making decisions after having the somatically marked options and outcomes made salient to us, we can learn to avoid these kinds of irrational behaviours. Indeed Damasio thinks that knowing
how emotion works in practical reasoning is important because we can then learn better how to regulate and train our emotions:

Knowing about the relevance of feelings in the processes of reason does not suggest that reason is less important than feelings, that it should take a backseat to them, or that it should be less cultivated. On the contrary, taking stock of the pervasive role of feelings may give us a chance of enhancing their positive effects and reducing their potential harm. Specifically, without diminishing the orienting value of normal feelings, one would want to protect reason from the weakness that abnormal feelings or the manipulation of normal feelings can introduce in the process of planning and deciding. (1994:214)

Like Damasio, de Sousa also thinks that the role emotions play in practical rationality has the potential to cause irrational decisions. Like many historical theorists, he implicated emotions as a major cause of weakness of will: “emotions affect the relative saliency of the two arguments, thus disconnecting the ‘stronger’ from the train of action” (1987:5). Emotions “give us frameworks in terms of which we perceive, desire, act, and explain,” which is necessary for practical rationality, but this can also cause irrationality “due to their power to reinterpret the world” (1987:24).

In other words, we need emotions so as not to be paralysed by ties of reason and by the overwhelming volume of potentially relevant knowledge and information, but the role that they play in picking out and highlighting relevant concerns can also highlight and pick out irrelevant things and lead us to ignore relevant ones. This is because we have mistakenly assimilated something to our
paradigm scenario that we should not have. However, there is some ability to exercise conscious control over our emotions, although it is very difficult.

Emotions are, in part, patterns of attention. Therefore, one might expect a change in patterns of attention to detail a change in emotion…To some degree, attention is in our power. But it is a lot easier to attend at will than to withdraw one’s attention at will...So we may expect great difficulty in trying to get rid of an unwanted emotion, but more success in working ourselves into one. (de Sousa, 1987:243)

3. Application to the Case Studies

In the previous sections I showed that perceptual pro-emotion theories can explain how emotions can cause irrational behaviour, and that they allow for the possibility of emotional training and regulation. In this section I will test the adequacy of their accounts by applying them to the case studies of recycled wastewater and emotional eating.

In Chapter Three, we noted that the Evolutionary Psychology explanation for disgust at recycled wastewater was superficially plausible but inadequate. The idea was that the disgust module has evolved to be activated by substances known to have come into contact with waste, or alternatively to be responsive to what others around us consider disgusting. The first proposal cannot explain the inconsistency of disgust reactions across cultures, and the fact that people who are disgusted by wastewater are not disgusted by other substances known to have come into contact with waste. The second is more promising, since it gives a significant role to cultural and social influences in shaping the ‘inputs’ to our
disgust response. However, as I argued, the role given to social and cultural factors in Evolutionary Psychology theories is not broad enough.

In particular, the point made by Jones and others about the need for flexibility of response applies here. There is some plausibility to the idea that there might be an adaptive advantage in children’s quickly acquiring the disgust responses of those around them, since this will often be the safest and most efficient way to acquire the knowledge that has built up in a society about those things that are actually likely to cause contagion or illness. However, further thought indicates that it would in fact be equally important that those acquired disgust responses be able to be revised on the basis of new experience or exposure to new environments. This must involve not just the ability to add new things to the eliciting conditions of disgust, but also the ability to remove things that are already there. A feeling of disgust for pork may be adaptive in a context where pigs are in fact likely to carry diseases, but if food becomes otherwise scarce, it would be wiser to put this disgust aside. The modularity thesis, however, does not appear to give sufficient room for this kind of modification of disgust responses by reason and experience.

Perceptual theories can do much better on this account. They can accommodate the plausible idea that we are predisposed to acquire certain disgust responses – for example, that it is much easier for us to learn a disgust response to things that are known to have been contaminated by waste, or to things that the adults around us find disgusting. On de Sousa’s account, there will be an evolved tendency to be disgusted by certain things, but we will also acquire a paradigm scenario for disgust very early in life from our parents and other adults. If they exhibit disgust at an object, or declare something to be disgusting, we learn what kinds of things it is appropriate to respond to with disgust. In particular, it is quite
plausible that human waste is something we are set up to include in the paradigm scenario for ‘disgusting’, and that we will learn the ‘law of contagion’ very early in childhood and hence assimilate this into our paradigm scenario. We can then acquire disgust at recycled wastewater by seeing that the water fits our paradigm scenario for disgust due to its being contaminated, through the law of contagion, by its past history of being in contact with human waste products.

This theory also explains why people who are disgusted by recycled wastewater are not similarly disgusted by regular tap water, even when they are told that it has a similar history of contamination. Recycled wastewater is a substance that we have not commonly encountered, and so we are more likely to imagine its abstract properties (in particular, where it came from) because we have not been exposed to its more concrete properties. On the other hand, we have a long history of use and exposure to regular tap water, and we have learned that it is not an appropriate object of disgust, since we drink it, bathe in it, and so on. This means that we would resist the conclusion that normal tap water is appropriately similar to our paradigm scenario for disgust.

An additional point to note is that the idea that disgust is appropriate to recycled wastewater is something that is reinforced, for example in media reports and by protest groups using slogans like ‘toilet to tap’. The media, advertisers and cultural norms also have a role in the more general ‘hyper-cognizing’ of disgust in our society, where our attention is continually drawn to the possibility of ‘germs’ and contamination, with the message that this is an intolerable state of affairs. This creates an environment in which we are repeatedly exposed to the message that disgust is an appropriate response to recycled wastewater. The ‘wastewater’ part is made over-salient, while the ‘treated’ part is ignored. In de Sousa’s terms, the emotional response (disgust) plugs its inappropriate conclusion (that we
cannot support recycled wastewater schemes) into the decision-making process in place of the whole argument, which is ignored. In Damasio's terms, the idea of having recycled wastewater has been tagged with a very powerful negative somatic marker which causes us to dismiss this option from serious consideration.

Moreover, because perceptual theories acknowledge the role of cultural and social factors in shaping and reinforcing people's paradigm scenarios, and in making it more likely that certain objects will be seen as relevantly similar to those paradigm scenarios, they can account for the difference in disgust reactions seen in different societies. It is true that the general disposition to find human waste disgusting, and the learning of the 'law of contagion', will probably be more or less universal. However, this doesn't mean that any object which could be seen as disgusting according to the law of contagion will be, just as people do not draw all of the possible inferences from their existing beliefs. As we have seen, people who are disgusted by wastewater are not disgusted by normal tap water, because other factors (their long experience with the substance itself, or perhaps an element of self-deception) override any tendency to draw that 'conclusion'. In other cultures and societies people haven't assimilated recycled wastewater to their disgust paradigm scenario in the first place, or they have learned to disconnect them.

There are various ways in which cultural and environmental factors might contribute to the reduced tendency to be disgusted by recycled wastewater. Firstly, a lack of media and activist group attention to the 'disgustingness' of recycled wastewater and much more focus on other aspects (such as purity, 'hi-tech', and self-sufficiency of supply) works to encourage the idea that recycled wastewater is an appropriate substance for consumption, which is incompatible with a disgust response. Furthermore, exposure to the substance itself will reduce
people’s tendency to react to the idea of recycled wastewater over the concrete properties of the water itself. Singapore runs tours of its NeWater facility, and the public is accustomed to the idea that it is used for their domestic water supply.

The case of emotional eating differs from that of recycled wastewater in that it is not the emotional response that is inappropriate, but rather the behaviour that goes with it. In the last chapter I gave evidence that this is not plausibly considered an adaptive response, and it appears to be part of a more general syndrome of maladaptive coping mechanisms. Evolutionary Psychology theories cannot account for the fact that only some people exhibit this behaviour, that it is not adaptive, and that the mechanisms under which it might be adaptive (serotonergic activity, e.g.) are not supported by the evidence. The behaviour appears to be learned. However this is difficult for Evolutionary Psychology theories, since they only allow for the effects of learning and culture on the ‘input’ side – that is, on eliciting conditions for the emotion. In this case, though, it is the behaviour resulting from the negative emotion that is irrational, and this is an output. Evolutionary Psychology theories, as previously argued, do not have the ability to explain flexibility of emotional response related to the ‘output’ side.

Perceptual theorists, however, can account for this. In fact de Sousa explicitly allows for emotional expression and behaviour to be learned, not just the ‘input’ side of eliciting conditions. Along with paradigm scenarios giving us appropriateness conditions for when to experience the emotion, we learn sets of appropriate behaviour and expression that characterise having the emotion. Thus, we could explain eating in response to negative emotion as a behaviour pattern which is learned in early childhood and reinforced through cultural and social influences. The child may be fed when distressed, may see parents and other adults modelling this same behaviour pattern; then later they will encounter media
stereotypes along the same lines (the heartbroken just-dumped woman ‘pigging out’ on ice cream; advertising of junk foods, especially sweets, with an explicit message that they can be used to deal with negative feelings).

It is important to note, however, that the situation is slightly more complicated than just that of eating as a ‘programmed’ and automatic response to experiencing negative emotion. As noted in Chapter Two, the eating behaviour appears to be an attempt to alleviate the negative emotion. Thus, eating is not associated with general negative emotion in the way that fleeing is associated with fear; rather, the behaviour constitutes a (poor) attempt to regulate emotion. However, behaviours that aim to regulate emotion can obviously be learned in the same way as expressive behaviours and other emotional ‘outputs’. In this case, it is a bad strategy for emotional regulation, which has occurred because of a mistaken identification of the relevant property of formative comforting situations, such as being taken care of by parents, or getting a treat, which involved food. Thus the food comes to be added to the paradigm scenario for ‘comforting’, simply because it was often present in situations which actually were comforting. According to Damasio’s theory, we may have tagged food with positive somatic markers because of repeated association with comforting situations.

Perceptual theories can therefore account for how strategies for emotional regulation can be acquired or learned, as well as how we can learn both eliciting conditions and behaviours for our emotions. These accounts give a good explanation of the emotionally-driven irrational behaviour in the case studies, which is consistent with the observed evidence. Moreover, perceptual theories seem to give a satisfactory account more generally of the role of emotions in practical rationality.
Conclusion

Perceptual accounts explain how irrational decisions and behaviour can be caused by emotions. Unlike Evolutionary Psychology theories, they do this in a way which does not threaten the central pro-emotion thesis that emotions are necessary for practical rationality, and which is consistent with the observed evidence regarding actual cases of emotionally driven irrational behaviour. For de Sousa, we need emotions to solve the ‘philosopher’s frame problem’ and sift all potentially relevant considerations down to a manageable set; but precisely because emotions focus our attention on some aspects of a situation and away from others, they can cause us to ignore relevant factors and place too much weight on irrelevant ones. According to Damasio, we need somatic markers to sift out the most important potential outcomes of our decisions from amongst an indefinitely large set: these somatic markers cause us to immediately discard certain options (as leading to a negatively marked outcome) and to be drawn towards others (as leading to a positively marked outcome). But again, if our somatic markers are mistaken or inappropriate, we may discard good options and focus on bad ones.

Perceptual accounts therefore provide a complete pro-emotion account which can explain why emotions are necessary for practical rationality, but at the same time how they can cause irrational behaviour and decisions. They also allow for the possibility of emotional regulation, and can explain how this can occur. This means that perceptual accounts can, at least in theory, provide advice on how we might enhance the helpful aspects of emotions, and avoid the harmful aspects, in practical rationality. An account of how eliciting conditions and behaviours associated with emotions are shaped can provide us with suggestions for making sure that we avoid inappropriate eliciting conditions and connect helpful
behaviours to our emotions, which would make sure that emotions can better
perform their functions as reason-trackers, aids to decision-theoretic reasoning,
and motivators to correct action. But to do this in adults (rather than those in
their formative years) requires the use of emotional regulation to decouple
emotional experience from habitual behaviour, to up-regulate or down-regulate
emotional experience itself, and to modify our existing emotional dispositions.
Again, the perceptual theory’s account of how emotional regulation might occur
will give rise to more specific suggestions for how we ought to engage in the
management of our own emotions. In the next chapter I will explore the question
of emotional management, from the point of view of both the individual, and
public policy.
Chapter Five: Dealing with Emotionally-Driven Irrational Behaviour

In the previous chapter I argued in favour of perceptual pro-emotion theories as giving a satisfactory explanation of emotionally-driven irrational behaviour and emotional training and regulation, while still retaining the crucial claim that emotions are necessary for rational decision-making. Indeed, these theories allow that emotionally-driven irrational behaviour can arise directly out of the roles played by emotions in facilitating rational behaviour. Emotions facilitate rational behaviour, essentially, by rendering our ‘decision space’ manageable: they highlight certain features of a situation and allow us to quickly focus on those while ignoring others, and at the same time provide motivational priming for certain actions. This is helpful when the features picked out are those that are actually relevant, since it prevents us from having to consider an indefinitely large number of possibilities; it also facilitates quick responses to the features that have been identified as salient. On the other hand, emotions can also cause us to give too much weight to irrelevant features or to ignore relevant ones; and the motivational aspect of emotion means that these mistakes may be translated into inappropriate action.

At the same time, perceptual pro-emotion theories have the advantage of allowing the theoretical possibility that we might do something about emotionally-driven irrational behaviour, because training and regulation of our emotions is possible throughout life. Unlike Evolutionary Psychology accounts, emotions are not informationally encapsulated modules with mandatory outputs; they are (at least in principle) open to cognitive influence. Our emotional
responses are learned, although we may be innately primed to learn some responses more easily than others, and so it is possible that they can also be unlearned or modified.

In this chapter I will talk about ways in which we might deal with emotionally-driven irrational behaviour, if we take the perceptual pro-emotion theory as a reasonably accurate picture of how emotions work in decision-making. Perceptual pro-emotion theorists themselves have not devoted much discussion to ways in which the theoretical possibility of emotional training and regulation might be realized in practice. As noted in Chapter Two, however, the need for emotions to be controlled and regulated by reason was a major theme in most historical theories of emotion. Moreover, as I have argued, many historical theories are quite compatible with the pro-emotion view, and even hold that emotions play roles in practical rationality quite similar to those accepted by the pro-emotion consensus. If these historical theorists had the same view as perceptual pro-emotion theorists of the way that emotions function in practical rationality, then their suggestions for how we might successfully regulate and manage our emotions are likely to be useful in the current context: in other words, perceptual pro-emotion theorists can make use of historical insights on emotional management. Perceptual pro-emotion theorists can supply the explanation of the mechanisms behind emotional regulation, and historical theorists can supply some valuable suggestions for practical approaches to the regulation of emotion.

The particular historical theories I will draw on in this chapter are those of the Stoics, Aristotle and Descartes. We saw in Chapter Two that the Stoics did advocate the elimination of emotion, and hence their view is not strictly speaking compatible with the perceptual pro-emotion account, according to which this is not only undesirable but impossible. On the other hand, the Stoics shared with
perceptual pro-emotion theorists the idea that emotions were appraisals (or more specifically, evaluative judgements); their advocacy of elimination of emotions followed from their claim that all of the evaluative judgements involved in emotional response were false: both those which ascribed evaluative properties to objects, and those which held that it was appropriate to react. We need not share this view in order to accept the Stoics may have had some useful insights on how to deal with emotions that are inappropriate to their objects, or which lead to inappropriate action. And Aristotle and Descartes both had views of emotions, and their functions in practical rationality, which are quite congenial to modern perceptual pro-emotion theories. They also devoted substantial attention to the question of how we might ensure that our emotional responses were correct. Aristotle also considered how we might regulate the emotions of others – a question which will become relevant in the discussion of public policy in cases of emotionally-driven irrational behaviour.

This chapter in fact has two overall aims. The first is to show how perceptual pro-emotion theories can give us useful suggestions for engaging in emotional management based on their account of the function of emotions in practical rationality. While it is beyond the scope of this thesis to demonstrate that these suggestions are efficacious, I will provide some support for this conclusion by showing that the suggestions cohere well with research on emotional regulation in the cognitive sciences. The second aim is to effect a sort of reconciliation between the modern pro-emotion theories and those historical theories that anticipated their views. In the second chapter I argued that there is little support in historical theories of emotion for the notion of a historical ‘anti-emotion’ consensus. Here I will try to merge modern and historical discussion of
emotional regulation to show that the modern pro-emotion consensus is the continuation of, rather than a break from, certain strands in historical thought.

1. Dealing with Emotionally-Driven Irrational Behaviour – Theory

As mentioned, perceptual pro-emotion theories take emotions to be analogous to perception; they are, more specifically, perceptions of value-laden aspects of the world. Just as we automatically focus on movement in our visual field without conscious decision, so our emotions focus our attention on those aspects of a situation or decision problem which are salient in the sense of being relevant to our concerns. In de Sousa’s terms, this happens by way of paradigm scenarios, which give us a kind of ‘baseline’ for appropriate eliciting conditions for particular emotions; we experience an emotion when we encounter situations that are relevantly similar to a paradigm scenario, and this helps us to quickly see what is important in that situation, and to be primed to act accordingly. According to Damasio, this happens by way of ‘somatic markers’ that are attached to potential outcomes of the options we are considering. The somatic markers come to be attached to certain features by learning that they are associated with things that are good or bad in particular ways; when we are deciding what to do, the presence of these somatic markers causes us to immediately discard some options as aversive, or to be attracted to others.

I have already mentioned the role of learning and training in the development of paradigm scenarios or the attachment of somatic markers. This is extremely important in considering whether (and how) we can deal with emotionally-driven irrational behaviour. Equally important, however, is the fact that according to perceptual pro-emotion theories the role of emotions is basically
to deliver a manageable choice set from which we can decide what to do using ‘conventional’ decision-theoretic or rational choice methods. Of course, emotions deliver a (sometimes very strong) ‘recommendation’ at the same time, in the form of their motivational strength, physiological arousal and action tendencies. But there is space in between emotional input and decision output for us to use other methods of deciding.

This view is consistent with the way cognitive scientists conceive of the operation of emotions. For example, Gross and Thompson (2006: 5) write that while emotions “do possess an imperative quality … [in that] they can interrupt what we are doing and force themselves on our awareness,” there is considerable scope for working with this: in fact, “emotions often must compete with other responses that are also occasioned by the social matrix within which our emotions typically play out.” Thus, our emotions are ‘malleable’ and can be “modulated in a large number of ways.” These methods can be conscious or can occur without our conscious awareness; in fact, the former kind of regulation can be transformed into the latter with sufficient repetition (Gross and Thompson, 2006: 6).

In fact, there are three different ways in which we might attempt to intervene in cases of emotionally-driven irrational behaviour. The first is attempting to alter the disposition to feel the emotion in certain circumstances, by removing the somatic marker or undoing the assimilation of the situation to one’s paradigm scenario, so that it does not arise at all. The second is attempting to moderate, extinguish or even increase an occurrent emotion, after it has already arisen. The third is attempting to interrupt the connection between emotion and behaviour, without attempting to moderate the emotion or the disposition to feel it in such situations. All of these are possible on the perceptual pro-emotion view,
but which is the best approach will depend on the nature of the emotionally-driven irrational behaviour, the nature of the agent, features of the situation, and so on. Moreover, we will need some ideas about how to attempt these kinds of intervention, and when or whether they are likely to succeed. This is where the insights of traditional theories, as well as those of modern cognitive science, might be useful.

The idea that we might use the insights from historical theories of emotion in order to flesh out the types of emotional regulation that are possible on a perceptual pro-emotion view might be thought to need some defence. After all, there are still some important differences between the general conception of emotions on historical accounts and the way in which they are regarded by pro-emotion theorists. Even those philosophers such as Descartes, who disagreed with the Stoics about the elimination of emotion completely, still took emotion and reason to be two different faculties. They therefore thought that we should be using one faculty – reason – to control another – emotion. Perceptual pro-emotion theorists, on the other hand, would mean something different by talking about using reason to regulate emotion. In this case, what we are trying to regulate is bound up in the resource we are using to regulate it. The recommendations one might make about how to get a reason faculty to regulate an emotion faculty would (it might be thought) be quite different to the ways in which we might try to regulate emotion when emotion is inevitably part of the regulating mechanism.

But in fact the differences are not as significant as all that. Historical theories such as Descartes’ or Plato’s, which took emotion and reason to be separate faculties thought of emotions as identifying things of concern to us, and then presenting this to the faculty of reason along with a (sometimes very strong) dose of motivation to act in a particular way. Perceptual pro-emotion theorists
likewise take emotions to pick out those features of a situation that are value-laden, and to present the results of this activity, along with corresponding motivations, as a set of feasible alternatives; they then appear to believe that we can use ‘pure’ rational choice or decision-theoretic methods to pick from these alternatives. The main difference between the two is that two-faculty theories think it would be possible to have ‘pure disembodied reason’, whereas perceptual-pro-emotion theorists would deny this. But the methods by which a faculty of reason might resist the influence of a separate faculty of emotion will be applicable to the perceptual situation, in which decision-making might resist the influence of emotion which is operating in the decision process.

The fact that the Stoics advocated eliminating all emotion does not mean that they cannot give us useful advice about how to manage emotions, or emotionally-driven behaviour that is inappropriate or irrational. The Stoics thought that all evaluative judgements that formed the basis of emotions were false, and the reason for eliminating emotions was therefore because we ought not to have, or act on, false evaluative judgements. Perceptual theorists would disagree that all the evaluative judgements (or appraisals) are false. But if the Stoics had effective methods to reduce or eliminate certain kinds of occurrent emotions or emotional dispositions, or if they taught us how to stop emotions causing irrational behaviour, these would be useful on the perceptual pro-emotion view as well. In any case, as Richard Sorabji (2000:5) notes, the Stoics’ proposals for dealing with unhelpful emotions were intended to be of assistance to anyone, including those who did not accept the Stoic theory.

2. **Breaking the Connection Between Emotion and Behaviour**
Perhaps the most straightforward way of trying to handle emotionally-driven irrational behaviour is by trying to disrupt the connection between emotion and behaviour. Indeed, according to Descartes, it is the only thing we can do in the face of very strong emotional experience: “the most the will can do while this commotion is at its full strength is not to go along with its effects, inhibiting many of the movements to which it disposes the body” (Descartes, 1985:345). This is obviously possible on all pro-emotion accounts (and not just the perceptual theories), since not even the crudest versions of Evolutionary Psychology theories will hold that emotions necessarily cause their characteristic expressive behaviour. Nevertheless, we might need some training and practice in learning to identify and utilise the space between being presented with emotional ‘conclusions’ and motivations, and acting.

But perceptual pro-emotion theories also allow that we might, to a certain extent, change the habitual or characteristic behaviours associated with emotions. This means that trying to break a connection between emotion and irrational behaviour need not consist entirely in suppressing the behavioural ‘output’ of an emotion. Rather, we might learn to modify that behavioural output so that the characteristic or habitual behaviour corresponding to a particular emotion is no longer irrational.

This approach will be most appropriate when the emotionally-driven irrational behaviour is irrational because of the behaviour rather than the emotion itself: that is, when it is not the emotion that is inappropriate or irrational, but just the behaviour that we are motivated to perform as a result of experiencing the emotion. The case of emotional eating is an example of this: the negative emotions which cause emotional eating are nor (or need not be) inappropriate or in need of modification. Rather, what is inappropriate is the eating behaviour that the person
is habitually motivated to engage in when experiencing negative emotions. The goal will therefore be to disconnect this behaviour from the experience of negative emotion, and to replace it with something that is a more constructive or adaptive response.

Historical theories tended to focus on modifying or eliminating the emotion itself as a way to affect emotionally-driven irrational behaviour. Some of their advice and suggestions for training are useful in thinking about how we can attenuate the link between emotion and characteristic behaviour. The Stoics, for example, believed that emotions involved two evaluative judgements: firstly, that there is something good or bad about the situation, and secondly, that it is appropriate to react to it (by experiencing the emotion, and expressing it). According to Sorabji, while some Stoics tried to work on the tendency to make the first kind of judgement, Chrysippus targeted the second, and tried to train people to withhold the judgement that it is appropriate to react. As Sorabji writes (2000:42), “Untrained people in an emotional state, the Stoics would say, are likely to have assented to appearances so automatically as not to recognize assent as a distinct operation” – in other words, they endorsed their emotional responses, and the appropriateness of their actions, without thought. Being trained to withhold assent was supposed to prevent one from experiencing the emotion. However, we can also see how such training would help us to better control the characteristic behaviour resulting from the experience of emotion, and eventually to instil a new habitual behaviour.

One of the elements of Stoic training was to have people continually question ‘appearances’, by asking ourselves whether it was really good or bad, or really appropriate to react. Since perceptual pro-emotion theorists do not aim to eliminate all emotion, it is probably not necessary to question our emotions and
motivations relentlessly. But when we have identified an emotionally-driven
behaviour as undesirable or irrational, it would be compatible with the perceptual
pro-emotion theory to question, not the experience of negative emotion itself, but
the motivation to engage in the relevant behaviour. If we can learn to continually
stop ourselves and reflect when motivated to engage in emotional eating
behaviour, and to deliberately engage in a substitute behaviour, then we may (over
time, with effort) succeed in disengaging eating behaviour from the experience of
negative emotion, and to instead adopt a new way of coping.

Similarly, Descartes tells us to train ourselves in such a way that we can
avoid acting irrationally when in the sway of strong emotion,

when we feel our blood agitated in this way, we should be on our guard
and bear in mind that everything presented to the imagination tends to
mislead the soul and make the reasons for pursuing the object of its passion
appear much stronger than they are, and the reasons for not pursuing it
much weaker….when a passion is pushing us towards some course of
action that has to be started right away, impelling us to actions that
require an immediate decision, the will should devote itself mainly to
considering and following the reasons for not acting in that way, even if
they appear less strong. (Descartes, 1985:403-404)

However, it should be noted that attempts to change the behaviour
associated with emotions are often not the most promising way to deal with
emotionally-driven irrational behaviour. Even in the case of emotional eating, it
might be argued that the problem is not so much that negative emotion causes
inappropriate eating behaviour, as it is that eating behaviour has been learned as a
form of emotional regulation as an attempt to alleviate the emotional distress.
That is, we do not (only) want to break the connection between the experience of negative emotion and an ‘output’ behaviour, we also want to unlearn this behaviour as an attempted form of emotional regulation.

More importantly, cognitive researchers have found that attempts to suppress the expression of emotion are generally inferior to attempts to avoid having the emotion. Attempts to disrupt the expression or behaviour associated with emotion requires a great deal of effort, and will tend to adversely affect our cognitive performance on other tasks (Faucher & Tampolet, 2008:109). As John and Gross (2004: 1305) write:

Because suppression comes late in the emotion-generative process, it requires the individual to effortfully manage emotion response tendencies as they arise continually. These repeated efforts should consume cognitive resources that could otherwise be used for optimal performance in the social context in which the emotions arise. Moreover, suppression may create in the individual a sense of discrepancy between inner experience and outer expression.

It therefore seems that the other two strategies – attempting to modify occurrent emotions, or attempting to change our emotional dispositions – are likely to be more useful in emotional management. Nevertheless, there is an important place for this strategy, and it is important that we learn the ability to regulate our emotional expressions if needed. No matter how ideal our emotional dispositions, there will always be potential situations in which it is undesirable to express our emotions or behave in their characteristic ways. And, as we have seen, the behavioural ‘outputs’ of our emotions can sometimes be maladaptive; in these cases it is best to try to replace these behaviours with more productive ones. In the
case of emotional eating, we can train ourselves to stop before we engage in the eating behaviour and ‘redirect’ ourselves to other consciously chosen behaviours that are more useful – for example, solving the problems that caused the negative emotions. Over time, the alleviation of emotional distress associated with these new behaviours will cause them to become habitual as regulatory strategies.

3. Managing Occurrent Emotions

The management of occurrent emotions involves an attempt to moderate or even get rid of an emotion once it has already arisen. We might try to do this when the emotion itself is inappropriate. Obviously it would be even better not to experience an inappropriate emotion at all, and thus training emotional dispositions (to be discussed in the next section) will be an essential way of managing our emotions. Moreover, as Faucher and Tappolet point out, “although we can have some control over our occurrent emotions, most of the long-term and deep impact we can hope for depends on the plasticity of our emotional dispositions” (Faucher & Tappolet, 2008: 106). But such change of emotional dispositions cannot happen at once, and cannot reasonably be expected to result in ‘emotional perfection’ in the sense that we have only those emotions that are appropriate. So there is still a place for attempts to manage occurrent emotions. In fact, such attempts will almost certainly be the starting point for attempts to alter our emotional dispositions.

The case of disgust at recycled wastewater is one where the emotion itself is both inappropriate and imprudent, and should therefore be either eliminated altogether or reduced in strength to the point where it does not prevent our taking the appropriate action. Although it should involve attempts to change the
disposition, this will (as noted above) start with attempts to change our reaction of
disgust at the thought of recycled wastewater.

According to perceptual pro-emotion theorists, emotions arise because of
(quasi-)perceptions of evaluative properties: because we are seeing a feature of a
situation as dangerous, or disgusting, or wonderful. Hence, any attempt to manage
our experience of an emotion should involve attempting to change our perception
of the situation as involving those evaluative features. We should, if possible, get
ourselves to stop seeing the recycled wastewater as disgusting (or to stop seeing
the piece of chocolate cake as comforting), and hopefully the inappropriate
emotion will disappear when our evaluative perceptions change.

This sort of attempt is where Stoic advice is most likely to come in useful: where an emotion is inappropriate, this is because the appraisal or evaluative
perception is false. As noted, Stoics believed that all such evaluative judgements
were false, and their solution was to learn to question appearances: that is, to
question whether the thing in question was really good or bad, as well as whether
it was appropriate to react to it. According to Sorabji, the Stoics advised that
people learn to change their evaluative judgements by “saying things to
themselves” (Sorabji, 2000:152). For example, when we are distressed, we are
supposed to say to ourselves ‘you are not the only one’. When we are angry, we
should remind ourselves that we have also committed offenses against others
(Sorabji, 2000:161-162). In the case of distress, the underlying judgement is
presumably that our situation is insupportably unfair or unable to be borne; telling
ourselves that we’re not the only ones to go through such things helps us to see
that others get on with their lives and deal with the same kind of problems.
Similarly, thinking about offenses we have committed against others undermines a
feeling of self-righteousness which might underlie anger: if we think of ourselves
as imperfect and prone to wronging others, we will be more likely to be forgiving and less likely to be angry when we ourselves are wronged. Although we might think of this as being purely prudential – a way of avoiding unpleasant or inadvisable emotions – we have to remember that the Stoics thought the underlying evaluative judgements were false. So when they advise us to 'say things to ourselves', they are really advising us to remind ourselves of the truth of the situation, to counteract the false impressions given us by our emotions. So we can try to make use of this advice in tackling emotions that really are inappropriate to their objects.

Similar advice is given by other historical theorists of emotion. For example, Descartes advocated “cognitive exercises” which trained us to be able to pay attention to different aspects of situations, so that we can “connect new thoughts with stimuli received through the senses” (Brown, 2006:35.) But Descartes also identified the importance of understanding the origin of an inappropriate emotion. In saying to ourselves that an emotion is inappropriate, it is even more helpful to be able to tell ourselves why we are nevertheless experiencing it. Brown writes that, for Descartes, the joining of new thoughts with emotions “depends on first understanding something about the causal genesis of a passion and second on techniques for reconditioning one's responses” (2006:199). This provides us with another specific cognitive technique we can apply in the case of inappropriate occurrent emotion: along with telling ourselves that the object does not really have the properties attributed to it by the emotion, we can remind ourselves of the causal story underlying our coming to experience the emotion in this inappropriate way.

Again, these suggestions are consistent with modern empirical research on emotion regulation. Wranik, Barrett and Salovey (2006: 399) tell us that,
according to the balance of evidence, ‘antecedent strategies’ such as reappraisal are more effective methods of regulation than ‘response-focused’ strategies (e.g., trying to suppress emotional expression). They emphasize the importance, in such antecedent-focused strategies, not only of reappraising the situation, but also of understanding how the inappropriate emotion was generated:

If we correctly perceive our emotions as they occur and know how they are generated by our appraisal of a situation, then we can (a) decide against expressing the emotions until we have more thoroughly evaluated the situation, and (b) ask ourselves whether we have correctly appraised the event, and look for alternative factors.

Thus, if we wish to alter our disgust at the notion of recycled wastewater, we can tell ourselves the causal story of how we came to acquire this reaction (the law of contagion, social reinforcement, and so on) and why it is inappropriate to its object as well as strategically irrational. These sorts of techniques are more likely to be helpful in the case of emotional eating, where eating is mistakenly seen as a regulatory technique. In this case, we have misattributed an evaluative property (being comforting) to the food, when it is really a property of other things with which the food has been associated in our experience. Brian Wansink cites research showing that most people’s ‘comfort foods’ become such because of “the memories linked to them” rather than the nature of the food as such. He writes, “Past associations are the most common reason a food becomes a comfort food. Some of these associations can be linked to specific individuals…or specific events….They’re also associated with specific feelings that the person likes to recall or wants to recapture” (2006:148). These associations need not be formed in childhood; they can be, and often are, formed later in life.
So we can say to ourselves that the food is not really comforting, or that it is unhealthy, as the Stoics suggest. But if we properly understand the origin of the misattribution, we can also tell ourselves that we are only attributing comforting qualities to the food because of an association with other aspects of previous situations (being taken care of by our mothers, being amongst friends, etc.). This can help us to devise new strategies for behavioural diversion and more successful alleviation of emotional distress.

There is in fact evidence that such techniques may work, even without very much effort. In particular, understanding the real reasons for our occurrent emotions – by deliberately making salient the feature that is actually responsible for them – seems to automatically lower the chances of misattribution and thereby lowers the chances we will make irrational decisions based on that misattribution. For example, Clore (1994:106) describes a study in which subjects were given life satisfaction interviews, either on a cold and rainy day, or on a warm and sunny day. Those subjects interviewed on the warm and sunny day tended to report greater happiness and life satisfaction. However, when the weather was mentioned by the interviewer, in order to make this factor salient as a cause of their moods, this effect disappeared. Thus, when the subjects could identify the weather – something irrelevant to overall life satisfaction – as the explanation for their mood, their moods had less influence over their subsequent decisions about whether they were satisfied with their lives.

The Stoics thought that deliberately ‘telling ourselves things’ about the way things really are would help us to stop feeling inappropriate emotions. We can see that there is some empirical support for this idea. Nevertheless, we should not overestimate the power of these techniques to eliminate inappropriate emotion. It is a well known fact about emotions that they tend to be ‘recalcitrant’ to
judgements: that is, they do not always change in line with our beliefs, so that simply removing the belief that the emotion is warranted will not always result in removing the emotion. This is especially the case when the occurrent emotion is the result of an ingrained disposition to react in that way. As Nussbaum notes (2001:114-119), emotional dispositions are very ‘tenacious’: once they have been learned, they are very difficult to unlearn. In such cases, we would not expect that telling ourselves things would work as a way of moderating or removing occurrent emotion. The suggestion above for alleviating disgust at recycled wastewater, is likely one such case: attempting to reappraise the situation and remind ourselves of the origin of our disgust will probably not enable us to down a glass of recycled wastewater right then and there. However, repeated efforts to alleviate our occurrent disgust might eventually work to alter our emotional disposition towards disgust at recycled wastewater.

4. Changing dispositions to feel emotions

The most successful strategy for avoiding emotionally-driven irrational behaviour (at least where it results from emotions that are inappropriate, rather than merely behaviour that is inappropriate to the emotion) is likely to be that of changing our emotional dispositions, so that we are no longer disposed to have inappropriate emotional reactions. However, this is a longer-term project, and will require effort and persistence. Nevertheless, according to perceptual pro-emotion accounts, it is possible to change emotional dispositions: according to de Sousa, we can even question and change our paradigm scenarios if we come to believe they are not in fact appropriate to their emotions. Historical theorists of emotions also believed that changing emotional dispositions was both possible and important.
As noted, the attempt to change emotional dispositions will likely begin with attempts to moderate or change occurring emotions. Essentially, the goal is to change one’s automatic evaluative appraisals or perceptions of certain kinds of situation. In de Sousa’s terms, we are trying to change the perception that certain features of a situation bear a relevant similarity to the paradigm scenario of an emotion, and therefore to change the perception that that emotion is an appropriate response to those features. According to Damasio, we are trying to remove a somatic marker from a particular feature. (We could also be trying to add a somatic marker or to start perceiving certain features as relevantly similar to a paradigm scenario).

Sorabji writes that “students in the Stoic tradition were trained to question appearances by continually asking themselves questions about what they see, or experience,” and that this training was reinforced through behavioural exercises which aimed at halting automatic behaviour caused by emotion. (Sorabji, 2000:215-216). This is likely to be useful training for us in everyday life as well, but if we do not agree with the Stoics that anything not under our voluntary control is indifferent (neither good nor bad), we can concentrate on learning to question appearances only in the case of particular emotions we have come to identify (or suspect) as inappropriate.

Descartes had a similar suggestion for training emotional dispositions. He thought that a “remedy for all the disorders of the passions” was to cultivate the virtue of “generosity.” Those who have this virtue will have “complete command over their passions” (Descartes, 1985:385). As a virtue, generosity is a “habit of the soul” which disposes the soul to “have certain thoughts” (Descartes, 1985:387-388). Essentially, generosity for Descartes appears to be about placing appropriate value on things: in other words, once we know what is really good or bad, and
exactly how good or bad it is, we will not be troubled by emotions being elicited inappropriately. There is not very much advice about exactly how to get into this state, but as noted above, the idea seems to be that of habituation by continual practice.

In the case of emotional eating, then, we can keep telling ourselves that the behaviour is inappropriate and irrational, as well as deliberately diverting ourselves to other, more useful, behaviours. If we persist in doing this, we will eventually attenuate the attraction and habitual pull of eating behaviour as a response to emotional distress. We will also habituate ourselves to the new behaviours that are more useful.

The attempt to change emotional dispositions is also likely to work in the case of recycled wastewater. As noted above, disgust is a very strong and highly recalcitrant emotion, so we are not likely to succeed in getting rid of any individual occurrence of disgust by telling ourselves that its object is not really disgusting, or that we are only disgusted because of social and cultural influence. On the other hand, we might eventually succeed in changing our disgust reactions through a long process of doing this repeatedly, over time. We might also exercise our imagination in various ways (e.g. imagining ourselves drinking the water, trying to picture clean pure sparkling glasses of water when thinking of recycled wastewater, imagining all of the impurities being removed) so as to try to rid ourselves of the automatic association with disgustingness and contamination. And we might slowly expose ourselves to recycled water itself: visiting a water treatment plant, looking at the treated water, smelling it, touching it, and eventually drinking it – so as to expose ourselves to situations involving recycled water which are not associated with elicitors of disgust.
There is some evidence that this strategy will work. Gross and John (2004: 1306) cite studies which found that while emotional suppression (trying to inhibit the expression of emotion) had no effect on subjective emotional experience, reappraisal “decreased both expression and subjective experience, and did not increase physiological activation.” This means that attempting to reappraise the situation does in fact attenuate how much the emotion itself is felt, as well as our tendency to express that emotion; moreover, it does not seem to require a great deal of cognitive effort. Thus, even if this is not sufficient to avert irrational behaviour in the case of a particular episode, it may be that repeated use of reappraisal will eventually change our emotional dispositions.

Nevertheless, the idea of individuals trying to change their disposition to feel disgust at recycled wastewater seems rather far-fetched. (It is more plausible to suppose that individuals might try to change their disposition toward emotional eating). This is because most people either believe their disgust reaction to be appropriate (as the Stoics put it, they do not question the appearances their emotions put in front of them), or if they do not, they do not see it as so inappropriate or undesirable that it needs changing. If people are not likely to work on their own dispositions toward disgust, then it will not matter much whether such methods are available and effective: we still will not be able to avoid the consequences of widespread emotionally-driven irrational behaviour. As a result, it is important that we also consider how we might manage emotionally-driven irrational behaviour from a public policy viewpoint – that is, how we might try to regulate the emotions of others.

5. Emotion regulation in public policy
The suggestions and techniques discussed above are all about what *individuals* can do to deal with and manage emotionally-driven irrational behaviour. However, an equally – if not more – important question is how we might deal with emotionally-driven irrational behaviour on a *public policy* level. After all, such behaviour can have widespread negative effects in society, whether in the form of public health consequences of unhealthy diet and overeating, or in the form of general public opposition to sustainable water management policies. There is therefore good reason to think about how public policy, or social and cultural influences more generally, might help or hinder people in trying to address emotionally-driven irrational behaviour.

Moreover, there is another potential role for public policy in helping people to recognize that their emotional responses, or behaviours, are irrational. In the case of much emotionally-driven irrational behaviour, people may not recognize that the emotion is inappropriate to the situation, or that the behaviour is inappropriate to the emotion. The case of recycled wastewater is a good example of this. While it is possible that people are being led by their disgust to oppose wastewater recycling while also sincerely believing that recycled wastewater is safe and not *really* disgusting, it is more plausible to suppose that public opposition (and that of activist groups) is due to people’s not being prepared to accept that the wastewater is not disgusting. People will not attempt to change their emotional reactions unless they are convinced that the reactions are in fact inappropriate. As Menegaki et al note, one of the most effective methods of reducing opposition to recycled wastewater would be to increase people’s exposure to it (Menegaki et al, 2009:286); this is consistent with our discussions above. However, people will not increase their exposure unless they first believe there is a reason to do so. Where there is a negative social effect resulting from widespread
emotionally-driven irrational behaviour, then, it is important to first persuade
people that the behaviour is both emotionally-driven and irrational. Alternatively,
we can use other, less direct means to manage people’s emotions (for example, by
secretly replacing all of the tap water with recycled water and only telling people
years later), but this would be politically problematic!

In this context, it is important to re-emphasize the crucial role in
perceptual pro-emotion theories of learning in the formation of our emotional
dispositions, and the fact that this learning is not limited to childhood but
continues throughout life. Perceptual pro-emotion theories therefore acknowledge
the importance of our culture, and the influence of others in forming, modifying or
reinforcing our emotional dispositions. We learn that wastewater is disgusting, or
that eating certain foods is comforting, not only by what our parents say and do,
but through wider social influences including peers, leaders or other prominent
members of society, or perfect strangers.

The role of the media will be particularly influential in most modern
societies, since it will influence our perceptions of what ‘other people’ think,
believe, or take to be appropriate. This means that the media can reinforce
emotional dispositions by representing certain emotional reactions as
inappropriate, or by drawing attention to features of situations which are likely to
elicit certain emotional responses. Martha Nussbaum writes in relation to
empathy and compassion that,

obviously television and the other mass media are also potent educators of
citizens and can nourish empathy or obtuseness, appropriate or
inappropriate compassion….Its choice of images and roles, in news stories,
advertising, and drama, will have important consequences for citizens’
moral abilities, for better or worse….If we think not just about empathy
but about getting the judgments right, we can see that the media have considerable power in that regard as well, portraying calamities as more and less grave. (Nussbaum, 2001:434)

In relation to emotional eating, Brian Wansink notes the tendency of magazines to feature pictures and articles about ‘comfort foods’, and to promote what he calls “common comfort-food myths,” namely that comfort foods must be unhealthy, that they are a common reaction to sadness, stress or boredom, and that comfort food preferences are fixed in childhood (Wansink, 2006:139). Such messages are also heavily reinforced in popular entertainment, for example in the common TV and movie trope of women eating ice cream or chocolate to cope with a break-up. And they are also clearly promoted by food companies in their advertising, portraying certain kinds of (usually fast/junk) foods as being associated with cheerfulness, fun, relaxing, and the presence of family and friends.

In relation to wastewater recycling, the media has played a prominent role in drawing attention to the ‘disgusting’ features of recycled wastewater, and in normalizing the disgust reaction. For example, Menegaki et al point to the use of the slogan “toilet to tap” by opponents of recycling schemes (Menegaki et al, 2009:286). Similarly, an article in the Singaporean media blames the failure of Australian wastewater recycling schemes on the mass media’s tendency to highlight the ‘yuck factor’ (Gazello, 2001). The emphasis on disgust-eliciting features, along with the implication that disgust at recycled wastewater is appropriate and normal, is likely to reinforce the irrational emotional response, along with the belief that the water is in fact disgusting. Thus, not only will it be harder to convince people that their disgust reactions are inappropriate; it will also be hard for people to move from this new conviction to an actual change in their
disgust response, since the response is continually being elicited in connection with recycled wastewater and being exhibited by other people.

Historical theorists did not tend to focus much attention on how people might influence the emotional dispositions of others. The exception is Aristotle, who dealt with the eliciting and management of an audience’s emotions in the Rhetoric. This might be seen as a form of non-rational manipulation. However, on the perceptual pro-emotion view it need not be so. A speaker can arouse emotions in listeners by drawing attention to certain features of a situation which elicit certain emotions. Where the emotional response is irrational, and attention is drawn to the eliciting features because the speaker knows that the emotion has been inappropriately associated with those features (or that an emotion will be elicited which causes the audience to overlook other, more important features), then the speaker is manipulating the audience. However, the speaker might also draw attention to features of a situation which quite properly elicit certain emotions as an appropriate response to the situation, or which lessen the emotions they had previously been feeling. In this case the speaker would be drawing attention to relevant evaluative features of the situation that had been overlooked – perhaps because inappropriate emotions were causing the audience to pay attention only to irrelevant or less important factors.

This focus on the ways in which we can regulate the emotions of others is likely to be crucial in public policy responses to emotionally-driven irrational behaviour. In this, the perceptual pro-emotion theorists have a useful contribution to make. The default policy response to widespread irrational behaviour tends to be an attempt to ‘educate’ people that the behaviour is indeed irrational: for example, that the recycled water really is clean, or that eating junk foods is unhealthy. As we have seen, convincing people that their behaviour is irrational
will have to be part of the story, since people will not try to change behaviour that they perceive as justified and appropriate. But the perceptual pro-emotion theory also helps us to see why a simple ‘education’ strategy is unlikely to succeed.

Firstly, emotional management strategies must deal with not just the behaviour, but the emotion underlying the behaviour: that is, they must convince people not just that their behaviour is irrational but that the emotion driving it is inappropriate. This is likely to be difficult, and not only because of the phenomenon of emotional recalcitrance. If people are convinced that others also react that way, and if the emotional reaction is continually elicited by the media, activist groups, or political leaders, then people will not be motivated to work to change their disposition toward the relevant emotional responses, even if they come to believe that the resulting behaviour is irrational. We must therefore persuade people that their emotional responses are inappropriate, and also put in place policies and strategies that help them in their efforts to change their emotional dispositions.

The question now is; how can we do this? As Sorabji notes, Aristotle’s advice to orators was to try to get rid of one emotion by arousing an opposing one, “knocking out anger by pleasure, and vice versa” (Sorabji, 2000:221). This meshes well with de Sousa’s observation that deliberately paying attention to something is much easier than deliberately not paying attention to it, and therefore that “we may expect great difficulty in trying to get rid of an unwanted emotion, but more success in working ourselves into one” (de Sousa, 1987: 243). Thus, if people tend to experience an inappropriate reaction, we could try to dilute it or force it out by eliciting another, more appropriate, opposing emotion. This opposing emotion could be directed towards the same object, or it could be directed towards a different object, so long as experiencing it tends to drive out
the original inappropriate one. Over time, if this is done consistently, or if it is culturally reinforced, the original emotion might become disconnected from the object.

For example, the Singaporean article mentioned earlier attributes the general acceptability of recycled wastewater use in Singapore to the fact that media messages are almost uniformly positive, and do not tend to mention the ‘yuck factor’. What is equally informative, however, is the tone of the article: the article clearly exhibits (and implies it is appropriate to experience) pride in the success of the wastewater scheme (noting that Australians are asking Singapore for advice), and emphasises the contribution of the scheme to Singapore’s goal of being self-sufficient in water. Whether or not this was the intention of the article, we can see that eliciting emotions such as pride would be a good way to dampen disgust reactions. If people are being encouraged to consider the fact of self-sufficiency, and being world-leading, succeeding where others have failed, then they are less likely to allow their disgust to control their responses. Pride need not be sufficient to completely overcome the disgust response: as noted above, if people were to be willing to gain familiarity and exposure to recycled wastewater, they will eventually learn from experience not to be so disgusted by it. Thus, if pride is just sufficient to interrupt the link between experiencing disgust and the resulting behaviour (opposition to recycled wastewater), it would allow people to learn more about recycled wastewater: for example, by visiting the treatment facility (which in Singapore is a tourist attraction). Finally, if a strong message is sent via media and other social influences that disgust is not an appropriate reaction, people will be more motivated to try not to express disgust, and to try to alter their dispositions so as not to feel disgusted.
Of course, there is a limit to how much such messages can be controlled by public policy. Martha Nussbaum notes the problem of market forces in the mass media, which encourages focus on any aspects of a situation that are likely to appeal to audiences, and which discourage responsible use of media to promote good deliberation (2001:434-5). Nussbaum thinks that popular pressure might be able to overcome this problem. However, at least in the case of recycled wastewater there is a chicken-and-egg problem: until people are persuaded that their disgust is inappropriate, they will not put pressure on the media to stop reinforcing disgust; but until the media stops reinforcing disgust, people will not be persuaded that their disgust is inappropriate.

Perhaps the best that can be done is to attempt to counter existing media and cultural reinforcers by expanding current educational efforts to place more emphasis on understanding emotional responses - both the origin of those responses, and the fact that they are inappropriate – as well as suggesting other, more appropriate, emotional responses. Politicians and other public figures could present an example, by publicly drinking recycled wastewater; and the general public could be encouraged to visit water treatment facilities, and/or to see, touch and taste the recycled water. The message of environmental sustainability and taking care of future generations (‘our grandchildren’) could be reinforced.

In the case of emotional eating, the idea of regulation to stop the reinforcement of misattribution is more promising. In the case of emotional eating, many people already see their behaviour as undesirable, and can also be easily persuaded that it is a result of misattribution of ‘comforting’ characteristics to food which is not really comforting. On the other hand, the continual reinforcement of the message that eating behaviour is an appropriate and normal response to distress may undermine people’s efforts to change their dispositions. Regulation of
junk food advertising may help to remove some of this reinforcement, by not focusing attention on the possibility of eating to alleviate emotional distress, and campaigns which emphasise alternatives (for example, going for a walk, or phoning a friend) might help to introduce new, more constructive, emotional associations.

**Conclusion**

Within this thesis I have accepted the central pro-emotion claim that emotions are necessary for rational decision making. In particular, I have argued that perceptual pro-emotion theories provide a satisfactory account of the role of emotions in practical rationality, while also explaining how emotional regulation by reason is possible. This gives us a number of suggestions for how we might engage in emotional management both of ourselves and of others. It should be possible, on the perceptual pro-emotion view, to regulate our emotional behaviours, our occurrent emotions, and our emotional dispositions. In this way, we can harness the helpful roles played by emotions in rational decision, while avoiding as much as possible of their harmful potential.

I have also tried to show that the modern pro-emotion consensus is not in fact an overturning of a ‘traditional’ consensus which took emotions to be entirely detrimental to rationality. Rather, there is wide variation in historical theories. However, many of them say quite similar things to the modern pro-emotion theorists concerning the functions of emotions in practical rationality. Moreover, the insight which is widespread amongst historical theorists – that emotion will tend to cause irrational behaviour unless controlled and regulated by reason – is
compatible with the central pro-emotion thesis; it can and should be accommodated within pro-emotion theories.

Strategic theories and Evolutionary Psychology theories are, I have argued, unable to meet this challenge. Strategic theories offer only a partial account of the role of emotions in practical rationality, and their account of how emotions cause irrational behaviour raises doubts about their ability to remain committed to the central pro-emotion thesis. Evolutionary Psychology theories offer a more complete account, but their explanation of emotionally-driven irrational behaviour is not consistent with the empirical evidence, and – like the strategic theories – there is a tension between this explanation and their supposed commitment to the central pro-emotion thesis. Moreover, their commitment to modularity prevents them from being able to accommodate the degree of flexibility of emotional response, and influence of environment, experience and reason, that is required for emotions to be helpful to practical rationality.

Fortunately, perceptual pro-emotion theories can provide a complete and satisfactory account of the function of emotions in practical rationality, which is consistent both with empirical evidence and with the central pro-emotion claim. Moreover, perceptual theories allow for a significant degree of flexibility in both the input and output sides of emotional response, which means that emotions can be responsive to learning, experience and reason in a way which allows them to be on balance helpful to rational decision-making.

In this final chapter, I have argued that the perceptual pro-emotion account has some useful suggestions to make regarding the management and regulation of emotions on both a personal and a public policy level. These suggestions are drawn from the theoretical discussion, and have not been proven to work;
however, I have given some reason to think that these suggestions are consistent with work that has been done on emotional regulation in the cognitive sciences.

Finally, I have tried to show that reconciliation between modern pro-emotion theorists and some of the historical theorists of emotions is in order. Many historical theorists gave accounts of emotions that are compatible with the perceptual pro-emotion theory. More importantly, their focus on the need to regulate emotions may provide some useful insights in our own attempts to devise strategies for managing emotionally-driven irrational behaviour. It is my view that the modern pro-emotion consensus members (or at least, the perceptual theorist members of this consensus) are correct regarding their view of the functions of the emotions in practical rationality; however, it is wrong to think that they are overturning centuries of historical thought. A more nuanced view of both the modern pro-emotion consensus, and of historical theories, will likely prove useful in our attempts to manage emotionally-driven irrational behaviour.
Bibliography


146


London.


Harlow, J. M. (1868) ‘Recovery from passage of an iron bar through the head’.


Philosophy,


Accessed April 2012.


