An interdisciplinary approach to belief formation and resistance to change

Peter Frank Thompson

A thesis in fulfilment of the requirements for the degree of
Doctor of Philosophy

Please note: This thesis is currently under examination.
As it may change, please refrain from citing it until the final version is released.

School of Humanities and Languages
Faculty of Arts and Social Sciences

August 2014

THE UNIVERSITY OF NEW SOUTH WALES
Thesis/Dissertation Sheet

Surname or Family name: Thompson

First name: Peter | Other names: Frank

Abbreviation for degree as given in the University calendar: PhD

School: Humanities and Languages | Faculty: Arts & Social Sciences

Title: An interdisciplinary approach to belief formation and resistance to change

Abstract

One of the most remarked on, and perplexing, aspects of beliefs is their resistance to change as discussed by researchers such as Frijda, Mesquita, Manstead and Bem, amongst others. This is especially the case for false beliefs. The thesis examines the many influences on belief formation in general, but with a special interest in false beliefs. While the basic definition of belief used is partly derived from philosophical literature, including Clifford, Russell, Goldman, Garfield, and Morton, it is shaped by an interdisciplinary approach to understanding belief formation and resistance to change, predominantly involving cognitive psychology, neuroscience and social psychology. An extensive review of literature on heuristics and biases—particularly the work of Tversky and Kahneman, and Gilovich and Griffin—reveals effects on perceptions that affect beliefs, and leads to a similar review of the interactions of emotions and feelings with beliefs. The 'feelings-as-evidence' hypothesis proposed by Clore and Gasper emerges as having significant explanatory power with respect to feelings contributing to establishing and maintaining beliefs. To further understand possible mechanisms at play, Antonio Damasio's definitions of emotion and feeling, and his somatic marker hypothesis, are proffered as a solid foundation for examining how emotion and feelings interact with the numerous influences on beliefs and how they contribute to feelings being such a dominant source of 'evidence' used by people in a lay context to establish that, for them, a belief is true. On the basis of this foundation and literature on levels of consciousness (e.g. Damasio, 2010; Gallagher, 2000; Northoff et al., 2006; Panksepp, 2005, 2012; Panksepp and Northoff, 2009), the thesis develops the proposal that emotions and feelings act through a lens of self-enhancement and self-protection processes to support beliefs that protect a person's neural construct of their 'self' with feelings being the key perceptual indicator of whether a belief serves the self-construct or not. If a belief, whether true or false, is experienced as serving the interests of the 'self', then it is likely to be highly resistant to change.

Declaration relating to disposition of project thesis/dissertation

I hereby grant to the University of New South Wales or its agents the right to archive and to make available my thesis or dissertation in whole or in part in the University libraries in all forms of media, now or here after known, subject to the provisions of the Copyright Act 1968. I retain all property rights, such as patent rights. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

I also authorise University Microfilms to use the 350 word abstract of my thesis in Dissertation Abstracts International (this is applicable to doctoral theses only).

....	....	....
Signature	Witness	Date

The University recognises that there may be exceptional circumstances requiring restrictions on copying or conditions on use. Requests for restriction for a period of up to 2 years must be made in writing. Requests for a longer period of restriction may be considered in exceptional circumstances and require the approval of the Dean of Graduate Research.

FOR OFFICE USE ONLY Date of completion of requirements for Award:

THIS SHEET IS TO BE GLUED TO THE INSIDE FRONT COVER OF THE THESIS

ORIGINALITY STATEMENT

'I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.'

Signed .....

Date .....

Contents

Table of Figures

Acknowledgements

Who in their right mind would return to university as a mature-age student while running their own business? More than that, while running a business—training, facilitation and coaching—that is at the mercy of business fashion and spending cycles? And, during a time that included the years of the 'great financial crisis'? Answer: me.

While I could anticipate difficulties juggling workload, I wasn't deterred because I could see the relevance of the research to my work. Also, I have been infected with an irrepressible love of learning. Still, even with these pre-existing support structures, one needs encouragement from others. Dr Peter Slezak is to be thanked for many things but especially for: encouraging me to enrol in the first place; sharing his enthusiasm for enquiry and for precision in thinking; demonstrating knowledge across a number of disciplines (a style after my own heart); and for his support through the particularly challenging times when demands of life and running a business were competing with the research. Dr Anthony Corones was similarly supportive through these times. I also owe particular thanks to Tony for assisting me in navigating the administrative requirements through the years. However, I am most indebted to Tony for his assiduous reading of my work through the years and the advice that flowed as a result.

I am unlikely to be alone in muttering about the tasks required for annual progress reviews. The Postgraduate Research Coordinators have the unenviable task of guiding the postgraduate research student through this process: encouraging but also being vigilant to areas they consider necessary to ensure success. For this I must thank Dr Anthony Corones (who served not only as co-supervisor, but also as Postgraduate Research Coordinator for a time), Dr Mina Roces, Dr Geoff Nathan, and Dr Stefania Bernini. I am also indebted to Sally Pearson for her patience and charming assistance each year with administrative requirements.

Introduction

In reviewing three books about the threats to public health from 'vaccine deniers', Chris Mooney joins a growing lament from some quarters about the perceived growth of 'cognitive relativism' and 'truthiness'. The books, he claims, are more than a 'wake up call'; they are:

> … a call to arms against the broader phenomenon of tilting against reality, or making up one's own version of it, and clinging to it fiercely despite all evidence and consequences – a condition also referred to as human nature.
>
> Irrationality can be a very dangerous and communicable disease – and we still don't know how to adequately inoculate against it.
>
> (Mooney 2011, p. 47)

Such exhortation is indicative of the passions that strongly held beliefs arouse. On the other hand, it may be that strong passions give birth to the beliefs—and this points to one of the key influences on the formation and maintenance of beliefs that becomes a central focus of this thesis.

Mooney is likely to have found himself in sympathy with W. K. Clifford (1999) whose concern for the basis of a belief leads to this claim:

> The question of right or wrong has to do with the origin of his belief, not the matter of it; *not what it was, but how he got it*; not whether it turned out to be true of false, but whether he had a right to believe on such evidence as was before him.
>
> (ibid., p. 71, emphasis added)

Such a stringent view of belief may appeal to a passion for logic. However, there are beliefs that many might argue are of no consequence if they satisfy the holder of the belief and cause no harm to others. But, Clifford does not view a belief as some trivial thing aimed primarily at providing comfort and solace for the individual. In his view each one of us has a duty to humankind, and erroneous or ill-proved beliefs work against that duty:

> But forasmuch as no belief held by one man, however seemingly trivial the belief, and however obscure the effect on the fate of mankind, we have no choice but to extend our judgment to all cases of belief whatever. Belief, that sacred faculty which prompts the decisions of our will, and knits into harmonious working all the compacted energies of our being, is ours not for ourselves, but for humanity.
>
> (ibid., p. 74)

It is easy to imagine Mooney (2011) chiming in here with something like *'Yes, see, that's exactly the point – the vaccine deniers risk harming not only their own children but also whole*

societies'. Such a point might be easy to establish; however, given the *range* of type and complexities of beliefs, others might point to beliefs with little apparent consequence to any other than the holder of the belief. They could argue, for example, that there is little effect on society if a person believes broccoli to be a more palatable choice of phytonutrients and antioxidants than is spinach, even though this position is taken with little or no supporting evidence.

William James (1956) acknowledges Clifford's exhortation to ensure that beliefs are not founded on insufficient, or no, evidence as seeming 'healthy' (ibid., p. 8). However, he goes on to suggest that it flies 'in the teeth of the facts' (ibid.) to imagine that once wishful thinking and 'sentimental preference' (ibid.) have been removed all that remains is 'pure reason' (ibid.) to settle opinions. In this, James demonstrates a deep insight into—or, perhaps more accurately, prescience about—the nature of the processes of belief formation and maintenance that we are now beginning to understand from the developments in cognitive psychology and cognitive neuroscience. James considers it evident that:

> … our non-intellectual nature does influence our convictions. There are passional tendencies and volitions which run before and others which come after beliefs, and it is only the latter that are too late for the fair; and they are not too late when the previous passional work has been already in their own direction.
>
> (ibid., p. 11)

For James, in practice, *'one's conviction that the evidence one goes by is of the real objective brand, is only one more subjective opinion added to the lot'* (ibid., p. 16). He goes on to claim that he is not shunning an empirical approach to human knowledge or the 'chase for truth' (ibid., p. 18), but he cannot go with Clifford who, he considers, would prefer to believe nothing rather than risk believing on the basis of insufficient evidence. James suggests:

> We must remember that these feelings of our duty about either truth or error are in any case only expressions of our passional life. Biologically considered, our minds are as ready to grind out falsehood as veracity, and he who says, 'Better go without belief forever than believe a lie!' merely shows his own preponderant private horror of becoming a dupe.
>
> (ibid., p. 18)

More recently, Antonio Damasio highlights the profound role that beliefs—however they may eventually be defined—play in the lives of individuals, societies, and countries:

> ... we tend to hold beliefs about certain kinds of matters—mostly, as it turns out, about matters of life and death. The proper subject matter [for a belief] is almost any topic that gravely affects our well-being, in both the physical and spiritual senses.

> This range includes matters of moral behavior, matters of faith, matters of life ... Lastly, belief is about matters that have to do with the idea we hold of ourselves.
>
> (Damasio 2000b, p. 326)

Damasio (ibid., p. 300) suggests that there 'are a number of possibilities for studying belief from a neurobiological point of view' and hopes (ibid., p. 332) that 'our understanding of the cognitive neuroscience of belief will make considerable progress as we advance our understanding of the underpinnnings of memory, of emotion and feeling, and of the biological nature of the self'.

The times appear to be ripe for such advancements. As Damasio points out:

> This is a time of explosive developments in neuroscience and cognitive science. It is likely that such developments, along with the refined scholarship the humanities are offering, will permit new progress on the issues of belief and memory ...
>
> (ibid., p. 325)

This research does not address psychotic delusions. Rather, it focuses on the everyday irrational—and therefore potentially limiting—beliefs that 'normal' people seem to be so adept at generating for themselves. The field of 'normal' beliefs and their departure from normative rationaliy has become a major field of research in cognitive science in the tradition of 'biases and heuristics' following the work of Kahneman and Tversky (for example: Kahneman 2003a; Kahneman 2003b; Kahneman & Tversky 2000; Kahneman *et al.* 1982). Besides its intrinsic theoretical importance such work may have implications for counselling as well as for education and societal development.

Central to this project are the 'beliefs' that people generate, maintain, and even strengthen that have the following characteristics:

- Each is expressed as a *belief*[1]. That is, it is thought to be the 'truth' regarding a particular situation.
- It is a belief:
  - which is clearly falsifiable by observable evidence **or**,
  - to which an equally logical and plausible counter position can be proposed, in the absence of observable falsifying evidence
- The belief is maintained despite the clear presentation of contradictory evidence.

1. Note that we still have not attempted to define belief. It is being used here, again, in the common lay sense.

Of particular interest are those beliefs that cause concern, difficulty and disappointment in a person's endeavours. Typically, these are beliefs[2] that relate to such areas as:

- their perception of their efficacy in some endeavours. For example:

  I simply can never get myself organised to arrive on time.

  I always seem to make some mess with my reporting.

  No matter what I try, I always seem to say something embarrassing.

- their perception of their potential (or lack of it). For example:

  I would never be able to learn to be a good manager.

  I'll never be able to understand how to use Excel.

- their perception of themselves. For example:

  I'll always be shy.

  I'm just hopeless in relationships.

  I am not arrogant, I'm just assertive.

In daily discourse, such expressions—whether we ultimately term them 'beliefs', or not—are relatively common. Often, they might occur as a passing comment and be treated as such. Where they do become significant is when they are expressed in the context of someone seeking advice, encouragement, reassurance or solace. Typically, this might be with some confidant—friend or professional. The focus here is on the potential use of the research in some professional context (for example, coaching, counselling, or education).

The research focuses on eliciting greater understanding of how someone can maintain a belief which is falsifiable (especially if they express a wish that whatever is the predicate of the belief were not the case). The future value of this research may lie in possible new approaches to effectively assisting people in removing self-imposed limitations which affect the quality of their work and lives.

For example, consider the person who says *'I'll never be any good as a manager but I wish I could be'*. On inquiry we find that they would like to perform well as a manager for some of the usual reasons: better career prospects, improved financial circumstances, a greater number of better choices for such activities as holidays, and so on. We might also learn that they have good literacy and numeracy skills and that it is easy to demonstrate that to them; yet they maintain their insistence that they would never succeed as a manager. It is pointed out to them that they have excellent 'interpersonal skills' and that, with those and

2. However these might ultimately be defined. Here, the term used in the general lay sense.

good literacy and numeracy, all the requisite management skills could be easily acquired and developed by them.[3] Yet, the person insists on the inevitability of their failure.

Those working in areas such as education and coaching are well familiar with the debilitating effect that beliefs such as those outlined above can have on people in their efforts to advance in life—particularly those beliefs relating to the *self*. While there are existing psychotherapeutic processes for changing beliefs, the enterprise here is to bring additional scientific rigour to the area and to open the way for non-faddish interventions that are easily and appropriately adaptable to such contexts as schools, business training, and executive coaching.

As might be anticipated there is a broad range of influences on the formation and maintenance of beliefs. This range of influences extends the investigation beyond the solely philosophical. As is argued by both philosophers and psychologists[4], understanding the complexities of belief formation and maintenance requires an interdisciplinary approach in order to elucidate in more detail this intriguing area of human cognitive and emotional behaviour. Therefore, on the advice of great names in the various fields, an interdisciplinary approach is taken. Some of the connections between the various disciplines are relatively obvious; others are not. Moreover, the hypothesis developed relies to a large extent on the relatively recent, but burgeoning, interest in emotions. This field of study is still very active in even agreeing on definitions while simultaneously seeking to advance our understanding of influences on beliefs and gaining ground on understanding the underlying mechanisms for belief formation. Because of the very active debate in this area, both it and the other areas addressed in this thesis are subject to considerable literature review. The literature reviews not only discuss the debates that are germane to this thesis but, more importantly, are required in order to establish the grounds for selecting the key hypotheses and models developed in this thesis. These hypotheses and models are also used to draw together elements from disciplines that might otherwise seem unrelated.

Outline of Chapters

Chapter One opens with acknowledgement of the focus that philosophy has on beliefs and examines aspects that are relevant to this thesis. The investigation shows that to improve our understanding of the role of beliefs in affecting actions and decisions it is useful, and arguably necessary, to assume an interdisciplinary approach. In doing this, the

3. This would be justifiable on the basis of the experience of other people with similar attributes and abilities having developed the requisite skills for managing well.

4. In Chapter 1

exhortations of philosophers such as Garfield (1988) are accepted. Consequently, the chapter is able to present a definition for beliefs that incorporates a principal element from philosophical discussion but which also includes two elements that address the psychological interest in the role of beliefs.

The chapter then briefly addresses debate over the difference between knowledge and belief and it notes the conflation of these terms in general lay use. Two criteria for 'knowledge' are accepted but what emerges, particularly from taking the psychological point of view, is the need to view 'knowledge' from the perspective of the speaker. Throughout this part of the investigation, the argument is strengthened for an interdisciplinary approach with support from both the philosophical side (for example, Russell 1921/2005) and the psychological side (for example, Eichenbaum & Bodkin 2000).

Given the accepted interdiscplinary approach, Chapter Two begins at a natural starting point (from both an epistemological and psychological viewpoint)—perceptions. It is noted that people, especially in dispute, will refer to their claimed perceptual accuracy as a means of establishing 'truth' and 'knowledge'. Mechanisms that lead to different perceptions, and consequently different beliefs, amongst people are discussed: for example, genetic differences affecting perceptions, illusions.

As well as perceptual differences and errors, Chapter Two extensively reviews the profoundly influential work of Daniel Kahneman and Amos Tversky that exposes the biases to which humans are subject in making judgments. This is developed to argue that many of these judgments are, in fact, beliefs.

The chapter then moves to briefly discuss the interaction between perception and emotions and to show that perceptions can be affected by emotions, beliefs and language. From this, it establishes the need to examine in detail the role of emotions. What is implicit in Chapter Two is the requirement that for someone to rely on their perceptions and judgments there must be a belief (albeit unconsciously held) that their perceptions are reliable. This aspect re-emerges in Chapters Four and Five.

Chapter Three establishes the general agreement that emotions affect beliefs (and vice versa) and, indeed, it is the resistance of beliefs to change that suggests the need to account for the effect of emotions. This need to examine the role of emotions necessitates acknowledgement that there is an untidy range of definitions for the terms *emotion* and *feeling*, and proper examination does require relatively extensive literature review. However, the review provides a platform for arguing the case for accepting the definitions offered by Antonio Damasio which have grown in both acceptance and influence over the last 20 years. Beyond providing definitions that allow for more rigorous testing of

hypotheses, Damasio's definitions demonstrate profoundly why a purely philosophical approach may struggle to provide the depth of insight into the lay person's view of 'belief' and their consequent actions.

The chapter transitions through an examination of appraisal hypotheses to highlight the significance of feelings, particularly as defined by Damasio, in creating and maintaining beliefs.

Chapter Four follows with an examination of the feelings-as-information hypothesis. This leads to an examination of Damasio's Somatic Marker Hypothesis which, in turn, provides an hypothesis for both the strength and centrality of the influence of emotions on beliefs. Damasio's work also provides a supportable argument for the introduction of the 'self' concept and how Damasio's model for this phenomenon provides a neuroanatomical explanation of how levels of consciousness arise. Beyond explaining the genesis of levels of consciousness and the emergence of the 'self', Damasio's hypothesis provides the bridge between feelings and autobiographical beliefs which are an intrinsic component of the self. From this I develop an argument for the primacy of feelings in guiding attention and the style of cognitive processing. This argument suggests that beliefs, and belief systems, are formed and maintained to support the structural integrity of the self construct.

A significant claim emerges through development of the hypothesis presented: that the mechanisms for forming beliefs do not appear to be different for true or false beliefs. Therefore, while the initial focus of the thesis is on false beliefs, the findings apply to beliefs in general. What is suggested to be profoundly significant is that in many instances feelings are the final arbiter of truth and, by extension, reality.

Chapter Five reviews the contributions from the different disciplines and emphasises the benefit of taking an interdisciplinary approach. The suggestion is made that it is the very range of influences on beliefs—from perceptions through to existing beliefs—that is one of principal elements in rendering beliefs immune to logic in so many instances. This chapter distils the findings on the significance of feelings and, how, through the lens of self-enhancement and self-protection processes, feelings support beliefs that protect a person's neural construct of their 'self'. It is shown that the literature supports a view that there is a psychological 'homeostatic' rebalancing that will accept, or even generate, beliefs of the type given on page 10 in order to preserve a domain of concern that is more important to the overall integrity of the self than the one being sacrificed.

Finally, it is suggested that the understanding of this primacy of feelings offers much for the potential refinement and development of counselling, coaching and educative methodologies.

Chapter 1: Belief, Knowledge and Memory

In *Memory, Brain, and Belief ,* Schacter and Scarry (2000b, p. 4) acknowledge very early the significant problem with the very term *belief:*

> … we are still some distance from an adequate working definition of belief that is shared across scientists and scholars.

However, while defining beliefs is significantly problematic, there is no escaping the use of the term[5]. At this stage, the term is embedded in our language. If, in the long run, the Material Eliminativists have their way and the term is cast aside, it will be interesting to observe how people will replace utterances such as the following:

> Adams: It seems to me … that there's this feverish desire to believe … But other people do want to believe, don't they?
>
> Rushdie: They do, and I think there is a great need in people to know what they are for, as well as to know what they are against.
>
> (Adams 2007, approx. 29 min 30 sec)

> Belief can make people do the strangest things. At one level, it provides a moral framework, sets preferences and steers relationships. On another, it can be devastating. Belief can manifest itself as prejudice or persuade someone to blow up themselves and others in the name of a political cause.
>
> (Jha 2005)

While an extensive epistemological inquiry into the nature of *belief* is outside the scope of this thesis, there is a need to establish an understanding of what is meant by *belief* to satisfy both scientific and philosophical inquiry in (Schacter & Scarry 2000b), and to establish a suitable context for examining the processes that give rise to beliefs and which affect beliefs. *Belief*, however, is not the only term requiring special attention. A brief investigation of the difference between *belief* and *knowledge* is conducted because of the bearing that both these concepts have on *memory*: whether belief and knowledge are forms of memory, and how they interact with memory if (or when) they are *not* forms of memory (Eakin 2000; Eichenbaum & Bodkin 2000; Nelson 2000; Westbury & Dennett 2000). What emerges in the literature is a general consensus that there is some *relationship* between memory, belief and knowledge. Before examining this relationship closely, the manner in which 'belief' is being used throughout is examined.

5. Unless the Material Eliminativists have their way!

1.1 Belief

The prominence of beliefs in human mental activity is suggested by Russell (1921/2005, p. 139) in claiming that belief is the 'central problem in the analysis of mind' and that beliefs essentially comprise the bulk of our intellectual activity.

> The whole intellectual life consists of beliefs, and of the passage from one belief to another by what is called 'reasoning'. Beliefs give knowledge and error; they are the vehicles of truth and falsehood. Psychology, theory of knowledge and metaphysics revolve about belief, and on the view we take of belief our philosophical outlook largely depends.
>
> (ibid., p. 139)

Quine and Ullian also note the constant mental activity associated with beliefs or, as they later suggest, with *believing*:

> One's repertoire of beliefs changes in nearly every waking moment. The merest chirp of a bird or chug of a passing motor, when recognised as such, adds a belief to our fluctuating store. These are trivial beliefs, quickly acquired and as quickly dropped, crowded out, forgotten. Other beliefs endure: the belief that Hannibal crossed the Alps, the belief that Neptune is a planet
>
> (Quine & Ullian 1978, p. 9)

James (1890/1950a, p. 283) takes belief to be 'the mental state or function of cognising reality' and that it means 'every degree of assurance, including the highest possible certainty and conviction'. And, Schwitzgebel's (2010) summary of the topic of *Belief* notes that belief formation is 'one of the most basic and important features of the mind'. While these views assent to the central role and the pervasiveness of beliefs in human mental activity, they have not approached a definition of *belief* that serves for this thesis.

Dennett (1990, p. 91) indicates a direction from a folk psychology perspective which, he suggests, has beliefs as 'information-bearing states of people that arise from perceptions and that, together with appropriately related *desires*, lead to intelligent *action*'. In this, the connection between belief and actions is introduced; a connection which is discussed in more detail later. Dennett goes on to pose questions relating to the nature of beliefs: for example, can they be held by non-human animals; does language have a role in beliefs? For now, we have arrived only at a relatively nebulous concept that we perceive things that give rise to states that may lead to action.

Much earlier than Dennett, Hume (1739,1740/1985) was suggesting that belief requires, in part, the *idea* of an object. Both *impressions* and *ideas* derive from 'all the perceptions of the human mind' (ibid., p. 49) and Hume distinguishes them thus:

> Those perceptions, which enter with most force and violence, we may name *impressions*; and under this name I comprehend all our sensations, passions and emotions, as they make their first appearance in the soul. By *ideas* I mean the faint images of these in thinking and reasoning.
>
> (ibid., p. 49)

Russell (1921/2005) highlights the difficulty in applying the criterion that Hume uses to distinguish impressions from ideas, especially in light of Hume's offering an exception immediately after suggesting the criterion. What is relevant here is the early suggestion that perceptions have been considered to be involved in providing the base information for belief formation, at least as far back as Hume. In a forecast of modern discussion of the role of feelings and emotions in affecting beliefs (Clore & Gasper 2000; Forgas 2000b; Frijda *et al.* 2000b; Frijda & Mesquita 2000), Hume goes on to suggest that:

> An idea assented to *feels* different from a fictitious idea, that the fancy alone presents to us: And this different feeling I endeavour to explain by calling it a superior *force*, or *vivacity*, or *solidity*, or *firmness*, or *steadiness*. This variety of terms, which may seem so unphilosophical, is intended only to express that act of the mind which renders realities more present to us than fictions, causes them to weigh more in the thought and gives them a superior influence on the passions and imagination … I confess, that 'tis impossible to explain perfectly this feeling or manner of conception. We may make use of words, that express something near it. But its true and proper name is *belief*, which is a term every one sufficiently understands in common life, And in philosophy we can go no farther, than assert, that it is something *felt* by the mind, which distinguishes the ideas of the judgment from the fictions of the imagination. It gives them more force and influence; makes them appear of greater importance; infixes them in the mind; and renders them the governing principles of all our actions.
>
> (Hume 1739,1740/1985, pp. 146–147)

So here is presented an earlier view that beliefs, however they may ultimately be defined, determine all our actions. Similarly, this view is taken up by Pylyshyn (1999, p. 3) in proposing that 'people do things because of what they know or believe and because of what they want …'. This, Pylyshyn claims, is a truism that cannot be denied without finding oneself unable to explain the simplest behaviour. Additionally, Hume preempts recent research that points to the most significant role of feelings determining the strength of a belief and the tenacity with which it is held.[6]

That beliefs direct actions is an essential aspect for Clifford as he declares that which 'has not some influence on the action of him who holds it' is not 'truly a belief at all' (1999,

6. This is a focus of Chapter 4.

73). Because he sees beliefs as being 'realised' either in the present or in the future through *actions,* his concern with beliefs is that there is none that is 'truly insignificant' (ibid., 73).

Russell (1921/2005), in hypothesising about the causal efficacy of belief, first examines whether we can say that belief causes voluntary physical movement (as distinct from reflex action). While that may be, it does not account for beliefs that simply involve thinking. What he concludes, then, is that 'a belief *may* influence action' (ibid., p. 148) but that belief cannot be defined by its effect on voluntary movements. So, again, there is agreement about beliefs instigating action or at least establishing the potential for action when circumstances are relevant.

Whether a belief is considered to necessarily influence, or even direct, actions is especially germane to this thesis. Here, the subject is beliefs that are either unsupported by evidence or are directly contradicted by evidence. If patently false beliefs are guiding actions then there is concern for the consequences at all levels from the individual through to nations, as suggested early in this chapter. Clifford takes what might be considered a high moral ground here:

> … if I let myself believe anything on insufficient evidence, there may be no great harm done by the mere belief; it may be true after all, or I may never have occasion to exhibit it in outward acts. But I cannot help doing this great wrong towards Man, that I make myself credulous. The danger to society is not merely that it should believe wrong things, though that is great enough; but that it should become credulous, and lose the habit of testing things and inquiring into them; for then it must sink back into savagery.
>
> (Clifford 1999, p. 76)

Russell appears not so strict in his censure regarding the requirement for evidence for believing. In presenting required characteristics of belief he first suggests truth or falsehood and considers 'fact' the thing which makes a belief true or not. He can suggest therefore:

> You may believe that such-and-such a horse will win the Derby. The time comes, and your horse wins or does not win; according to the outcome, your belief was true or false. You may believe that six times nine is fifty-six; in this case also there is a fact which makes your belief false.
>
> (Russell 1921/2005, p. 139)

Clifford's (1999) counter to this is most likely captured in his insistence that the sincerity of a person's conviction is no help because they have no right to believe on insufficient or no evidence and that they must be held responsible for holding a belief that is supported only by 'stifling' (ibid., p. 70) doubts and not by 'patient investigation'. His insistence on founding beliefs on evidence offers little scope for accepting the believing that a particular

horse would win a race, for example. In Russell's terms the belief may prove to be true after the event, but the holding of it is has only the slightest chance of being acceptable to Clifford. We could suppose that the racing fan knows an individual (e.g. trainer, jockey, vet) who claims to have certain knowledge that the race is fixed. Clifford does allow for belief based on authority, but there still has to be an evidential test of the authority.

This raises the issue, however, of whether adherence to truth is a necessary characteristic for accepting a person's stated or implied belief. Russell (1921/2005) is concerned here with the *intrinsic nature of belief* and not with its truth or falsehood which he treats separately. This line of inquiry is more promising for the type of belief being studied. Despite Clifford's protestations, people *do* hold and act on beliefs such as which particular horse will win a race.

As discussed earlier, James (1956) argues that it is fanciful to consider that pure reason can settle opinions once sentiment and wishful thinking are removed. He goes on to argue that even a belief that there is a truth is 'but a passionate affirmation of desire' (ibid., p. 9). In James' view, it is the 'passional nature' (ibid., p. 11) that decides on which proposition to assign one's support. Again, as mentioned earlier, with this view he heralds the views present and emerging in the current literature on the significant role of feelings in creating and maintaining beliefs (whether true or false).

But what of the view from epistemology? Before landing on a view of belief that is an acceptable one for this inquiry, it is appropriate to examine the line of enquiry that would emerge from following the logicians and epistemologists.

Blackburn (2005) draws on the dispute between James and Clifford in examining absolutism and relativism. James confuses the discussion according to Blackburn who claims that James sees belief as a 'choice between options' (ibid., p. 7). The choice depends on thinking that the belief has some chance of being true, that the choice is a forced choice between believing or not believing, and that there will be important consequences from assuming the belief. If all three conditions are met, the choice, according to James, is a 'genuine option' (ibid.; James 1956). James concludes that when a genuine option (in James' terms) cannot be solved on 'intellectual grounds', the 'passional nature' must be employed to choose.

> *Our passional nature not only lawfully may, but must, decide an option between propositions, whenever it is a genuine option that cannot by its nature be decided on intellectual grounds; for to say, under such circumstances, 'Do not decide but leave the question open', is itself a passional decision—just like deciding yes or no—and is attended with the same risk of losing the truth.*
>
> (James 1956, p. 11)

Blackburn (2005) objects on several counts. First, what may be a 'dead' issue for one person may be a 'live' one—that is, it has a chance of being true—for another. Second, we cannot be sure, in advance, that the matter cannot be solved intellectually: for example, the validity of astrology was once accepted widely but, now, for many, its lack of validity is clear both from evidence and from intellectual reasoning. Third, Blackburn attacks the notion of forced choice; and this seems to be on the interpretation that James' forced choice is predicated on there being equal probability of each choice. Blackburn says 'we should certainly not accept that whenever something is not decidable, its probability is evens' (ibid., p. 8) and gives the example of the probability of a roulette wheel showing a specific number in a spin. Finally, he is suspicious of the role of passions and is disturbed by the interaction between what people want to believe and what they finally believe. He suggests that James is 'objectifying' belief: viz. 'does this suit me; is it a good thing to wear to the social party?'.

For Blackburn, beliefs are based on truth and falsity. He sees James as 'privatising' (ibid., p. 9) belief in a way that leads to relativism and that is contrary to the idea that 'truth has rights and privileges of its own, and they are not just the same of those of utility' (ibid.). It is not clear that Blackburn has distinguished some of the nuance in James' suggestions. For example, James suggests that options may be 'forced or avoidable' (James 1956, p. 3). Avoidable options suggested by James include:

> 'Choose between going out with your umbrella or without it' …'Either love me or hate me' … 'Either call my theory true or call it false'
>
> (ibid., p. 3)

All these, James suggests, are not genuine options because one can remain indifferent to each. Similarly, one could remain indifferent to a suggestion that the roulette ball will land on number 34; there is no need to believe that it will or that it won't, one could simply wait and observe the result. What is a forced choice for James is 'Either accept this truth or go without it' (ibid., p. 3). Blackburn's example of the roulette wheel does not address this type of example from James or, at least, only partly addresses it. There may be some consideration of relative probability implied—James later alludes to weighing losses and gains in considering belief in God—but it seems Blackburn has inferred more than was intended. James is discussing solving the dilemma posed with his example: choosing to avoid a decision is effectively going without the truth offered. His point is aimed at uncovering how such a decision is made as, according to him, it must.

Ideally, for Blackburn and from a general philosophical view, believing a proposition results from establishing the truth of the proposition. Where historical or scientific fact can inform, that settles the matter. Then, as Blackburn (2005, p. 9) notes, it is not relevant whether the truth is a matter of satisfaction for people. However, he goes on to admit that

satisfactions are not the point 'unless the issue itself is one about those very satisfactions' (ibid.). This last point is not explicitly developed; and this seems a lost opportunity in understanding what James may be proposing. It may be that Blackburn is pointing, albeit unwittingly, to recent literature (discussed in Chapter 4) suggesting that ensuring that a belief has factual support in order to be viewed as 'true' will only get attention if a basis of truth for a belief is a domain of concern for a person. Nevertheless, while establishing the truth of a proposition to justify believing it is desirable from both a folk psychological perspective and a philosophical one, we cannot escape the endless set of examples of what people *express* as a belief with insufficient evidence or indeed even when confronted with counter evidence: for example, 'But I *know* he's innocent', despite the unchallengeable evidence in court; 'I believe it's my lucky night for the casino'.

In examining whether there is a difference between 'animation' (ibid., p. 19) and belief, Blackburn turns to Clifford's view and agrees with the connection between beliefs and action.

> Beliefs are functional states. They are preparations for action. This is biological function, or what they are for. So perhaps true believers should be distinguished from those who are animated by stories, poems, or metaphors precisely by the nature of this preparation.
>
> (ibid., p. 19)

So, from some philosophers' views (Clifford 1999; Hume 1739,1740/1985; James 1890/1950a, 1956; Russell 1921/2005), we have agreement with the suggestion that action flows from belief. Later, it will be demonstrated that there is general agreement from a psychology perspective of the connection between emotion and beliefs. The more rigorous philosophical approach does not seem to admit an emotional component in searching for the truth or falsity of a proposition (though, admittedly, the philosophical literature is not being covered in depth here).

Morton (1997) provides two comments that help clarify why a thorough epistemological enquiry into defining belief would be a diversion from the principal enquiry.

> Until recently, most philosophers working in the theory of knowledge have not paid much attention to the different ways in which beliefs and the ways we acquire them can be satisfactory or deficient. They have not asked : What qualities do we want our beliefs to have? And what qualities do we want them not to have? … focusing on the desire for truth, most philosophers until recently have described various ideals for beliefs: ways in which our beliefs and the ways we get them could be perfectly organised …
>
> One important question to ask about each of [the philosophers' ideals] is: Are human beings capable of satisfying this ideal? Can we have beliefs that are like this? But another equally important question is: What would be the price for satisfying this

> ideas? In order to have beliefs like this, would we have to lose something else of value?
>
> (ibid., pp. 2, 3)

> In discussing beliefs, philosophers use a number of central concepts. It is hard to define any of them in terms that all philosophers would accept …
>
> (ibid., p. 4)

The truth or falsity of a proposition that someone believes is at the core of this enquiry since the focus is on false beliefs that people hold despite contrary evidence or equally good, or better, counter propositions. But, it becomes apparent that understanding *how* it is that someone forms and maintains a false belief is unlikely to be served by extended debate over the ideals which we would like our beliefs to have. Morton (ibid.) takes note of how we report beliefs and why it is that we do so:

> When people disagree, they throw arguments, evidence and persuasion at another. Very often they apply abusive of flattering labels to the beliefs in question. 'That's false', 'That's irrational', 'You haven't got any evidence', or 'That's true', 'I have good reason to believe', 'I know it'.
>
> *We use these labels because there are properties we want our beliefs to have*; we want them to be true rather than false; we want to have good rather than bad reasons for believing them.
>
> (ibid., p. 1, emphasis added)

Regardless of what ideals we might want beliefs to adhere to, people continue to use such declarations as 'I believe that *p*', 'I know *p*', 'I'm certain of *p*', 'I'm absolutely sure that *p*' to announce that they hold a certain proposition to be true. Moreover, they do so despite counter evidence or superior counter propositions. They are not concerned with truth tables or various forms of logic; but this does not stop them using these beliefs as a guide for what action to take or as motivation for action when the circumstances are relevant. Clifford may well be horrified at this, but any such censure is of no interest to the person who, fired by some strong emotion, acts according to what they believe to be true. How they manage to do this is examined in Chapters 2 to 4.

As noted earlier, Russell (1921/2005) separates the discussion of the *nature* of belief from the discussion of truth and falsehood of a belief. He suggests:

> From the psychological standpoint … there cannot be any purely psychological means of distinguishing between true and false beliefs. A belief is rendered true or false by relation to a fact, which may lie outside the experience of the person entertaining the belief.
>
> (ibid., p. 153)

This clarifies Russell's earlier observation that 'the intrinsic nature of belief can be treated without reference to what makes it true or false' (ibid., p. 140). Russell, like Morton (ibid.), refers to our desires for how our beliefs should be:

> Nevertheless, we can hardly avoid the consideration of truth and falsehood. We wish to believe that our beliefs, sometimes at least, yield *knowledge*, and a belief does not yield knowledge unless it is true.
>
> (ibid., p. 140)

However, wanting a belief to be true provides no guarantee that a person will rigorously seek verifiable evidence and ensure that the argument for their belief is a sound one. One speculation here is that, for the non-philosophers, wanting a belief to be true could make the person even more likely to rely on processes of believing that promote the holding of a false belief: this speculation gets closer attention in later chapters. Clifford points to discussion in, or related to, current literature:

> We feel much happier and more secure when we think we know precisely what to do, no matter what happens, than when we have lost our way and do not know where to turn.
>
> (Clifford 1999, p. 75)

Clifford is developing the argument that there is great power and great satisfaction in the knowledge that comes from true (in Clifford's terms) beliefs. His use of 'feel' is perhaps significant in a way he has not considered. The role of emotions and feelings (discussed in Chapters 3 and 4) accounts for a great deal in establishing and maintaining beliefs that have no rigorous examination of their truth but which are firmly held to be true nevertheless.

If the notion that beliefs should meet certain ideals is set aside, the observed manifestation of beliefs is that they:

- involve a view that a proposition is true regardless of factual evidence
- are related to desires and wants
- are a preparation for action.

What is shown in the following chapters is there is much to conspire against a person's philosophical enquiry into the veracity of their beliefs.

1.1.1 The Cognitive Science View

Regardless of the ideals that might be proposed from a purely philosophical view, psychological research into beliefs has burgeoned in the last three decades. From offerings for definitions of 'belief' from the cognitive psychologists (predominantly) and neuroscientists, we find some considerable variation in the specifics. However, a general theme and, indeed, a general consensus can be found among these and many similar definitions.

Frijda et al. (2000b, p. 5) suggest that beliefs, along with emotions, are mental states, share certain qualities, can be distinguished by other qualities, and are 'intertwined'. Their definition of beliefs is:

> ... states that link a person or group or object or concept with one or more attributes, and this is held by the believer to be true.

A later definition (Frijda & Mesquita 2000, p. 45) is more helpful in holding a belief to be a proposition which is held to be true and noting that the affirming of beliefs usually goes beyond the occurrence of observable facts:

> We use the word 'belief' for a proposition that a person considers to be true. The truth of most propositions is not objectively fixed. Most propositions affirmed in beliefs go beyond the mere occurrence of observable facts. They extend to the causes or intentions behind events, their likely consequences, and the enduring properties of things or persons that make them behave in a particular manner.

Adding another dimension to the range of definitions, Forgas (2000b, p. 108) posits beliefs as 'stable, enduring cognitive representations that have a fundamental influence on the way people perceive, construct and interpret the social world'. In a move very distant from the spirit of Clifford's stance on beliefs, Schacter and Scarry (2000a) propose that memory is a form of belief when they cite a Webster Dictionary definition: 'confidence in the truth or existence of something not immediately susceptible to a rigorous proof' (ibid., p. 2) . More strongly, Eichenbaum and Bodkin (2000, p. 177) offer that belief is 'a disposition to behave in a manner that is resistant to correction by experience'.

In summarising the proceedings of a conference on memory and belief, Damasio (2000b, p. 326) suggests that a general consensus was reached and offers his understanding of belief:

> ... the attribution of truth value to a particular thought content, either perceived or recalled. It makes no difference whether the perception or recollection is that of an action, or an event, or an entity, or whether the perceived or recollected material is concrete or abstract ... to believe is to qualify a perception or recollection as true, false, or somewhere in between.

However, Damasio considers that the 'proper subject matter' for beliefs relates to matters that 'gravely affect our well-being'. Westbury and Dennett (2000) take some issue with this restriction and point out how a relatively 'simple' joke cannot be understood and appreciated without the existence of considerable propositions or beliefs: remove just one belief from an extensive list, and the joke *cannot* be understood. They initiate their discussion on belief by highlighting that 'belief' is difficult to define (ibid., p. 19) and that this is partly due to its use being contextual. Ultimately, they opt for a context-dependent use of the word:

> ... the definition of a belief in any particular circumstance is equivalent to the identification of the contingencies which allow that belief to be attributable.
>
> (ibid., p. 28)

This, perhaps, is related to what Blackburn claims when he says 'The issue determines its own epistemology' (2005, p. 9).

It is evident from even these few examples that psychological and neuroscientific research is already pursuing a greater understanding of beliefs. There is recognition that the holding of a proposition to be true is central to *believing* that proposition. However, these researchers are waiting on neither a definitive philosophical consensus on the ideals to which a belief should adhere nor a philosophical consensus on how it is that a person may come to hold a patently false proposition to be true.

As will be shown in in Chapter 2, it is very easy for a person to be misled by their perceptual processes into believing something that can be demonstrated to be false. A straightforward example of that is the phenomenon known as change blindness. If a person can start a conversation with one individual and complete the same conversation with a different person—notably of different appearance—without recognising it (Levin *et al.* 2000, 2002), we can suppose that they will believe that the entire conversation was with the same person. To omit considerations such as this from psychology studies in seeking to understand false beliefs would be, quite simply, wrong.

From the philosophers' camp, Goldman (1986, p. 9) argues strongly for an interdisciplinary approach to 'substantive evaluations of (primary) individual epistemology' especially requiring 'inputs from the psychology of cognition' (ibid., p. 9). Goldman's idea of *primary* individual epistemology is drawn from Francis Bacon:

> Neither the naked hand nor the understanding left to itself can effect much. It is by instruments and helps that the work is done, which are as much wanted for the understanding as for the hand. And as the instruments of the hand either give motion or guide, so the instruments of the mind supply either suggestions for the understanding or cautions.
>
> (Francis Bacon, *Novum Organum*, 1620, Bk 1, Section II, cited in Goldman, 1986, p. 4)

In accepting Hume's position, Goldman considers that an enquiry into cognitive processes to be 'fully appropriate for epistemology' (ibid., p. 5). With this challenge to philosophy's autonomy over epistemology, Goldman asks questions such as 'How can moral philsophy abstract from contingent facts of the human psyche; the sources of aggession and sympathy, the sense of fairness and reciprocity?' (ibid., p. 8). And, while granting that epistemology may be autonomous over 'foundational questions' he presents the view that 'substantive individual and social epistemology need help from other disciplines' (ibid., p. 9).

Goldman is not alone in looking beyond a singularly philosophical approach to understanding beliefs. Garfield (1988), at the time of writing, considered computational cognitive psychology as the 'best game in town' (ibid., p. 1) and sets out to use its paradigm to explore propositional attitudes. Forgas (2013, p. 3) suggests that, still, there is 'very little agreement among philosophers about the meaning of [belief]' and that, in contrast, psychology has 'done a great deal to define, operationalise and measure concepts such as belief in ways that are open to replication and verification'.

Schacter and Scarry (2000a) report on how interdiscplinary work that involved cognitive psychologists, neurobiologists, psychiatrists, and others on understanding memory distortion was confronted with the 'nature of belief and its role in memory' (ibid., p. 2). They note that the literature on false memories shows that:

> … they are often expressed as powerful, seemingly unshakable beliefs about the past.
>
> (ibid., p. 2)

This enquiry about belief and memory is an instance in which 'questions of science merge unavoidably with questions of epistemology' according to Westbury and Dennett (2000, p. 11). Damasio's (2000b) suggestion about the possible contribution from neuroscience and cognitive science to the understanding of belief and memory has already been noted (see page 9). These contributions from neuroscience and cognitive science, then, can be considered among the modern tools for the enquiry into the nature of belief and, in particular, false belief. There are ample examples from neuroscience and cognitive neuroscience suggesting how a person may come to hold a false belief (see for example: Bayne & Fernández 2009; Coltheart & Davies 2000; Hirstein 2000, 2005, 2009; Ramachandran & Blakeslee 1999).

Admittedly, while brain damage studies might not directly explain the false beliefs held by those who have no neuropathology, they are suggestive of possible mechanisms giving rise to false beliefs. For example, Gazzaniga (1998) and Ramachandran (Ramachandran & Blakeslee 1999) suggest that the left hemisphere of the brain has an interpreter role and

the way he hypothesises that it works accounts for how he can consider that 'autobiography is hopelessly inventive' (1998, p. 2).

Bortolotti (2010) also echoes the sentiments about the difficulty of defining beliefs and of providing 'an account of the necessary and sufficient conditions for believing that something is the case' (ibid., p. 11). She takes beliefs to be representations of 'how things are' (ibid.) and distinguishes between beliefs and other intentional states (e.g. desires, wishes) thus:

> 1 Beliefs have relations with the subject's other beliefs and other intentional states.
>
> 2 Beliefs are sensitive to the evidence available to the subject.
>
> 3 Beliefs are manifested in the subject's behaviour.
>
> (ibid., p. 12)

One aspect that is a limitation in Bortolotti's analysis, she acknowledges, is the 'affective dimension' (ibid., p. 12). In this, she draws on Stephens and Graham's (2004) analysis that includes an 'affect claim' (ibid., p. 237) along with a *content* claim, a *confidence* claim and a *reason-and-action* claim. Taking a broad view of intentional states, however, it is reasonable to propose that affective states supervene on manifested behaviour. (Chapter 3 examines this aspect more closely.) Bortolotti's view of Stephens and Graham's criteria is that they capture 'what *rational* beliefs should be like' (2010, p. 13); but she observes that not all beliefs behave in the way we would like. We find a form of agreement with others in her view that 'beliefs should not be idealised' (ibid., p. 13); especially, we should add here, for the purposes of cognitive psychological understanding.

To return to the focus of this thesis: that is to enquire into how it is that people experiencing no neuropathology or recognised psychopathology—that is, 'normal', everyday people—can develop and maintain demonstrably false or unsupportable beliefs. In this enterprise, it appears then that the approach of those including the cognitive psychologists, neuroscientists, and philosophers who engage in science, offers much more in discovering the underlying cognitive and neurological mechanisms than does a pure epistemological enquiry. So, despite the lack of consensus on the specifics on what is entailed in a belief, the messy collections of beliefs—tending more to the folk psychological type—are those being used henceforth.

At this stage, the literature suggests that for the purposes of this enterprise there is no need to go beyond the view of belief outlined on page 22 (repeated below for convenience).

A belief:

- involves a view that a proposition is true regardless of factual evidence
- is often related to desires and wants
- is a preparation for action.

1.2 Belief and Knowledge

It was earlier foreshadowed (see page 14) that the difference between *belief* and *knowledge* requires examination, principally because of the bearing it has on our understanding of *memory* and, in turn, on the relationship between belief and memory. Russell (1921/2005) highlights the human concern with knowledge:

> We wish to believe that our beliefs, sometimes at least, yield *knowledge*, and a belief does not yield knowledge unless it is true. The question whether our minds are instruments of knowledge, and, if so, in what sense, is so vital that any suggested analysis of mind must be examined in relation to this question. To ignore this question would be like describing a chronometer without regard to its accuracy as a timekeeper, or a thermometer without mentioning the fact that it measures temperature.
>
> (ibid., p. 153)

Quine & Ullian (1978) note, as do others (see Morton 1997, for example), that *knowing* is a 'special kind of believing' (ibid., p. 13).

> Believing something does not count as knowing it unless what is believed is in fact true. And even if what is believed is true, believing it does not count as knowing it unless the believer has firm grounds for belief.
>
> (Quine & Ullian 1978, p. 13)

However, Russell (1921/2005), again, points out that there is considerable difficulty not only in defining knowledge but also in knowing whether we have knowledge. Morton (1997), too, draws attention to the controversial nature of defining knowledge. He suggests that to 'take a belief as knowledge is to take it as something that you can trust when forming other beliefs' (ibid., p. 130), but he hastens to disqualify this suggestion from being a definition of knowledge. However, relevant to this thesis is that he suggests that this 'formula' is 'about which of a person's beliefs that person or another will *treat* as knowledge' (ibid.). Quine and Ullian (1978), as well as Morton (op. cit.), acknowledge the perversity of language that was indicated earlier (see pages 20, 21). They also note especially the 'hyperbolic use of "know" as an emphatic variant of "believe"' (ibid., p. 14). Of course, at the other end of the belief–knowledge spectrum, to suggest a possible construct, is the very clear distinction a person can make when choosing between saying 'I

believe it' or 'I know it'. For example, if the question is 'Do you consider public figure X is dishonest', the response could be a very considered 'I believe so', being an elliptical sentence where the 'but I have no evidence' remains unuttered; this is distinct from the 'I know so' where the 'because I have tangible, indisputable evidence' is not said but is known to the speaker.

Quine and Ullian's acknowledgement of the hyperbolic use of *know* as an emphatic form of belief is relevant to the discussion pursued in Chapters 3 and 4 examining the effect of emotions and body states—particularly the role of feelings—on believing. For now, it points to the challenge of distinguishing knowledge from belief and the additional challenges that arise when taking the psychological approach rather than a purely philosophical approach.

Damasio (2000b) has a somewhat awkward attempt at explaining his idea for separating belief from knowledge. Returning to his cited definition (see page 23), we recall he considers 'to believe is to qualify a perception or recollection as true, false, or somewhere in between' (ibid., p. 326). He adds:

> Whereas the term 'knowledge' refers to a datum in and of itself, the word 'belief' refers to a datum qualified. The qualification is based on the use of a scale on which truth and falsity are at opposing poles, and on which, in between, is a greater or smaller probability of being close to one or the other pole. The believer is confident of the qualification. Belief belongs in the same family of meanings as conviction and certainty.
>
> (ibid., p. 326)

Apart from acknowledging that data are involved in knowledge, this attempt appears to achieve little in clarifying the difference between belief and knowledge. It may be that by referring to 'conviction and certainty' Damasio is alluding to the belief being justified and true, conditions which, in one view, qualify the belief as knowledge. In more general terms, perhaps he is suggesting that in qualifying the belief to reach a state of conviction, the person has examined the available data and supporting evidence as valid. However, this attempt does highlight the challenges noted above.

Drawing on neuroscience studies, Eichenbaum and Bodkin (2000) diverge from the usual philosophical approach to relate belief and knowledge to *processes*:

> We ... propose that knowledge and belief may instead reflect distinct processing modes of the brain, that they utilize memory differently and reflect the activity of different brain systems. The distinction we make between knowledge and belief turns on the concept that knowledge is, as it were, *justified* through a particular type of representation and processing, whereas belief is not.
>
> (ibid., p. 177)

This anticipates the discussion in Chapters 3 and 4, in which the emotional response accompanying an utterance might correlate with the *strength* of a belief—how firmly that belief is held and defended—and also in separating belief and knowledge in the mind of the speaker. What remains problematic, at this stage, is when an utterance of *'I **know** it'* is said in defence of something which is clearly *believed* by the person. Is such an utterance an expression of strong belief, or is it a metacognitive assessment—which may also be a belief—of the state of the speaker's knowledge? While Damasio's explanation (above) alone is not sufficiently well clarified to be helpful with this problem, his views on the role of emotion and consciousness[7] might well align with Eichenbaum and Bodkin's intuitions regarding different brain systems.

Eichenbaum and Bodkin continue by suggesting that 'knowledge is a disposition to behave that is constantly subject to corrective modification and updating by experience, while belief is a disposition to behave in a manner that is resistant to correction by experience' (ibid., p. 177). Damasio's view that knowledge involves data gives an area of agreement between these two views, since data can be used to modify and update experience and thereby alter what one knows. These suggestions from Eichenbaum and Bodkin[8] and from Damasio add weight to the suggestion from Morton:

> A much more controversial conclusion … would be that the concept of knowledge is independent of that of justification: much of what we know is not justified and what is valuable about knowledge is not a matter of evidence and good reasoning.
>
> (Morton 2003, p. 106)

Since the focus of this thesis is the phenomenon of a 'natural' false belief, Morton's (ibid.) suggestion supports the avoiding of extended philosophical discussion of the nature of 'knowledge'. When variants of *'I know'* are used in lay speech, the person is usually not concerned with an epistemologically rigorous defence of their 'knowledge'. If they are concerned with 'evidence', then we see later (especially in Chapter 4) that what may be used as evidence is very much aligned with Morton's suggestion.

Examples can be: *I know I could never learn another language, I know there's no chance of my getting a promotion, I know politician X is definitely a crook.* Utterances such as these are ambiguous: is 'I know …' being used to express a belief that is unsupported by evidence or does the person have evidence (or, at least, the perception of evidence)? Where they consider they have evidence, it raises the question of what the speaker is choosing to use as data to support such 'knowledge'. There is the trivial possibility that the speaker is correct: for example, they might have made an earnest and concerted effort to learn three

7. To be discussed at some length in Chapters 3–5.

8. This work is discussed further in Section 1.2.1

languages, failed miserably, and be supported by a brain scan showing physical impairment in the areas of the brain which are requisite to developing a command of a foreign language. When supported by demonstrable evidence such beliefs are not false, and the holding of them might be a disappointing acknowledgement of the state of affairs, while at the same time being useful in that they prevent further wasted endeavour on a futile exercise. In the context of this thesis, such supportable claims are trivial. What is of interest are those utterances for which there is clear evidence which falsifies the utterance, or for which there is neither supporting nor falsifying evidence but for which we could propose ample reason why the utterance might be false. With these, it is suggested that such mechanisms as the following could be at play:

- observable falsifying evidence is ignored or is somehow 'filtered' from awareness
- observable falsifying evidence is noted but subsequently dismissed and the dismissal justified in some manner
- inappropriate data are chosen to support the belief: for example, emotional response and the strength of that response.

The first two points above will receive significant discussion in Chapter 2. The third point will be the subject of attention in various parts of Chapters 3 and 4.

Hirstein (2005) is another strong advocate of a multidisciplinary approach, drawing especially on neuroscientific findings. He follows Goldman and Quine in wanting to 'naturalise' epistemology and in the view that 'how we do obtain knowledge is highly relevant to the question of how we should obtain it' (ibid., p. 179). In an extensive discussion of knowledge in the context of confabulation, Hirstein points to the importance of realising that the falsity of a belief tells us nothing about how the falsity was arrived at. *Understanding how the falsity is arrived at and transformed into something held to be true is central to this thesis.*

At this point, it is useful to note that in Hirstein's discussion of criteria for confabulation, he argues that 'confidence is both an emotional and an epistemic concept' (ibid., p. 193). This points, first, to the acknowledged relation of emotions to beliefs and, second, to the mechanism used by the person to establish some idea or proposition as 'knowledge'. Both these aspects of confidence in the truth of a proposition are particularly salient in establishing a false belief. Where recognition of them helps here is in proposing an early view of 'knowledge' that can be developed later in the examination of how an individual might distinguish for themselves between what they consider to be a belief and what is knowledge.

Returning to Quine and Ullian's (op. cit.) example of the hyperbolic use of 'know' for 'believe'. For this thesis, it is appropriate to take up the view that when the person says 'I

know it', it can be an expression that the belief is strongly held. Leaving aside any rigorous epistemological distinctions between 'believe' and 'know', there are other possibilities that can occur at an individual level. The individual may have distinct criteria for how they distinguish a belief from knowledge. For example, something known is something they have witnessed (*I know it because I saw it happen.*): their own perceiving serves sufficiently, for them, as proof that the event or idea need not be taken on trust but can be accepted as a certainty, something known to be or to have occurred. An extension of this is that they may even be considered enough to limit the use of 'believe' to concepts or events of which they do *not* have direct evidence and 'know' to concepts or events of which they *do* have experience and/or verifiable evidence. In such a case, we might hear them say *'No, I don't believe it, I **know** it*': that is, it is not being taken on faith but rather on the basis of what they consider valid evidence.

A broad definition such as Bortolotti's (op. cit.) of beliefs as being representations of how things are would still lead to classifying something 'known' as a belief. In this case, it is possible to consider things known to be a subset of beliefs.

Rather than pursue rigid classifications of belief and knowledge, a more relevant question for this thesis is that raised by Hirstein (op. cit.). How does a person come to accept something as knowledge?

Where there is consensus in the literature examined is that what a person considers to be knowledge is considered true. A weaker consensus appears to exist with respect to how the truth of the knowledge is established. By weaker consensus, it is meant that there is agreement that knowledge is 'justified' but varying ways of arriving at the justifcation are offered. Eichenbaum and Bodkin (2000) propose 'that a cental feature of this required justificaton is the capacity to recover both direct and indirect associates of the known item, or have command of a body of facts and principles' (ibid., p. 204). Keeping in mind that they distinguish belief and knowledge as distinct forms of memory they suggest that:

> Knowledge-driven memory processing is 'bottom up', in that new experiences are paramount in forcing novel bits of information together to build or modify a memory scheme … By contrast, belief-driven memory processing is 'top down', in that the general schema is paramount in guiding the interpretation of new experience to confirm convictions and to specify actions consistent with those convictions.
>
> (ibid.. p. 204)

What this approach still does not address is that discussed above, in which the person indicates clearly that they are distinguishing belief from knowledge, but what they are claiming to be knowledge is far from justified in the eyes of the others. If, for example, a person claims that they *know* that one of their work colleagues is untrustworthy, it may well be that they have evidence, or what they believe to be evidence, for them to say

'know' rather than 'believe' or 'suspect'. In the case of evidence, however, there is a vast difference between their observing the colleague falsifying documents and their forming a judgment on the basis of 'intuition' which might be attributable to the accurate, albeit unconscious, interpretation of microexpressions. Even if the person has been trained by Paul Ekman (2001, 2003) we may not agree that the intuitive judgment is knowledge, while it is much easier to accept that the judgment is knowledge on the production of the falsified documents. A different example highlights the challenge: suppose a family member has been missing for some months. Despite the disappearance occurring in circumstances that suggest an unhappy ending, it is not uncommon to hear a family member say *'I just know they're still alive'.* In this instance, there appears to be no justification for considering the proposition 'they're still alive' to be knowledge; it is a strongly held belief. If challenged on this, they could reply *'This is not a belief, I know ... I just know they're still alive'*.

So far, the remaining issue is still the person who has made the distinction between believing and knowing and then claims knowledge in some matter for which they can produce no evidence or argument that (fully) satisfies others. They feel justified in treating the proposition as knowledge, and for this to have occurred we can hypothesise that there has been some conscious or unconscious reference to criteria for establishing truth. Eichenbaum and Bodkin's (op. cit.) hypothesis of different neural pathways may go some considerable way to explaining how either false knowledge or a false belief might be acquired. A further possibility comes from Damasio's (2000b) explanation of how a belief comes to be qualified.

> I suggest that either covert bioregulatory signals or overt emotions play the role of qualifiers in the ongoing mental formulation process. I also assume that certain kinds of internal biological states have a positive value that is aligned with consonance and truth, or a negative value that is aligned with dissonance and falsity (and every possible pairing in between). The sense of truth, the sense of falsity, and the confidence invested in that sense are related to bioregulatory operations including, of course, those we know as emotions and feelings.
>
> (ibid., p. 330)

It does not seem necessary for the purposes of this thesis to further pursue the epistemological clarity that may be desirable in distinguishing belief from knowledge. What is observed is that people both confuse the two (for example, in the hyperbolic use of 'know' to express a strongly held belief) and sometimes separate the two. In both cases, it seems that some mechanism is at play that creates the experience of knowing versus believing.

At this stage, there appear to be two criteria for a proposition to be accepted as knowledge:

1 the belief or the proposition is held to be true

2 available data are examined or interpreted to justify the claim as knowledge rather than belief. The data may be perceived external events or may be internal processes. The internal processes are discussed in the following chapters.

These two criteria fit well with Quine and Ullian's definition (see page 27) of knowledge. What it does leave open is what the person considers 'firm grounds for belief' (op. cit.), and that is an integral part of the discussion in Chapters 2 to 4. Also taken into consideration in settling on these two criteria is Eichenbaum and Bodkin's distinction between knowledge and belief (see page 29).

The view that knowledge is a disposition that is subject to experience and modification allows for a person to have the view that they have 'evidence' and, in that, being able to distinguish knowing from believing. Of course, it says nothing about either the quality of the evidence or of the deductive or inferential processes employed to conclude that a particular proposal is something known (that is, true and justified). Their recognition of belief as being resistant to correction fits with the type of belief that is the focus of this thesis, and goes some way to accounting for some of the vagaries of language noted by Quine and Ulian and by Morton. For example, the use of 'I know' as an expression of strong conviction of belief at least suggests that there is some qualification of what is thought to be true in order for there to be such emphasis in the expression. There still remains the investigation of how this qualification occurs.

Continuing the focus on what people do, rather than what they would ideally do, allows the two criteria above to serve as a working distinction for knowledge. While not a rigorous epistemological distinction, it follows the path of Quine and Ullian, Golman, Hirstein, and Eichenbaum and Bodkin in drawing on epistemology but recognising the findings from neuroscience and cognitive psychology should be allowed to shape the considerations of how a person distinguishes knowledge from belief in day-to-day thinking.

1.2.1 Knowledge and memory

There is a collection of terms which become problematic from both philosophical and semantic viewpoints. Terms such as *remember, know* (as in know that an event occurred without recalling all details), *autobiographical memory, autobiographical knowledge,* and *autobiographical belief*, are sometimes used with precision, and sometimes used with rather messy synonymy. Use can vary from one author to another and it takes some considered

reading to determine, with different authors, whether some of these terms are being used interchangeably or not.

In an early preemption of current debate, Russell (1921/2005) discusses 'memory-belief' and 'memory-knowledge'. When it comes to examining the degree to which memory gives us reliable knowledge of the past, Russell opts for the psychological approach (again preempting the approach being advocated by many current researchers):

> In regard to memory, as throughout the analysis of knowledge, there are two very distinct problems, namely: (1) as to the nature of the present occurrence in knowing; (2) as to the relation of this occurrence to what is known. When we remember, the knowing is now, while what is known is in the past. Our two questions are, in the case of memory:
>
> 1 What is the present occurrence when we remember?
>
> 2 What is the relation of this present occurrence to the past event which is remembered?
>
> Of these two questions, only the first concerns the psychologist; the second belongs to theory of knowledge. At the same time, if we accept the vague datum with which we began, to the effect that, in some sense, there is knowledge of the past, we shall have to find, if we can, such an account of the present occurrence in remembering as will make it not impossible for remembering to give us knowledge of the past. For the present, however, we shall do well to forget the problems concerning theory of knowledge, and concentrate upon the purely psychological problem of memory.
>
> (ibid., pp. 102–103)

Rolls (2000) observes that 'Neuroscience has reached the stage where it is possible to understand how parts of the brain actually work, by combining approaches from many disciplines' (ibid., p. 600). He goes on to examine neural systems involved in learning and in memory and concludes:

> It is now possible not only to delineate brain systems involved in different types of memory, but also to have some insight into the neuronal operations that take place in each of these brain systems, and thus to be able to produce quantitative biologically plausible neuronal network models of how each of these memory systems actually operates …
>
> (ibid., pp. 623–624)

A contemporary view of belief and knowledge as different forms of memory (Eichenbaum & Bodkin 2000; Eichenbaum 2002) may point to a way of resolving some of the problems posed by terms such as *remember* and *know* as used by Schacter (1996). Schacter (ibid., pp. 22–26) uses *know* in the sense of being able to be certain that an event occurred but being unable to record the details ('knowing' that you've met a person but being

unable to recall what they look like), or the 'tip-of-the-tongue' experience in which we know that we know someone's name without being able to recall it in the moment. These researchers seem to be reserving *remember* for being able to recall detail: for example, 'the subjective sense of remembering almost invariably involved some sort of visual reexperiencing of an event' (ibid., p. 23). Schacter does extend the use of *remember*, however, to include experiencing confidence in remembering because of the recollection of thoughts and reactions:

> … As I read the article, I made many mental notes about points where I agreed with the analysis and points where I disagreed. If you now ask me about the article, I can state with full confidence that I clearly remember having read it. But I do not have any specific visual images of exactly where I was when I read it, what the article looked like, and so forth. I remember having read the article because I recall my specific thoughts and reactions to it.
>
> (ibid., p. 24)

As suggested above, a potentially useful development in distinguishing knowledge from belief lies in the work of Eichenbaum (2002) and Eichenbaum and Bodkin (2000). Eichenbaum and Bodkin suggest that 'belief and knowledge are different forms of memory' (ibid., p. 176). In suggesting this, they do acknowledge that the position is 'admittedly simplistic' (ibid.); but, they do offer compelling evidence from cognitive and neuropsychological studies for this position. While many of their key points have been represented in the previous section, there is still more to be gained from their discussion that can help with understanding what *could* be occurring the in the following example.

When a person declares, *'I know that's what happened'*, the question arises as to whether this expression of a memory can be considered as a belief, as knowledge, or as both. Even though they employ 'know', we do not know whether it is a considered use of 'know' or simply expressing 'believe' with some greater degree of conviction than might be present for other recollections. They may be reporting accurate facts about a specific event, or they may simply strongly believe both that the event occurred as recounted and that they have evidence (even if the evidence gathering, the evidence itself, or the inference from the evidence is flawed). If the speaker consciously distinguishes 'know' from 'believe' and continues to provide evidence that can be verified by others, then to consider this as knowledge may be appropriate. On the other hand, if they say something like *'I can't recall the exact details, but I know that they argued'* it seems more appropriate to consider this a statement of belief.

Eichenbaum and Bodkin do acknowledge that 'the separation between belief and knowledge is not complete' (ibid., p. 179).

> [Belief] leads to a feeling, acceptance, or conviction … Knowledge involves continuous testing and updating of our organisation of information, and using these information structures to guide new decisions and insights. But beliefs can 'drive' knowledge processing as well, in that a belief will guide the recall of specific instances or facts that support the beliefs, and a belief can be used to manipulate and take command of a set of facts and rules. Thus, we think of belief-driven and knowledge-driven information processing as fundamentally interactive rather than operating in parallel.
>
> (ibid., p. 179)

If this view is right, then it is a key point in understanding the lack of precision in the choice of words used in recalling memories. Their view that knowledge and belief use different neural systems points to possible elucidation of mechanisms that can lead to (unconscious) confusion between belief and knowledge: for example, the use of *'I know it happened'* that is reporting a strongly held belief.

In an attitude similar to Rolls (2000), Markowitsch (2000) draws attention to the contribution of several disciplines in understanding memory:

> The combined application of modern neuroradiological, neuropsychological, and neurological methods, which allow the in vivo study of the healthy as well as the damaged brain, has resulting in a degree of functional localisation that until recently was unpredictable … Generally, the division into short-term and long-term memory has its continuing value. Short-term memory is seen as a mechanism of cortical association regions, particularly within prefrontal and parietal areas. The limbic system is viewed as engaged in the transfer of episodes and fact for long-term storage in cortical networks. Unimodal cortical regions process, and most likely store, primed information. And subcortical, in particular basal ganglia and cerebellar structures, process, and most likely store, procedural information. Retrieval is seen as engaging a combination of fronto-temporo-polar regions: the left hemisphere dominates retrieval of factual information and the right dominate retrieval of episodic information.
>
> (ibid., p. 478)

Drawing on the work of McDonald and White (1993), Eichenbaum and Bodkin (op. cit.) report three different types of memory resulting from activity in separate anatomical pathways: procedural memory from the 'convergence of sensory inputs and motor system connections in the neostriatum' (Eichenbaum & Bodkin 2000, p. 197), emotional memory from 'the convergence of sensory and affective inputs in the amygdala' (ibid., p. 197), and declarative memory from 'the convergence of higher-order sensory inputs in the hippocampus' (ibid.). They label procedural memory and emotional memory as beliefs, while declarative memory is considered to be knowledge. The distinction for Eichenbaum and Bodkin is that knowledge is representative in nature and, more importantly, allows for

inferential flexibility because of that representative nature. Normal memory, they suggest, results from a balance of the influences of the three parallel pathways and their associated forms of memory.

Eichenbaum and Bodkin (ibid.) hypothesise that defects in the cortical-hippocampal systems that produce the 'representational flexibility' (ibid., p. 201) they call knowledge allows the procedural and emotional memory systems—that is, beliefs—to dominate. This provides another way of viewing the problem of whether the statement *'I know that's what happened'* is knowledge or a belief. They link their findings to the philosophers' distinction that knowledge must be justified and propose that:

> … a central feature of this required justification is the capacity to recover both direct and indirect associates of the known item, or have command of a body of facts and principles. Conversely … memory without this feature of knowledge is 'pure' belief.
>
> (ibid., p. 204)

To reiterate an earlier point, if the person is able to fulfil Eichenbaum and Bodkin's criteria—that is, recover associates of the known item and have reference to facts and principles—the listener should have no problem in accepting this memory as knowledge. But, this has still not addressed the issue of the person who declares it as knowledge but who is unable to satisfy the philosophical listener by providing any verifiable facts or evidence. In their own mind, the speaker may well consider that they have satisfied the criteria for knowledge, while the listener considers the statement to be a belief.

For example, it is possible that the speaker has made an error of correlation in the past and established a faulty principle on which to assess whether they 'know' something: they attempt to remember a specific event and because of particular feelings (that is, body sensations) they declare that they are 'certain' of their recollection. Later, they check a reliable record (for example, diary, photos) and find that their recollection was correct. If they now decide that the presence of those particular body sensations (that is, the feelings of 'certainty') is a sufficient criterion to declare knowledge, they have established a guiding principle based on an heuristic that is most likely faulty. But, that does not stop them using the principle and declaring to possess knowledge while the patient philosophical listener, from a perspective external to the neural processes of the speaker, declares what is said to be belief only.

A key problem here is the point of reference. The definition of knowledge that the philosopher or philosophically-inclined neuroscientist may wish to invoke, can also be invoked by someone whose processes for justifying a belief are faulty. A key interest of this thesis is how people can establish and maintain the use of faulty criteria not only for

believing something but also for deciding that they 'know' something which for anyone else is merely a belief (and possibly a false one).

To avoid protracted epistemological consideration that will not further the thesis, it is proposed that the criteria for knowledge suggested earlier (see page 33) be used. However, the consideration required is that criterion two (that is, available data are examined or interpreted to justify the claim as knowledge) is carried out *from the perspective of the speaker*. The examination of the data may be faulty, the data may be misperceived initially, and the inferences drawn may be faulty; but, for many people, none of those errors will drive them from their certainty, principally because they are simply not aware of the errors. In discussing the degree to which rationality constraints should be applied to beliefs, Davies and Coltheart (2000) adopt a similar approach:

> If we cannot make any sense at all of how a certain person could reasonably have arrived at a particular belief on the basis of experience and inference then this counts, provisionally even if not decisively, against the attribution of that belief to that person. One way that we can apply this test of intelligibility is to imagine ourselves in the other person's situation, or even imagine being the other person, and then to consider what, from that person's point of view would be a reasonable thing to believe. We *simulate* the other person in imagination and seek to understand him or her 'from the inside'.
>
> (ibid., p. 2)

What this approach does admit is acceptance of Eichenbaum and Bodkin's view of memories as being belief or knowledge. A counter view that could be offered is that since one definition of knowledge is that it is justified belief, all memories that are knowledge are, by definition, beliefs. But, as noted earlier, despite any desire for epistemological purity, that is not how most people behave. In many situations, many people will declare to possess knowledge when, in fact, they have only a belief. Consequently, they can manifest different behaviour depending on whether they are relying on their 'knowledge' and what they consider a belief.

What is proposed here is that the use of *'know'* meaning *'believe'* must be determined contextually. Where someone says *I know it* with the meaning *I believe it (strongly)*, this will be treated as a belief. Contextual awareness should also include awareness of the perspective: that is, *from the speaker's perspective*, is the claim considered belief or knowledge regardless of whether it satisfies epistemological ideals. The chief advantage of assuming this approach lies in its being an opening to addressing factors such as the speaker's understanding of particular language elements (for example, 'know', 'believe'), the effects of emotions and accompanying feelings, the effects of other beliefs, and autobiographical memories.

This last factor, autobiographical memories (relevant again in Chapter 4), also leads other researchers to distinguish between memory, belief and knowledge. Conway and Pleydell-Pearce (2000), for example, distinguish between autobiographical memory and autobiographical knowledge; their basic premise is that 'autobiographical memories are transitory dynamic mental constructions generated from an underlying knowledge base' (ibid., p. 261). Acknowledging Conway and Pleydell-Pearce's work, Mazzoni and Kirsch (2002), also distinguish between autobiographical knowledge and autobiographical belief. These two approaches appear to differ from Eichenbaum (2002) and Eichenbaum and Bodkin (2000) in distinguishing belief and knowledge *from* memory whereas Eichenbaum and Bodkin categorise memories *as being* either belief or knowledge. Considering memories as being belief or knowledge arises from neuroscience findings, while the distinctions of memory, belief and knowledge arise from a different focus in the study and from a psychology platform. The memory-as-belief-or-knowledge approach appears more immediately useful in pursuing an understanding of how a person distinguishes—or confuses—knowledge and belief (autobiographical or otherwise). However, it is relevant to acknowledge the other views to emphasise how the framing of the enquiry and the discipline in which the enquiry occurs can legitimately employ different views of the same concepts. While this does not necessarily assist in clarifying terminology on a general scale across disciplines, it does draw attention to the need for clarity within the context of the investigation.

1.2.2 Knowledge, Belief and Memory

As discussed above, Eichenbaum and Bodkin (ibid.) view *both* knowledge and belief as forms of memory. In distinguishing between these two forms of memory they use the distinctions of *memory-with-knowledge* and *memory-without-knowledge*. They hypothesise that the knowledge system involves the cortical-hippocampal system and that the belief system involves the cortical-neostriatal and/or cortical-amygdalar systems. This latter system they consider rigid, habitual and emotion-driven and it is that which makes it a *belief* system (ibid., p. 201). They seek to align their view with a philosophical view that knowledge is 'justified'; so a memory which is knowledge derives from the capacity to recover both direct and indirect associates of the known item, or have command of a body of facts and principles' (ibid., p. 204).

Eichenbaum and Bodkin are not alone in viewing at least some types of memories as beliefs (Conway & Pleydell-Pearce 2000; Mazzoni & Kirsch 2002). So, too, do Schacter and Scarry:

> To the extent that beliefs are defined by a subjective conviction about the truth of an assertion that cannot be proven, then memory could be viewed as a type of belief: it

> is often difficult to offer any 'rigorous proof' of what did or did not happen in the past. And because memory is a fundamentally constructive process that is sometimes prone to error and distortion, it makes sense that such beliefs are occasionally misguided.
>
> (Schacter & Scarry 2000b, pp. 2-3)

Johnson and Raye (2000) take the view that memories are beliefs, and that beliefs are constructed from memories as well as being reinforced by memories. Their note on the lay use of *memory* and *belief* somewhat supports Eichenbaum and Bodkin's view of memories:

> People tend to use the word 'memory' when a mental experience or report of a mental experience is detailed, including information indicating that one experienced the event oneself, and they tend to use the word 'belief' when it does not have contextual details and for a broad range of mental experiences or reports that seem to assert present or past general events …
>
> (ibid., p. 36)

Here, the use of 'memory' is in accord with Eichenbaum and Bodkin's view of memory as knowledge while the use of 'belief' is somewhat in accord with their view of memory as belief (as discussed above). This might add weight to the distinctions that Eichenbaum and Bodkin draw, but final acceptance of terms used this way awaits verification of human correlates of their hypothesised knowledge and belief systems.

However, in reading of Ramachandran (2000) it seems possible to infer a more complex relationship between memory and belief than that suggested by Johnson and Raye. Findings such as those he reports in a case of anosognosia[9] and of a patient with Capgras Syndrome[10] raise the following points:

- In the case of anosognosia, the belief that a paralysed arm is still functional appears to stem from physical damage to the brain. Indeed, the awareness of the paralysis during cold water caloric irrigation to the left ear (but not the right), further strengthens the case for the activation of specific regions and/or neural circuits being instrumental in forming beliefs (false beliefs in this case). While the neural regions involved in this case are not exactly the same as those referred to by Eichenbaum and Bodkin (2000), that these false beliefs are region-specific does give some support to their view that knowledge systems and belief systems use different regions/circuits of the brain.

9. Lack of awareness or denial of some handicap (e.g. paralysis) following some brain injury.
10. The regarding of close acquaintances (e.g. parents, spouse, children) as impostors.

- Ramachandran notes that while Capgras Syndrome is seen in psychotic states, 'more than a third of the documented cases' have been in conjunction with traumatic brain lesions. He suggests that the syndrome has an organic basis.

  Some interesting points come out of the case he discusses (2000):

  - In one case at least, the syndrome was modality-specific: the patient could recognise his parents auditorily (on the phone) but not visually.
  - Despite affirming that pictures of his parents *looked* like his parents (although still perceived as other people), the patient showed no galvanic skin response (GSR) when viewing the photos. Ramachandran posits the possibility that there is a disconnect between the face area of the temporal lobes and the amygdala—even though both regions may be intact—resulting in a lack of emotional response when seeing his mother, for example. The lack of emotion does not prevent the patient from recognising that the face resembles his mother, but she is not experienced as being his mother. Hence the belief that the woman in front of him is not his mother.

    This work points to an argument that will be developed later (Chapter 3) in which the role emotion in forming and maintaining beliefs is examined.
  - Ramachandran also suggests (ibid., p. 109) that the brain might rely on specific signals from the amygdala to link successive episodes. So the Capgras Syndrome patient examined could meet the same person in two episodes separated by half an hour and, although acknowledging that the 'two people' look alike, believe that they are different people. If we refer back to Eichenbaum and Bodkin's suggestion, we have a possible example here of a case which demonstrates the inability to create associations and links which can be classed as 'knowledge'. Again, while not being conclusive it does add weight to the view that a memory which can be classed as knowledge derives from different neural activity to a memory which is classed as a belief—or, in this case, a belief resulting from a lack of memory.

Ramachandran briefly mentions (ibid., p. 108) antegrade amnesia[11] caused by damage to the hippocampus. He concludes that the hippocampus is required for acquiring new memory traces. Eichenbaum and Bodkin (op. cit.) draw on studies of patients with damaged or removed hippocampuses in establishing their view that there are 'knowledge-processing pathways' in the brain.

11. The inability to recall new information.

By now, the case is strengthened for the view that different brain regions or networks are responsible for phenomena that can be labelled as 'belief', 'knowledge' or 'memory'. It is not yet clear that all memories can be divided into either beliefs or knowledge (Eichenbaum & Bodkin 2000; Eichenbaum 2002), or whether, for example, a memory can be considered distinct from knowledge in some circumstances (Conway & Pleydell-Pearce 2000; Mazzoni & Kirsch 2002). However, the usefulness of making appropriate distinctions is apparent on recalling the focus of this thesis: how does a person establish and maintain a false, or easily falsifiable, belief? The vagaries of an individual's language have been noted on several occasions, but its relevance grows when we shift perspective from researcher/observer to being the person. When the person says *'I know that's how it happened'*, it is relevant to know whether they are making a ***personal*** distinction between *know* and *believe* or whether they are using *know* to be a hyperbolic form of *believe*. In either case, if the underlying neural mechanisms make a difference, then it can be anticipated, quite reasonably, that a greater understanding of these mechanisms will help researchers reach a collective understanding of the relationship between memory, belief and knowledge. But, the challenges go beyond just establishing the role of particular brain regions or neural circuits.

Frith and Dolan (2000) also suggest additional complexity in the relationship when they refer to the perceptual distortions experienced by some schizophrenia sufferers:

> The symptoms seem to lie in the middle of a continuum that runs from false perceptions to false beliefs. This middle ground of abnormal psychological experience suggests that perceptions and beliefs are not necessarily distinct categories.
>
> (ibid., p. 116)

A full discussion of the role of perception belongs in Chapter 2, but it is worth noting the importance of investigating the interplay between perception and beliefs. Frith and Dolan (ibid.) echo the earlier observations of James:

> … *whilst part of what we perceive comes through our senses from the object before us, another part* (and it may be the larger part) *always comes … out of our own head.*
>
> (James 1890/1950a, p. 103)

Frith and Dolan continue:

> … perceptions are partly a function of prior knowledge. This prior knowledge, which must be a form of memory, is often expressed as beliefs. It follows that both perceptions and beliefs must be intimately related to the operations of memory.
>
> (Frith & Dolan 2000, p. 117)

The interaction between beliefs, knowledge, and perception are further complicated by the role of emotions (Frijda *et al.* 2000a). Awareness of this goes back at least as far as Hume (1739,1740/1985), as noted earlier (see page 16). Similarly, Russell (1921/2005) refers frequently to the role of 'feeling' and James' (1956) view on the role of emotions has also been noted earlier (see page 18). Investigation of the role of emotions and feelings is reserved for Chapters 3 and 4. Here, it is worth noting that the extent of research into the effects of emotions on beliefs (and vice versa) indicates, first, the growing contributions from psychology and cognitive neuroscience and, second, the broad acceptance of the view that while such research benefits from philosophical guidance, philosophy alone is unlikely to help us understand how people act. Frijda et al. (2000b) put their position thus:

> In the philosophical tradition *belief* is distinguished from *knowledge* by reference to the truth value and claim to objectivity of knowledge: 'True' knowledge is distinguished from 'mere' belief. Psychology is less interested in this question of *de iure*, the question of justification of the proposition; it is more concerned about the question of *de facto*, the psychological reality. Thus whether Dracula exists or not is less important for psychology than the fact that Rachel believes and hopes that he will pay her a visit tonight. If there is a difference between knowledge and belief that is of psychological significance, it is the way in which they vary with respect to preparing the individual to act.
>
> (ibid., p. 4)

With respect to the complexity of the interplay between knowledge, belief and memory, Westbury's and Dennett's (2000, p. 29) comment is highly suggestive of the scope for further research:

> The complex network of implicit and explicit knowledge that underlies the categories of both 'belief' and 'memory' rests on the ability of that network to both define and recognise its own coherence.

And they further suggest the challenge involved by cautioning against 'any definition that treats memory and belief as if they could be simple primitives existing only inside the brain' (ibid., p. 29). From the same conference as Westbury and Dennett, Damasio (2000b) hypothesises processes which affect belief and knowledge (or perception of knowledge).[12] If Damasio is right, the effects of these 'qualifiers' in their many forms and varying intensities may contribute greatly to such phenomena as the degree of belief or conviction. Combine Damasio's qualifiers with the possibility of different neural regions or circuits acting to produce either beliefs or knowledge (and possibly memory as a

12. See quote on page 32.

separate phenomenon) and it becomes significantly easier to imagine how such complexity can lead people to easily create and maintain false, or falsifiable, beliefs.

1.3 Summary

It is very easy to find instances of false belief and irrationality; though it can be much harder to gain agreement about what is false or irrational! We noted at the beginning of the chapter that there is a rise in educated popular literature drawing attention to the ubiquity and potential dangers of irrational beliefs.

While the nature of belief has long been a concern of the philosophers, one of the most strident and stringent philosophers reviewed with respect to placing constraints on belief is Clifford. A more tempered approach from James has been examined, and his work is one of those (along with Russell and many contemporary researchers) who recognise the role of feelings (that is, emotions) in establishing and shaping beliefs. Damasio's suggests that what we recognise as beliefs are in the domain of those things that are particularly important to us. This, along with the proposed role of emotions, indicates a path of inquiry that may help explain how it is that normal[13] people can cling to beliefs that are either demonstrably false or for which there is no available evidence and which are falsifiable through logically sound argument.

The challenge in defining beliefs has been evident in the literature reviewed. More particularly, in understanding what people actually believe and how that affects their decisions and actions, an approach that informs philosophy from the findings of psychology and neuroscience is becoming the established practice with many researchers. Garfield (1988) puts this approach succinctly when saying that 'the philosophy of cognitive science must pay closer attention to the practice of cognitive science' (ibid., p. 156). Whether stated explicitly or not, that view appears to be guiding much contemporary research into beliefs; however, such an approach is even evident in the work of philosophers such as Hume and Russell.

In this vein, the focus of the thesis is served by leaving aside the pursuit of a rigorous epistemological view of 'belief' and working with a view that relates to how people speak and act. For this purpose, the key common elements in the reviewed literature are that a belief:

1 involves a view that a proposition is true regardless of factual evidence

2 is often related to desires and wants

13. That is, people with no identifiable psychopathy, neuropharmacological abnormalities or neurophysiological damage.

3 is a preparation for action.

The first characteristic encapsulates the disregard for factual evidence that is the prime interest of this thesis. Points two and three acknowledge the complex relationship between belief, other propositional attitudes (and, by implication, emotions), and decision making and actions.

With respect to distinguishing belief and knowledge, the philosophical pursuit is complex and difficult. Morton (1997), for example, when discussing Lehrer's principle, notes that 'most beliefs result in part from processes that are not explicit reasoning' (ibid., p. 116) and thus fail to meet standards required by some philosophers. While the pursuit of philosophical precision is beyond the scope of this thesis, the aspect which is particularly relevant is the broadly agreed condition that knowledge is a 'justified' belief. There can be a significant difference for an individual between saying *'I know that p'* and *'I believe that p'*. If the issue of semantic laxity is put aside and we grant that *'I know that p'*, from the speaker's perspective, represents a qualification that imparts a greater degree of certainty on the proposition, we face the enquiry of determining what processes and criteria were employed to decide that they possess knowledge rather than mere belief.

It is suggested that an important approach for this enquiry is considering the matter from the speaker's perspective. We can then settle on knowledge being a belief that is qualified by the believer, but not necessarily qualified in an objective, rational way. For the observer, the professed knowledge might be quite demonstrably wrong and therefore be classed as an irrational belief, but that does not prevent the speaker from maintaining the view that they possess knowledge.

Of course, there remains the possibility that some will distinguish, consciously, between *believe* and *know*: for example, *believe* being to make decisions and determine actions on the basis of something being considered true but with an acknowledged lack of proof (e.g. the sun will rise tomorrow); *know* being something confirmed as fact (e.g. the temperature outside is –40°C because that's what the thermometer is reading; she arrived at exactly 3.00 pm because the time signal was playing on the radio as she walked through the door).

Returning to the point of semantic laxity, there remains the acknowledged possibility that when a person says 'I know it', it can be an expression that the belief is strongly held. While this adds a degree of complexity to understanding the mind of the speaker, it points to aspects of the enquiry under focus, particularly the role of feelings and emotions.

For the purpose of this thesis, there appear to be two principal criteria for a proposition to be accepted as knowledge:

1 The belief or the proposition is held to be true.

2 Available data are examined or interpreted to justify the claim as knowledge rather than belief. The data may be perceived external events or may be internal processes such as emotional responses. What is acknowledged here is that there is not necessarily an inherently rational process used to justify the claim of knowledge.

Developing out of the challenge of distinguishing belief from knowledge is the enquiry into the relationship of both these with memory. Suggestions from both philosophers such as Russell (1921/2005) and contemporary researchers (for example, Eichenbaum & Bodkin 2000) that it may be possible to separate memories into being either beliefs or knowledge open a valuable line of enquiry into both psychological and neurological mechanisms that give rise to such distinctions.

Beliefs are formed in a complex system involving emotions, perceptual processes, interaction with other beliefs and propositional attitudes, and metacognitive beliefs. Given this complexity, Westbury and Dennett's (2000) exhortation to account for **context** in defining and understanding belief will be shown to be particularly relevant. The following chapters examine how these respective elements of this system help form and maintain beliefs and, in particular, how people can hold false beliefs.

Chapter 2: Perception, Heuristics and Biases, and Emotion

At the time of writing, the recognition of the scale of the problem of irrationality is increasingly evident in the leakage of writing from research domains into the popular literature (Gigerenzer *et al.* 1999; Gigerenzer 2007; Gilovich 1991; Klein 2005; Macknik & Martinez-Conde 2010; Sutherland 2007; Taleb 2004, 2007). Topics range from how our perceptual processes mislead us through to biases that affect judgments and, at the other end of the continuum, how unconscious processes can contribute to effective 'intuitive' judgment. Titles such as *Simple Heuristics That Make Us Smart* (Gigerenzer *et al.* 1999) examine how 'fast and frugal heuristics' (ibid., p. 5) can work well for making decisions. Prior to work such as this there has been an extensive literature on how decision making and judgments can go wrong. It is this literature that is of more interest here. As will be shown there is a strong connection between irrational beliefs and the problems of perception and our biases in our judgments that appear to be an evolutionary legacy.

2.1 Perception and Perceptual Mechanisms

What people presume to have perceived is frequently used as grounds for belief or for claiming knowledge. The question that is posed here, but addressed more fully in Chapter 4, is: what is the source of the certainty, the apparent strength of conviction, that allows people to hold fast to expressions of belief such as the following, even in the total absence of any objective evidence?

- *I know he's the one who did it, I **saw** him leaving the store.*
- *Just **listen** to the way they talk about my presentation; I can't help believing they really hated/liked it.*
- *You only have to **look** at the Balance Sheet to tell that this company is worth investing in.*
- *This suit is really top quality, just **feel** the quality of the fabric.*
- *I know I was there, I can **see** it clearly (in my mind). I remember you **saying** what a relaxing place it was.*
- *Just **look** at her eyes, you can tell she's as dishonest as they come.*

Expressions such as the above abound in everyday language, as casual observation attests. The unquestioned acceptance of such claims is most likely what Berkeley refers to when he says:

> … we see the illiterate bulk of mankind that walk the high-road of Nature, for the most part easy and undisturbed. To them nothing that's familiar appears

> unaccountable or difficult to comprehend. They complain not of any want of evidence in their senses, and are out of all danger of becoming *sceptics*.
>
> (Berkeley 2009, p. 7)

Berkeley goes on to point out that 'uncouth paradoxes, difficulties, and inconsistencies' (ibid., p. 7) appear as we commence a deeper, philosophical, examination of such apparently straightforwardly supportable claims. Similarly, from a contemporary viewpoint, Wheatley (2009) suggests:

> We assume that what we see and feel is an objective read-out of the world, unvarnished by personal biases, contexts, and assumptions. Our perceptual experience is taken as read and our explanations tailored to fit. As such our explanations are only as good as our perceptions.
>
> (ibid., p. 208)

When considered more rigorously, sense-based claims point to key considerations in how a person can form and maintain a belief or claim knowledge, namely:

- the validity of perceptions
- the cognitive processes for making judgments from perceptions
- the role of existing beliefs in (1) affecting how events are perceived, and (2) interpreting those perceptions
- the effect of emotions and feelings.

The last consideration relating to the effect of emotions and feelings is predominantly discussed in Chapters 3 and 4; however, some treatment is given later in this chapter in relation to their effect on perception. The other points comprise the principal content of this chapter.

To the layperson, it may seem common sense to start with perception as the source for forming beliefs and proclaiming knowledge. However, although what is claimed to have been perceived is often the ultimate evidence provided for a judgment or belief, perception is problematic from the perspectives of several disciplines.

While Goldman (1986) suggests that perception is 'a suitable starting point in exploring the interface between psychology and epistemology' (ibid., p. 184), he adds the qualification that the psychology of perception is both extensive and complex. Here, then, the focus remains on how perception processes by themselves can contribute to false belief.

Audi (2011) introduces the discussion of perception thus:

> Before me is a grassy green field. A line of trees marks its far edge, which is punctuated by a spruce on its left side and a maple on its right. Birds are singing. A

> warm breeze brings the smell of roses from a nearby trellis. I reach for a glass of iced tea, still cold to the touch and flavoured by fresh mint. I am alert, the air is clear, the scene is quiet. My perceptions are quite distinct.
>
> It is altogether natural to think that from perceptions like these, we come to know a great deal—enough to guide us through much of daily life. But we sometimes make mistakes about what we perceive, just as we sometimes misremember what we have done, or infer false conclusions from what we believe. We may then think we know something when in fact we do not, as when we make errors through inattention or are deceived by vivid dreams. And is it not possible that we are mistaken more often than we think?
>
> ....
>
> In approaching these topics in epistemology … it is appropriate to begin with perception.
>
> (ibid., pp. 1–2)

Morton (1997), too, says that 'any account of the ways we can and could acquire our beliefs must discuss perception, the use of our senses to learn things about the world around us' (ibid., p. 23). He goes on to note that philosophers have run into 'a very basic problem' (ibid.) when trying to answer questions about how we gain information about our environment, when we can know that information is accurate, and how the information from our senses provides us with evidence for beliefs. According to Morton (ibid.), Locke, Berkeley and Hume are among the philosophers who developed empiricism. Locke (1690) indicates his view of the key role of perceptions thus:

> First, our Senses, conversant about particular sensible objects, do convey into the mind several distinct perceptions of things, according to those various ways wherein those objects do affect them. And thus we come by those *ideas* we have of yellow, white, heat, cold, soft, hard, bitter, sweet, and all those which we call sensible qualities …
>
> (ibid., Book II, Chapter I)

The 'ideas' of Locke and those that followed are referred to as 'sense-data' by later empiricists (Morton 1997). Morton argues that the early empiricists had good reason to rely on sense-data for two reasons: first, they are 'at the end of the line in terms of evidence' (ibid., p.28); second, the evidence 'cannot be doubted by the person who has it' (ibid., p.28). Hence, resort to sense-data is a natural course when asked to justify a claim or belief. In fact, Morton argues that, under challenge, people tend to defend their views using even more perceptual beliefs based on sense-data. As will be shown in Chapter 4, despite what a person might *see, hear, taste* or *smell*, the sense-data of *feelings* are particularly significant. Morton also suggests that perceptual beliefs are 'relatively

independent of other beliefs' (ibid., p. 29), thereby reducing the potential for circular reasoning.

Morton (ibid., 2003) raises important philosophical issues with the empiricist position: for example, the pursuit of certainty leading to weaker evidence—'more certainty means weaker evidence' (1997, p. 30)—with an apt example being the limits to scientific measurement. More broadly, taking the empiricist position to its extreme means justifying 'such general and fundamental beliefs as the belief that there is a physical world which is the cause of one's perceptions' (2003, p. 24). There is also the 'other minds problem' (ibid., p. 24) in which others might lie, be mistaken, or simply be providing meaningless utterances.

Although Audi (2011) offers extended discussion of perceptual beliefs and issues relating to such areas as sense-data evidence, an extended pursuit of these deeper philosophical issues is not relevant to this thesis. Where Audi and Morton agree, and suggest that there is general agreement, is in the fallibility of our senses and associated perception:

> Even well-grounded beliefs can be mistaken. We can be deceived by our senses.
>
> (ibid., p. 8)

Styles (2005) notes that Locke drew attention to the complexities of sensation. She paraphrases one of Locke's examples as follows:

> Locke gave the example of three bowls of water—one hot, one cold and one tepid. Put one hand in the hot water and the other in the cold water, and wait for a few minutes. Then put both hands into the tepid water. To the hand that was in the cold water, the tepid water feels hot; to the hand that was in the hot water, the tepid water feels cold. How can this be? You have two different sensations arising from the same stimulus. Is the tepid water cold or hot? How can we rely on our senses to tell us about the world 'out there' if, as in this case, they are giving us conflicting information. Lock used this example to make clear the difference between the real world and the subjective, experienced world—the difference between sensation and perception.
>
> (ibid., pp. 43–44)

On the basis that the non-philosopher appears particularly inclined to rely on their senses to justify beliefs or to claim knowledge—without concern for such distinctions as doxastic justification or situational justification, perceptual beliefs, objectual beliefs or propositional beliefs—an examination of how sensation and perception can lead to beliefs that seem grounded in objective data and yet are false is in order. However, the background question that remains is that concerning the ability to remain fast to a false belief based on perceptions. We might hope that even a simple example such as Locke's, as presented by Styles (op. cit.), should help educate a person who is exposed to the example

regarding the fallibility of perceptual processes. From that, we might further hope that the person would be more willing to question their judgments based on perceptions and seek quality objective evidence. So, although it is shown in this chapter that a key source of false beliefs is what is deemed to have been perceived, the thesis ultimately moves to examining the influences driving the resistance to examination of and change of belief.

2.1.1 Sensory Dysfunction

In the literature reviewed there is little or no mention of the effect of damage or pathology or functional decline in the sensory apparatuses. There is ample discussion of the effects of brain damage (via trauma or pathology), but not of damage to the 'front line' of sensation (the eyes, ear, skin, nose and mouth, and the associated sensory neurones). However, damage to the 'front line' seems an appropriate place to start enquiry.

It is easy to imagine situations in which two or more people engage in a dispute that arises from different perceptions of sense-data.[14] More specifically, if one person has damage or decline in some function without realising it, a difference of perception, and therefore of opinion is all the more likely. For example:

- Person A and Person B go to the opera. After the event, Person A complains that the lead soprano was too shrill and had an edge to her voice. Person B protests that the soprano sang 'superbly' and while having a 'robust' sound had no hint of the alleged shrillness. What Person B has not realised is that they have developed high frequency loss progressively over a number of years, to the point where the decline is greater than normally expected for their age: they are unaware that they are missing auditory information that accounts for their perception of the singing being different from Person A's.
- Persons C and D go to enjoy a restaurant meal together. Person C complains that the food is tasteless and arrives at the conclusion that the chef must have changed. Person C does admit to being a bit 'sniffly' but protests that 'a slight sniffle doesn't take *all* the flavour from food'. Person D, on the other hand, finds the food 'delicious' and possibly even better than they had recalled. Unknown to Person C is that have contracted a nasal infection which is causing a mild case of anosmia[15].

14. Audi (2011) notes that proponents of the identity theory generally oppose the sense-datum theory, and cites Smart (1959) as a key discussion paper. That debate is beyond the scope of this thesis.

15. Loss of the sense of smell. Can be caused by infection or by brain injury. Can be temporary or permanent.

Similar examples can be imagined for visual, tactile or gustatory dysfunction. Whether there is damage to the receptor cells or to the associated specialised nerve cells or neural pathways, the lack of awareness of diminished function could easily lead a person to draw erroneous conclusions because the perceptual data on which they are relying are severely compromised.

For these examples to be valid the assumption is being made that the people are relying on sensory data and the resulting perception as being valid evidence for their position. That, in turn, presupposes that they are relying on other beliefs, albeit unconsciously held, that sensory data is an appropriate source of objective evidence and that their perceptions are accurate. As Berkeley says, 'They complain not of any want of evidence in their senses, and are out of all danger of becoming sceptics' (op. cit.).

Berkeley is comparing those who simply accept things as they perceive them with those who accept the challenges of reflecting on 'the nature of things', and he suggests the latter do not produce the calm and 'serenity of mind' that might be the anticipated reward for the philosopher. In the highly personal context of challenging one's ability to rely on one's perception, such a challenge is not a comfortable or comforting thing to do for most people. Doing so can be disquieting because of concern for one's perceptual faculties and the implications that can have for one's future. If acknowledged, the recognition of a failing in one's perceptual acuity necessarily challenges one's belief not only in being able to rely on one's senses but also regarding what is possible for oneself in the future. This connection between beliefs, and between expectations and beliefs, points to the value of a closer inspection of the role of other beliefs or meta-beliefs in shaping perception.[16]

That a person could be oblivious to neural damage or to diminished sensory function, at least in the early stages, is clearly possible; perhaps not so common, but unarguably possible. There are, however, sources of individual variation in sensation and perception that are not related to dysfunction and of which people are far more likely to be unaware.

2.1.2 Individual Variation

Research in recent years has been uncovering an answer to a long-standing question. Hollingham (2004) introduces relatively recent research results thus:

> It's a classic philosophical conundrum: how does my perception of the world differ from yours? Take a red rose, for example. We can probably agree it's red rather than blue, but what exactly is "red", and do I see the same red as you? And what about the distinctive smell—is my sense of what constitutes a rose's scent the same as yours?

16. Reserved for Chapter 4

> Philosophers have been wrestling with this question for centuries. Sensory scientists, too, have long been interested in why people report such different experiences of the same odours or flavours. Is it purely subjective, or based on some objective difference in their sensory experiences?
>
> (ibid.)

Part of the answer to just how differently people can perceive the same sensory stimulus is now being shown to have a genetic base.

> In the past few years geneticists have unearthed huge numbers of genes involved in the perception of taste, smell, touch and vision.
>
> (ibid.)

Taste varies greatly between individuals. Hollingham quotes neuroscientist Paul Breslin's example of different responses to the quinine used in tonic water:

> Most will tell you it's moderately bitter, some will say it's not bitter at all and a few will yell at you for trying to poison them.
>
> (ibid.)

Experiments with phenylthiocarbamide (PTC), a bitter flavour, revealed that even though the trait has 'reasonably distinct taster and non-taster groups' (ibid.), considerable variation was found within each group. The receptor gene was found to give rise to 'five different "flavours"', accounting for the variations found in the groups. This one gene codes 'for only one of the 23 different bitterness receptors' (ibid.); early indications suggest equally wide variance in other taste receptors, according to Hollingham's report.

Now, it is easy to conceive a situation in which one person (A) forms the belief, justified on the perception of extreme bitterness, that a gin and tonic is a 'disgusting' drink and disputes this point energetically with someone (B) at the other end of the taste spectrum who finds a gin and tonic close to 'perfection'. In this case, A has no idea that their genetic coding for taste has resulted in such a vastly different experience of gin and tonic to B's. Not having been schooled in genetics and not circulating in society that is inclined to be fascinated by the scientific explanation of the possible genesis of such experiences, A could take refuge in the secure 'knowledge' that what they taste is what they taste and it can only be right because that is their experience. Combined with other established beliefs about the matters on which people can and should have opinions at variance, it is easy to extend

the scenario to A's leaving the interaction with the opinion—that is, belief—that B is a person who has 'no taste'.[17]

As well as the differences in taste and olfaction, significant differences have been found in vision and in pain reception. Around eight per cent of people have some degree of colour blindness (ibid.), for example. Even with normal colour vision, there are significant differences in how colours are seen, and the differences appear to be largely genetic (ibid.; Paramel *et al.* 2004).

For an additional scenario, suppose Person L and Person C witness a vehicle speeding from the site of a store robbery. Person L is genetically coded to make less colour distinctions than Person C, and they are unaware of the possibility of such a phenomenon and of the fact that they are affected. In witness statements Person L is adamant—that is, strongly believes—that the car was dark brown while Person C is certain that the car was plum red.

With olfaction, humans have around 1,000 receptor genes (Hollingham 2004; Menashe *et al.* 2003). Of these, around 600 are 'pseudogenes' which are 'sequences that look like genes and are inherited like genes but have lost their function' (Hollingham 2004). However, the function has been lost relatively recently in evolutionary terms and Menashe et al. (2003), in examining an ethnically diverse group, found individual repertoires of smell receptors. While not conceding that the differences are genetically based, the research clearly establishes measurable differences in the sense of smell between individuals.

Further examples are unnecessary for establishing that individuals are frequently going to be unaware of some inherent differences in the way they perceive sensory stimuli compared to others. Clearly, this is more likely to be the case where the differences are more subtle. The person who is manifestly red–green colour blind is far more likely to become aware of it at some stage in their life compared to the one who fails to distinguish between deep shades of plum red and dark brown. It is reasonable to propose that, in everyday discourse, the latter difference will often be left as an agreed difference of opinion with respect to what the colour is *called* rather than realising that there is a genuine difference in perception. Still, the subtle differences could conceivably lead to significant effects in the person's life. If, for example, an acquaintance has a predilection for decorating in various shades of red, but also close shades (for example, plum, maroon, deep cherry), a genetic coding eliminating the ability to distinguish these subtleties of hue

17. As noted earlier, the effect of existing beliefs is discussed in a later section; but it is appropriate to consider the effect in such an example as this of the development of dysfunctional beliefs from a simple, unrecognised genetic difference causing a markedly different phenotype.

could have a person assessing the decorating colours as 'dull, truly dull, everything is just browns and earthy red'. This is not a trivial example if we consider how it could lead to judging the decorator, or to an inability to share experiences that others take pleasure in.

In summary, people can develop beliefs that are at variance with others simply because the receptor cells of their sensory apparatuses are genetically coded for a greater or lesser range of distinctions than are made by others. Whether such beliefs are irrational or false is debatable in this case. Is the colour really dark brown or plum red, for example? The 'true' answer probably lies with the majority perception or, preferably, with a spectroscopic measurement of the frequency of the reflected radiation. What is of greater significance is how that variant perceptual belief can lead to other truly irrational beliefs that can damage a person's relationships or quality of experience.

With respect to using sense-data as evidence for beliefs or for claiming knowledge, individual variation in perception—whether from genetic variation or some other source of variation—that the person is unaware of can be said to give rise to different perceptual beliefs. It may be difficult to classify many of these beliefs as false, but that they can vary so much from one individual to the next can arouse dispute and, in that, entail other irrational or false beliefs.

In these examples too, we see the reliance on what has been perceived—or believed to have been perceived, more precisely—as the arbiter of accuracy or truth. The object of interest lies in the response to a challenge. For example, *'How do you know that colour is plum red and not dark brown?'* could conceivably be answered with *'That's exactly the colour my Year 4 teacher showed us and called plum red'*. If the next challenge were simply, *'Are you sure?'*, it is not unusual to elicit the response *'Yes, of course'*. Morton (2003) proposes that 'sense data are well suited to be the ultimate evidence' because:

> It is that your perceptual beliefs are more certain than your other beliefs, and that your beliefs about your perceptual appearances are more certain than your other perceptual beliefs.
>
> (ibid., p. 23)

An additional challenge that could be posed in the example above is, *'Are you sure you're sure?'.* Polite responses might well include *'I know it'* as previously suggested by Morton (op. cit.), *'I'm absolutely certain',* and may be accompanied by some physical movement. Adhering to the belief requires some supporting evidence, and our interest is in what exactly that supporting evidence is. The proposal that emerges is that it is one form of sense data—one's own feelings (as discussed in Chapter 4).

2.1.3 Tricks of Perception

Beyond damage to sensory apparatuses or from genetic differences having the potential to lead to false beliefs, there is the general problem of perception; that is, false perceptions that can occur in the normal processes of processing sensory data.

A possible scenario: Person A, new to the city they find themselves in, asks Person B, a local resident, how long it should take to walk from their current position to one of the city's landmark buildings. Not previously having thought about it, Person B looks at the building in the distance. The weather is clear, there is bright sunlight, and the building is clearly visible. Person B estimates 20 minutes on the basis that the building 'looks about one kilometre away'.

While there are several assumptions Person B might be using that could be questioned, the principal matter here is their judgment of the distance. They formed a belief, based on their observation, that the walk should take about 20 minutes. Person A has arranged to meet a friend at the building and sets off happy in the expectation that they will make their appointment in good time.

Tversky and Kahneman (1982, p. 3) note a simple perceptual effect: that the more clearly an object is seen, the closer it appears to be.

Returning to the scenario, Person B is misled by the lighting conditions and is unaware that the landmark building in nearly two kilometres away. Consequently, Person A is left with a number of potentially unhappy outcomes including: taking much longer to walk the distance and therefore missing their appointment; or, having to catch a taxi to make the appointment on time.

It is reasonable to argue that a hypothetical, but likely, case such as that above easily demonstrates that perceptual effects could bring about, or interact with, emotions or moods and the combined effect can give rise to a belief or can change the 'intensity' with which the belief is held.[18]

For example, consider again the case of a famous landmark appearing closer because it can be clearly seen. A visitor to some land or city, without good geographical knowledge of the region, might easily judge, even *believe*, that it will be within easy walking distance and time. Some hours later, still walking, they realise that their judgment/belief was incorrect. While this is a relatively trivial example in one sense, it could turn out not to be. Giving this walker the additional attributes of an enthusiastic attitude, a high level of fitness, and an abundance of energy, we set them off on their journey to the landmark in the city they

18. The effect of emotions and feelings on beliefs is discussed predominantly in Chapter 3.

are visiting. Let's supply them with a perfect weather which gives excellent light falling on the landmark as well as crystal-clear air. Even though this enthusiastic walker takes three and a half hours to walk the 10 kilometres to the landmark, it is for them—because of all the attributes and environmental conditions we have given them—a 'short, easy' walk.

Now suppose that some months later they are helping a friend decide how to spend time in that same region. Because they had a wonderful walk to the famous landmark there, they recommend that the friend undertake the same expedition. The friend, being inclined to trust them and believe them, goes to the same region and undertakes the walk. However, the friend is not nearly so fit or energetic, and is beset with a cool breeze and dull, cloudy weather. Consequently their experience is quite contrary to that of the friend who recommended it. On return they utter expressions that are far removed from gratitude for the friend's advice and bluntly express their intention to pay no attention to any such advice in the future. Privately, they now harbour the new belief that you should not trust friends' advice about suitable activities when holidaying. The advice-giver now is convinced—viz. believes—that their (ex)friend is both ungrateful and lazy.

Here it is proposed that an everyday perceptual effect could affect a person's encoding of a memory, their emotional reaction to a scene (for example) which in turn could affect the memory coding, the person's ability to retrieve the memory, their emotional response to the recalled memory, the feelings arising from the emotional response, and their metacognitive assessment of their accuracy of the memory. Before examining some of these proposed effects, the simple challenges of perceptual illusions are briefly reviewed.

2.1.3.1 Visual Illusions

The cognitive psychology literature abounds with examples and discussion of perceptual illusions. Visual illusions grab most of the attention. While vision is not the only modality in which illusions occur, it is a modality on which people rely greatly and it dominates in the literature presumably because of this as well as possibly lending itself more readily to experiment design.

Reisberg (1997) states the issue which is a key part of this thesis:

> Pictures obviously have an existence independent of us, and independent of our understanding of perception. Thus, the raw material of the picture itself is separable from the interpretation we place on this raw material. As a consequence illusions are possible, in which the picture has one set of characteristics, but our perception of it has some other characteristics. Similarly, pictures can be ambiguous, with more than one interpretation possible.
>
> (ibid., p. 414)

Here, Reisberg is addressing the debate over the nature of mental imagery. He notes that psychologists such as Kosslyn (1980, 1983) and Finke (1980) have a view, as Reisberg put it, suggesting there is some 'raw material, akin to the picture, onto which we place an interpretation' (1997, p. 414). The view of others, including Reisberg's, is different. The debate about the nature of mental imagery is beyond the scope of this thesis.[19] However, that visual illusions occur is incontrovertible and the focus here is on how such illusions can be instrumental in leading to false beliefs when the observer relies on their perceptions as evidence.

Common examples of illusions include the Necker Cube (see Reisberg 1997, p.414), the duck/rabbit (ibid., p. 416), the Muller-Lyer illusion and the Ponzo illusion (Eysenck 2006; Styles 2005), and the Ebbinghaus illusion (Eysenck 2006). Reisberg observes that an image such as the duck/rabbit can be easily reinterpreted (for example, as a young girl with her hair tied back) and that this occurs frequently. By contrast, changing from seeing a duck to a rabbit—seeing a change of form—involves reference-frame changes, according to Reisberg (1997), and these are 'heavily dependent on hints and instructions' (ibid., 417). This conclusion is drawn largely from Reisberg and Chambers (1991) in which interpretation of shapes relied on an initial understanding of the forms. They enter into the debate (for example, with Finke *et al.* 1989) about whether mental images can be reconstrued. That debate, too, is outside the scope required here; but the relevant finding from Reisberg and Chambers is that 'what subjects can discover in their mental images seems bound by how the subjects understand their images' (1991, p. 347).

Necker Cubes and duck/rabbits might not be common outside the visual perception laboratories, but external images (or pictures)[20] or scenes that are open to more than one interpretation are common enough. In addition to the inherent illusory effect of some external images, others also note the importance of context in interpreting shapes (Palmer 1975; Styles 2005). For example, the light is on in an apartment bedroom, but a translucent blind is drawn. A silhouette of a shape that resembles a human torso can be seen and is positioned so that if a person's head were above the window frame, the silhouette could be interpreted as belonging to a body hanging from the ceiling.[21] The following day with the blind up, it can be seen that the image was formed by shirts being hung from the light fitting presumably so that they could dry. This could be anticipated by the observer by asking themself 'What *else* could cause such an image?'[22] The

19. Apart from the ongoing debate in the literature, it has been discussed in Thompson (2005).

20. Authors such as Reisberg used 'image' to refer to the mental image. To distinguish this, external image or picture is used henceforth to refer to the scene that is external to the observer.

21. Personal observation by the author.

22. The fortunate course of action taken by this observer.

translucence of the blind blurring the edges of the shirts, the light levels and the size of the window frame all contrived to allow for a possible interpretation of suicide. Now, place this illusion in the following context. The observer knows that a highly introverted young man lives alone in the unit; he rarely is seen with friends and is rarely seen looking 'happy' (however that is understood by the observer). With this context, it is easier for an observer to assume that they have witnessed a suicide and ring the police. The ambiguity of the image also has the potential to evoke biases of judgment.

James (1890/1950a, pp. 101–102) describes an illusion similar in nature in which he is certain that the chief engineer of the ship on which he is travelling has entered his state room. In fact, it was the reflection of his own hat and coat in the window.

A far more common illusion is that noted by James (ibid.) and familiar to modern train commuters. The train (A) on which one is passenger pulls in at a station next to another train (B) on the adjacent track. While train A is stationary, train B moves off and because it fills the view of the passenger window the passenger experiences a sensation of movement. Harmless enough, but still able to lead to false beliefs or inferences. Suppose the passenger on train A is only looking out the window and is also wearing headphones, as is very common now. As train A comes gently to a stop, the doors open and, at the same time, train B starts moving. Because the passenger is gazing only through the window to the view of train B and does not hear the door open, they believe the train is still moving. Now add a passenger who trips and injures themself badly as they are alighting from train A. When station staff ask for witnesses, our passenger is adamant the doors must have opened before the train stopped because they know they were still moving. The context matters, as may other assumptions held by the passenger; but the key element here is that the passenger has returned to what they believe to be irrefutable visual and kinaesthetic experience for evidence and are unaware that their perception is different from the external reality. Once again, this is a plausible scenario in which perceptual processes creating an illusory experience can lead to false beliefs.

The examples of visual illusions seem legion, but two more brief examples suffice to establish how visual illusion can lead to perceptions that, in turn, are used as a basis for false belief. Eysenck (2006), in discussing motion parallax, notes that 'the apparent speed of objects passing is faster the nearer they are to you' (ibid., p. 40). An observer, then, in an unfamiliar street is walking along as vehicle 1 passes them, going the opposite direction. The observer barely notices vehicle1 and is not attending to the road. They then hear tyres squealing behind them, from vehicle 1 attempting to avoid hitting a dog, and look up just in time to see vehicle 2 about 5 metres ahead of them beginning to break hard. Vehicle 2 hits the back of vehicle 1. The driver of vehicle 2 asks the pedestrian observer to verify that they were driving within the speed limit. However, with only seeing

vehicle 2 at a short distance, and in an unfamiliar street and therefore not having prior knowledge of distances, the pedestrian is certain—that is, believes—that vehicle 2 was exceeding the speed limit.

Both Eysenck (ibid.) and Styles (2005) discuss the effect of surface texture on the judgment of distance. Eysenck reports research showing:

> Observers were good at judging the distance of objects within 7 metres (23 feet) of them when the ground in between had a uniform texture. However, distances were systematically overestimated when there was a gap (e.g., a ditch) in the texture pattern. Distance judgments were also prone to error when the ground between the observer and the object was divided into two regions having very different textures (e.g., concrete, grass).
>
> (Eysenck 2006, p. 41)

Scenario: Two children are taunting each other, approximately 9 metres apart, with a ditch between them. Child A is standing on one side looking towards child B who is threatening to throw a projectile at child A. It is late afternoon, so the light is duller. Child A estimates that child B is too far away to be able to hit them from that distance, and consequently urges child B to 'have a go'. That poor estimation of distance in such a case could be comical or tragic, depending on the strength and accuracy of child B's projectile throwing. As with earlier examples, there are other beliefs involved; in this case, for example, a possible assumption about child B's strength. However, a key factor remains that the faulty perceptions arising from common visual illusions lead to beliefs which, in turn, have consequences when acted on.

Vision is not the only modality in which humans experience illusions. Styles (2005) reminds us that there are also kinaesthetic illusions such as the surface–weight illusion in which an object is perceived to be heavier when the surface is smoother. Cross-modal attention effects also produce illusions: for example, the speech seems to come from a ventriloquist's dummy when operated by skilled ventriloquist. Further examples of how these might lead to false beliefs are not needed. What they do point to is the role of attention in how our perceptions are shaped.

2.1.3.2 Attention, Inattention and Change Blindness

James (1890/1950b) laments the lack of interest from the English empiricists in the role of selective attention:

> Strange to say, so patent a fact as the perpetual presence of selective attention has received hardly any notice from psychologists of the English empiricist school.
>
> (ibid., p. 402)

In a foreshadowing of current findings he says:

> *My experience is what I agree to attend to.* Only those items which I *notice* shape my mind—without selective interest, experience is an utter chaos.
>
> (ibid., p. 402)

His continued upbraiding of the English empiricists is interesting to note because, once again, it points to findings that now occupy a great deal of current literature:

> These writers have, then, utterly ignored the glaring fact that subjective interest may, by laying its weighty index-finger on particular items of experience, so accent them as to give to the least frequent association far more power to shape our thought than the most frequent ones possess. The interest itself, though its genesis is doubtless perfectly *natural, makes* experience more than it is made by it.
>
> (ibid., p. 403, italics in original)

Selective attention is well recognised in cognitive psychology literature and Reisberg (1997) describes it as 'the processes through which you somehow select one input and "tune out" the rest' (ibid., pp. 81–82). Styles (2005) notes that selective attention is thought necessary simply because we cannot simultaneously process all the sensory input to which we are subject.

More relevant are the phenomena of *inattentional blindness* and *change blindness.* Eysenck (2006) defines inattentional blindness as the 'failure to notice an unexpected object appearing in a visual display' and change blindness as 'the failure to detect that an object has moved, changed, or disappeared' (p. 103). Both phenomena have received a great deal of attention in the recent times, particularly since the now-famous experiment conducted by Simons and Chabris (1999). In this experiment, subjects were asked, for example, to count the number of passes of a basketball between two teams; one team was wearing white shirts, the other black shirts. During the video one of two unexpected events occurred: either a woman carrying an open umbrella walked across the scene or a woman dressed in a full gorilla suit walked across the scene. Approximately 50% of the subjects failed to notice these events. In another experiment in more natural conditions, about 50% of pedestrians engaged in conversation by a person failed to notice that the person changed when a door was carried between them (the pedestrian) and the person who initially engaged them (Simons & Levin 1998).

One generally held false belief is that we have a 'clear and detailed visual representation of the world around us' (Eysenck 2006, p. 103). This phenomenon of overestimating the ability to accurately see events, and change, has been called *change blindness blindness* by Levin and others (Levin *et al.* 2000, 2002; Levin 2002). Levin summarises it thus:

> Many experiments have demonstrated that people fail to detect seemingly large visual changes in their environment. Despite these failures, most people confidently predict that they would see changes that are actually almost impossible to see. Therefore, in at least some situations visual experience is demonstrably not what people think it is.
>
> (Levin 2002, p. 111)

Chabris and Simons (2010) open *The Invisible Gorilla* with a striking example of how inattentional blindness and the beliefs both arising from it and associated with it can have a significant impact on a person's life. Kenny Conley was accused of perjury and obstructing justice following an incident in which police officers had mistakenly beaten another police officer while pursuing four men who had been involved in an armed robbery. The beaten police officer had been mistaken for one of the suspects by several police officers. While in pursuit of another suspect, Conley had had to run past the beating incident but had failed to notice it. Conley could not accept that he would not have noticed such an event. Neither could the jurors in the initial trial and convicted him.[23] Subsequently, Conley was tested by Chabris and Simons using the gorilla in the midst of the basketball game experiment, and was able to both count the passes and detect the gorilla.

Although this example does not show a false belief arising *directly* from a false perception, it does demonstrate that a critical *associated* and pre-existing belief can cause potentially false beliefs. Chabris and Simons' comment are pertinent:

> The illusion of attention is so ingrained and pervasive that everyone involved in the case of Kenny Conley was operating under a false notion of how the mind works: the mistaken belief that we pay attention to—and therefore should notice and remember—much more of the world around us than we actually do. Conley himself testified that he should have seen the brutal beating of Michael Cox had he actually run right past it.
>
> ....
>
> … if Conley was sufficiently focused on [the suspect he was chasing], in the way our subjects were focused on counting the basketball passes, it is entirely possible that he ran right past the assault and still failed to see it. If so, the only inaccurate part of Conley's testimony was his stated belief that he *should have* seen Cox. What is most striking about this case is that Conley's own testimony was the primary evidence that put him near the beating, and that evidence, combined with a misunderstanding of how the mind works, and the blue wall of silence erected by the other cops, led prosecutors to charge him with perjury and obstruction of justice.
>
> (ibid., Chap. 1, Section Kenny Conley's Invisible Gorilla)

23. This conviction was eventually overturned and the case finally dropped owing to legal technicalities.

Chabris and Simons' (ibid.) next example does demonstrate a false belief directly caused by expectations and inattention. In 2001 the captain of a nuclear submarine, the USS *Greeneville*, executed a rapid surfacing manoeuvre with the result that it sank a Japanese fishing vessel. Before executing the manoeuvre, the captain had checked the surface through the periscope, as is required, and believed there was no ship in the vicinity. Although the captain had looked in the correct direction, he did not see the fishing vessel. He is quoted by Chabris and Simons as saying 'I wasn't looking for it, nor did I expect it' (ibid., Chap. 1, Section The Nuclear Submarine and the Fishing Boat).

Many other examples are easy to find, or imagine. A common one, noted again by Chabris and Simons (ibid.), is that of the car that turns into a street and hits a bicyclist or a motorcyclist. After the event the driver will often say something like 'They simply weren't there when I looked' or 'It was clear when I pulled out'. An extension of this could be the driver then developing the view that the motorcyclist, for example, was obviously speeding excessively in order to arrive between the time when they were 'not there' (obviously, because the driver 'saw that it was clear') and the collision. Yet, as Chabris and Simons note: 'People don't see the motorcyclists because they aren't looking for motorcyclists' (ibid., Chap. 1, Section Ben Roethlisberger's Worst Interception).[24] If the driver blames the motorcyclist for speeding, we clearly have an instance of, first, motivations coming into play and, second, a belief being developed to satisfy the motivations. The subject of emotions and motivations is related to the discussion in Chapter 4; for now, it suffices that it is shown that inattentional blindness or change blindness *can lead* to the development of false beliefs or false claims of knowledge.[25]

'Blindness' to change is not restricted to vision. Vitevitch (2003) also reports on *change deafness*. In this paper he reports that 'at least 40% of the participants failed to detect the change in talker' (ibid., p. 333) where the talker was reading a list of words of varying lexical difficulty and the listener had to repeat the words. They specifically used an integral stimulus—spoken words—with attention drawn to the linguistic stimulus rather than the indexical stimulus. They conclude that their results 'support the hypothesis that attention must be directed toward a particular dimension of a stimulus for a change in that dimension to be detected' (ibid., p. 340).

Real-world experimentation would need to verify the following hypothetical example; however, based on experimental results in which one person is exchanged for another and the change is not noticed, it is plausible. Person A, discussing details of an important

24. Chabris and Simons report on studies showing that where bicyclists and pedestrians are more common (mainly in Europe) they are less likely to hit by a car.

25. In Chapter 4, there is a return to the metacognitive aspects of these phenomena.

contract on the phone, is arguing with Person B over payment arrangements. During the discussion there is a distraction and Person B is replaced by Person C (same sex as Person B and similar but, with attention, noticeably different vocal characteristics). Person A, focusing on getting their way in the negotiation, does not notice the change in voice. Person C agrees to a number of Person A's demands. Later, in the litigation arising from the failed contractual arrangements, Person A is adamant—that is, has a strongly held belief—that Person B agreed to the demands.

In summary, then, it is clear from the research that people are frequently unaware of how much they do not notice. A general false belief that appears to pervade our thinking is that we are able to see and hear far more than we do. It is not unreasonable to suggest that this belief is itself dependent on inattentional blindness and change blindness for its genesis. Once in place, it creates a foundation for entailing associated beliefs and motivations that, in turn, lead to incorrect inferences (for example, 'the motorcycle had to be speeding') or false claims of knowledge (for example, 'I know I wouldn't miss something like that').

2.1.4 Perceptions and False Beliefs – Summary

There is a vast literature on sensory mechanisms and perception. Examples of illusions—especially visual illusions—abound, and a great deal of research examines the nature of perceptual processes both in healthy people and cases where there is some disorder such as extinction.

This section (*Perception and Perceptual Mechanisms*) has examined the proposition that people will appeal to the apparently objective evidence obtained through their senses in order to justify a belief or to claim knowledge. The philosophical problems with justifying belief or knowledge with sense-data have been raised and it was argued that it is valid that current research from cognitive psychology and neuroscience inform the debate on how a belief might be justified, in an approach akin to what Morton calls 'naturalistic empiricism' (Morton 1997, p. 42).

General agreement exists in the literature about the resort to sense-data. However, the challenges with that course that have been demonstrated are:

1 Perceptions of the same sensory stimulus vary from one individual to the next from causes such as damage to sense organs or genetic differences causing different perceptions.

2 There are inherent illusions involved in some sets of external sensory stimuli (for example, the Ponzo illusion) and these can lead to false inferences and judgments.

3 The limitations of processing result in phenomena such as inattentional blindness and change blindness; but, the general lack of awareness of this phenomenon results

in the widespread misconception that people believe they see and hear far more than they actually do.

4 Different perceptions lead to different inferences. These inferences can be the basis for false beliefs and erroneous claims of knowledge.

In a thesis focused only on how perceptual mechanisms can lead to false inferences, much lengthier discussion of some of the perceptual phenomena also covered would be required, along with additional phenomena such as multitasking, adaptation and extinction. Returning to the focus of this thesis, however, it is recalled that the *maintenance* of false beliefs is of as much interest as is their genesis. That is, in the face of contradictory evidence or soundly argued counter proposals, how does a person maintain a false belief or a false claim to knowledge? It has been sufficiently established that phenomena such as visual illusions and variations in perceptual processing can lead to false beliefs or provide a cognitive foundation from which false beliefs can be generated. The questions that arise from these findings point to a key finding discussed in Chapter 4 relating to the ultimate reliance on sense data in the form of feelings.

In extending the scenarios offered in many of the examples, existing beliefs and motivations are frequently entailed. When considering not only the genesis and maintenance of false beliefs but also the effects of them, this entailment is a significant consideration which warrants the extended discussion in following chapters. A challenge to the belief in the ability to rely on one's own sense data can challenge a whole system of beliefs: this, too, is discussed in Chapter 4.

2.2 Heuristics and Biases

> The weighing of evidence and the formation of belief are basic elements of human thought. The question of how to evaluate evidence and assess confidence has been addressed from a normative perspective by philosophers and statisticians; it has also been investigated by psychologists and decision researchers. One of the major findings that has emerged from this research is that people are often more confident in their judgments than is warranted by the facts.
>
> (Griffin & Tversky 2002, p. 230)

Reisberg (1997) opens a discussion on judgment—especially as a prelude to decision making—by inviting consideration of what people *do* with knowledge once they have acquired it. Here, his use of the term 'knowledge' is not used to invite philosophical discussion of the nature of knowledge. Rather, it appears to be employed as a claim to knowledge based on sense-data, experience that provides memories considered to entail knowledge or be knowledge, or beliefs *considered* justified that can be used in the manner

suggested by Morton (2003) where the belief can be trusted as something that can be used in forming other beliefs. However, from the preceding discussion relating to false beliefs arising from perceptual mechanisms, it follows that many decisions and judgments made by people are made on the basis of false beliefs.

Tversky and Kahneman note some of the perceptual effects already discussed:

> … the apparent distance of an object is determined by its clarity. The more sharply the object is seen, the closer it appears to be … distances are often overestimated when visibility is poor because the contours of objects are blurred. On the other hand, distances are often underestimated when visibility is good because the objects are seen sharply. Thus, the reliance on clarity as an indication of distance leads to common biases. Such biases are also found in the intuitive judgment of probability.
>
> (Tversky & Kahneman 1982, p. 3)

If someone is relying on a false belief or a false claim to knowledge, but yet believing that their grounds are sound and that they are being rational, it is given that the derived judgments and decisions are at considerable risk of producing unwanted results.

It is probably fair to say that the work of Amos Tversky and Daniel Kahneman has led to a revolution in our understanding of rationality. For this reason, an extensive review of their findings is warranted for the development of this thesis. The connection of their work with this thesis is captured in one of their landmark publications:

> Many decisions are based on beliefs concerning the likelihood of uncertain events such as the outcome of an election, the guilt of a defendant, or the future value of the dollar. These beliefs are usually expressed in statements such as 'I think that …', 'chances are …', 'it is unlikely that …', and so forth. Occasionally, beliefs concerning uncertain events are expressed in numerical form as odds or subjective probabilities. What determines such beliefs? How do people assess the probability of an uncertain event or the values of an uncertain quantity?
>
> (ibid., p. 3)

The work by Tversky and Kahneman provides insight into systematic errors that are made unconsciously. While some are obviously directly related to perceptual effects, others interpret perceptions on the basis of other unconscious beliefs and interconnected memories. This point will be developed as the discussion develops, for it is clearly shown that false beliefs can develop as a result of systematic errors of judgment, but it is also proposed that there is often a base of beliefs associated with the systematic errors or which may be a conscious expression that 'explains' or justifies the systematic error process.

Although the thesis is examining how people create and maintain false beliefs, the motivation for that focus arises from the potential effects of those false beliefs. Making decisions can be a source of anxiety, disappointment, fear, concern, worry and interest for

many people. Decisions covering seemingly inconsequential events such as choosing the 'best' meal at a restaurant, can give rise to disappointment and annoyance if, after the fact, they prove to be the 'wrong' decision—for example, indigestion can assure the decider that they made a bad choice of meal and that the time in a restaurant was not nearly so well spent as it might have been. In the business world, making decisions is frequently made with the eye on 'the bottom line', and whether the bottom line was considered or not at the time of deciding, it inevitably will be affected. Wrong decisions can cost dearly in the business world. At the level of governmental policy, decisions have far-reaching effects on the health of society. Janis and Mann (1977) quote John F. Kennedy after his miscalculation in approving the Bay of Pigs invasion, and they suggest that everybody asks themselves the same question when confronted with the consequences of their own 'personal Bay of Pigs':

> How could I have been so stupid?
>
> (ibid., p. xv)

Gilovich and Griffin (2002) note that proponents of the notion of the 'rational actor' do allow that people make mistakes in judgments, but insist that the mistakes are unsystematic. The work of Tversky and Kahneman considers that intuitive judgment does not involve processes that are simpler than those of rational models but are of a categorically different kind (ibid.). Biases and heuristics lead to *systematic* errors. Both Eysenck (2006) and Gilovich and Griffin (2002) observe that debate remains relating to heuristics and biases program. However, despite airing considerable coverage of critiques of the program, Gilovich and Griffin suggest that:

> … the heuristics and biases program has weathered its various critiques and remains a vigorous and still developing perspective on human judgment.
>
> (ibid., p. 15)

With respect to the types of beliefs in focus here, it is proposed that the heuristics and biases program has significant explanatory power.

2.2.1 Judgments and Heuristics[26]

Tversky and Kahneman (1982) suggest that people rely on a 'limited number of heuristic principles which reduce the complex tasks of assessing probabilities and predicting values to simpler judgmental operations' (ibid., p. 3). In later work, Kahneman and Frederick (2002) emphasise that people do not normally deal in exhaustive probability analysis.

26. So much of the literature uses 'judgment' rather than 'judgement', that it is employed throughout. However, UK spellings are being used for all other instances for which there are US and UK variants.

A judgmental heuristic is 'a strategy–whether deliberate or not–that relies on a natural assessment to produce an estimation or a prediction' (Tversky & Kahneman 2002, p. 20). Tversky and Kahneman note that one of the 'manifestations of a heuristic is the *relative neglect of other considerations*' (ibid., p. 20, emphasis added). An example is giving too much weight to a stereotype that a child resembles when predicting the child's future vocation, without considering the base rate frequencies of the relevant occupations.

The heuristics are simple and efficient because they 'piggyback' on computational processes which the mind has evolved to make (Gilovich & Griffin 2002). Intuitive judgments are thought to occupy a position 'between the automatic parallel operations of perception and the controlled serial operations of reasoning' (Kahneman & Frederick 2002, p. 50). In evolutionary terms it is plausible that the resulting fuzzy boundary between perception and judgment could be helpful: a stranger who appears menacing, for example, is linked with a prediction of future possible harm thus instigating a suitably timed retreat or protective response. However, now, when mental representations are invoked by language, and the represented objects are not present, they still carry 'causal propensities and an affective charge' (ibid., p. 50).

So, even though the limiting of the heuristic principles employed can be quite useful, it can sometimes lead to errors which are severe and systematic (Tversky & Kahneman 1982). This latter feature stands in contrast to the classical models' view that errors are unsystematic.

In their 1974 paper Tversky and Kahneman (1974, 1982) present three principal heuristics:

1 representativeness

2 availability

3 anchoring and adjustment.

This literature on heuristics and biases is extensive and much of it is not required for this thesis (for example, the more involved statistical discussions). Principally, the focus below is on appropriate examples of the heuristics and biases and ways that they can lead to false beliefs or false claims to knowledge. Additionally, the inevitably of the heuristics' entailing of other beliefs will emerge.

2.2.1.1 Representativeness

Tversky and Kahneman define representativeness as:

> …an assessment of the degree of correspondence between a sample and a population, an instance and a category, an act and an actor, or more generally,

> between an outcome and a model. The model may refer to a person, a coin or the world economy, and the respective outcomes could be marital status, a sequence of heads and tails, or the current price of gold.
>
> (Tversky & Kahneman 2002, p. 22)

Where the probabilistic question is of the type *What is the probability that object A belongs to class B?*, people rely on this representativeness heuristic (Tversky & Kahneman 1982). That is, the degree to which A resembles B affects the evaluation of the degree to which A is representative of B.

> For example, when A is highly representative of B, the probability that A originates from B is judged to be high. On the other hand, if A is not similar to B, the probability that A originates from B is judged to be low.
>
> (ibid., p. 4)

Tversky and Kahneman (ibid.) report that errors arise because representativeness, or similarity, is not affected by several factors that do affect probability. These are:

- insensitivity to prior probability of outcomes
- insensitivity to sample size
- misconceptions of chance
- insensitivity to predictability
- the illusion of validity
- misconceptions of regression

For the purposes of the following discussion it is instructive to use the example provided by Tversky and Kahneman.

> For an illustration of judgment by representativeness, consider an individual who has been described by a former neighbour as follows: 'Steve is very shy and withdrawn, invariably helpful, but with little interest in people, or in the world of reality. A meek and tidy soul, he has a need for order and structure, and a passion for detail.' How do people assess the probability that Steve is engaged in a particular occupation from a list of possibilities (for example, farmer, salesman, airline pilot, librarian, or physician)? How do people order these occupations from most to least likely? In the representativeness heuristic, the probability that Steve is a librarian, for example, is assessed by the degree to which he is representative of, or similar to, the stereotype of a librarian. Indeed, research with problems of this type has shown that people order the occupations by probability and by similarity in exactly the same way (Kahneman, 1973, 4).
>
> (ibid., p. 4)

Insensitivity to prior probability of outcomes

Tversky and Kahneman (ibid.) note that prior probability, or base-rate frequency, of factors that affect outcomes should have a major effect on probability even though they have no effect on representativeness. If Steve, in the above example, is perceived as being representative of a stereotypical librarian, the probability of his being a librarian rather than a farmer will be evaluated as being greater even though there are many more farmers in the population.

Interestingly, in one experiment, when subjects were simply given the numbers of engineers and lawyers in a population they used the probabilities correctly. But, Tversky and Kahneman observed that once a description was provided, even if it was totally uninformative, prior probabilities were 'effectively ignored' (ibid., p. 5).

Two considerations arise from such experimental observations. First, it shows that, given a description, people will derive a view—most likely expressed as a belief—based on representativeness. They might not use the word 'believe', instead saying something such as *'There's no way he's a farmer', 'He's got to be a librarian',* etc. Such statements indicate that the speaker considers their proposition as being true (or, perhaps, as being highly likely to be true) and would inform any interaction with the person (at least initially), thereby satisfying criteria (1) and (3) of the proposed working definition of belief (see page 27).

The second issue is that the reliance on the stereotype for the librarian, in the example given, implies pre-existing beliefs about librarians and the validity of the related stereotype. Discussion of this entailment is better pursued later in Chapter 4. However, the indications are that although the heuristics and biases may help someone *create* a false belief, they also appear on many occasions to be associated with existing beliefs such as believing that a stereotype is a reliable description of a person, entity or event.

Insensitivity to sample size

In the studies of Tversky and Kahneman, when people are evaluating probabilities of results from a specified population, they appear to make their judgments in ways which were 'essentially independent of sample size' (ibid., p. 6). For example, in a scenario involving a large hospital (45 births per day) and a small hospital (15 births per day), subjects were asked to judge which of the two hospitals recorded more days in a year during which the percentage of male births was 60%[27]. Most subjects estimated the same in both hospitals, thereby ignoring the effect of sample size in which the larger of the two hospitals would be the one less likely to vary from 50%. Tversky and Kahneman observe

27. Normally, of course, the percentage of males born (in a suitably sized sample) is 50%.

that the 'fundamental notion of statistics is evidently not part of people's repertoire of intuitions' (ibid., p.6).

An inaccurate judgment such as that above may not in itself be particularly problematic. However, it is a useful activity to speculate on the conversations that could occur if, for example, the large hospital had several consecutive days in which the the percentage of male births exceeded 60%. Those who give credence to astrology, for example, might pore over charts to review planet alignment or misalignment at the times of conception. Others might suggest that Gaia is trying to redress the balance of sex proportion because of the number of males dying in wars and skirmishes or through suicide. While amusing to some, and the stuff of C-grade movies to others, such grossly unjustified beliefs do occur. Those coming to such conclusions most likely fail to notice the weeks in which female births greatly outnumber male births in the large hospital and that the overall proportion for the whole year is, as one would expect, around 50:50.

In a more disconcerting example involving posterior probability, Tversky and Kahneman (ibid.) demonstrate that people's intuitive judgments are dominated by the sample proportion and not by the size of the sample.

> Imagine an urn filled with balls, of which 2/3 are of one colour and 1/3 of another. One individual has drawn 5 balls from the urn, and found that 4 were red and 1 was white. Another individual has drawn 20 balls and found that 12 were red and 8 were white. Which of the two individuals should feel more confident that the urn contains 2/3 red balls and 1/3 white balls, rather than the opposite? What odds should each individual give?
>
> In this problem, the correct posterior odds are 8 to 1 for the 4:1 sample and 16 to 1 for the 12:8 sample, assuming equal prior probabilities. However, most people feel that the first sample provides much stronger evidence for the hypothesis that the urn is predominantly red, because the proportion of red balls is larger in the first than in the second sample … The underestimation of the impact of evidence has been observed repeatedly in problems of this type …
>
> (ibid., pp. 6–7)

According to Griffin and Tversky (2002) *strength* and *weight* of evidence are combined using rules prescribed by statistical theory and the calculus of change. With probability, posterior probability as applicable to the above example should be determined by correctly combining sample proportion and sample size. They note, however, that people generally do not follow the rules of probability and statistics when combining strength and weight of evidence. They also note that underconfidence is frequently observed as well as overconfidence. It is only fair to note here that many might consider the rules of

probability to be arcane and ultimately complex.[28] Most people will not even be aware of distinctions such as posterior probability, let alone know how to calculate it; so it is no wonder that they do not use the rules. What is significant, of course, is that there is some concept, process or belief that they resort to—albeit unconsciously—in order to make a judgment.

Griffin and Tversky's hypothesis is that people 'focus on the strength of evidence—at least, as they perceive it—and then make some adjustment in response to its weight' (ibid., p. 231).

> This hypothesis suggests a pattern of overconfidence and underconfidence. If people are highly sensitive to variations in the extremeness of evidence and not sufficiently sensitive to variations in its credence or predictive validity, then judgments will be overconfident when strength is high and weight is low, and they will be underconfident when weight is high and strength is low.
>
> (ibid., p. 232)

As an example, they suggest the process involved when considering a letter of recommendation for, say, accepting a student into a graduate program. The warmth of the recommendation, they suggest, is what takes the recipient's attention first and after that the recipient makes allowances for the writer's limited knowledge. Griffin and Tversky suggest, however, that the research shows that in cases such as this and many others that the adjustment made is 'generally insufficient' (ibid., p. 231). The possible problems here are easy to imagine. For whatever reason, the writer of the letter of recommendation has high expectations of the student's performance based on a perception of ability and further potential. Consequently, they write a glowing recommendation that does not take into account such factors as the student's true motivational levels, their ability to apply themselves under pressure, and their ability to establish the ongoing material and psychological support they might need during the course of their studies. If the student's motivation is poor and they give up when circumstances become demanding there has been a loss of opportunity all around.

As a further example, consider the person (A) whose friend (B) has had financial success with three purchases of shares in the last four months. Led on by B's celebratory exuberance, A accepts B's recommendation for a share purchase and invests a significant portion of his savings. Meanwhile, B has not mentioned the two purchases in the previous four months that resulted in significant losses; she has managed to attribute those losses to factors other than her own judgment. A has not asked B about losses because A likes and

28. The author would wager that few people find probability the most engaging topic during their schooling.

respects B as a friend and is duly influenced (Cialdini 1984), and it seems that three 'wins' in four months is so remarkable that B really is representative of a 'real trader'. The potential danger for A's investment is obvious.

Griffin and Tversky end their paper (2002) with the strong admonition that the significance of overconfidence cannot be overstated, principally because confidence leads to action. But, ill-founded overconfidence can lead to disastrous decisions. For example, at the time of writing, much of the world is staggering financially as a result of the unbridled overconfidence in notions such as the 'New Economy' and financial products that led to near financial collapse on a global scale.[29]

Misconceptions of chance

As a chance process proceeds—for example, tossing a coin or roulette—deviations are not 'corrected' but rather are diluted (Tversky & Kahneman 1982). One of Tversky and Kahneman's (ibid.) examples is that of people observing roulette and observing a run of red numbers. Most people will come to expect that a black number is 'due' even though the observed sequence is relatively short—this is the 'gambler's fallacy'.

People expect that the essential characteristics of a sequence of events generated by a random process will be evident even in a short sequence. Tversky and Kahneman give the example of people thinking that in six tosses of a coin the sequence H-T-H-T-T-H is more likely than H-H-H-T-T-T.

As with the examples in *Insensitivity to prior probability of outcomes*, this bias can lead to the creation of false beliefs and it also implies the existence of beliefs that influence the new belief. Examples of the false beliefs that can be formed are relatively straightforward and should need little discussion. For example:

- *'I know red is going to win next time, there have been 8 black numbers in a row.'*
- *'There has to be a Tail next time; the "law of averages" means you're never going to get that many Heads in a row.'*
- *'There can't possibly be another bad year for investment; you just don't get more than two bad years in a row. Besides, there's the economic cycle; things have to turn around.'*

At this stage, the potential dangers of beliefs such as these need no further discussion, the point having been established sufficiently above (in *Insensitivity to prior probability of outcomes* and *Insensitivity to sample size*). Again, propositions such as those given here imply that the speaker believes them true (or very likely true) and that they are guiding

29. And might still do (16 October 2011). Same sentiment in 2014.

actions. In this, then, they too are satisfying criteria (1) and (3) of the proposed belief definition (page 27).

Disturbingly, even experienced research psychologists fall prey to these misconceptions (ibid., p. 7). As a consequence, Tversky and Kahneman suggest, researchers have put too much store in small samples and overestimated the reproducibility or results.

As with earlier examples, it is interesting to observe that pre-existing beliefs seem to be involved. Perhaps they are being used in the way that Morton (1997, 2003) suggests, and discussed earlier, where beliefs are taken as a basis for forming other knowledge or where they are considered 'high quality' (1997, p. 112) and they are considered 'knowledge'. Again, as noted earlier, further discussion of the interaction of beliefs is reserved for Chapter 4. What is of interest here is that the heuristics and biases are the result of more than illusion or some perceptual dysfunction as well as the emerging evidence that there is a reliance on sense data in the form of feelings to support the judgment or belief formed.

Insensitivity to predictability

Favourable descriptions affect predictions (Tversky & Kahneman 1982), particularly numerical predictions such as the future value of a company or its shares, the value of a commodity, or the value of property. Tversky and Kahneman note that:

> The degree to which the description is favorable is unaffected by the reliability of that description or by the degree to which it permits accurate prediction.
>
> (ibid., p. 8)

When given a favourable description of a company, for example, subjects mainly predicted high share values regardless of whether the description imparted information relevant to profit or not. An enthusiastic investor, having been given a favourable description of the company, could easily develop a belief that the share value of the company will inevitably increase; all the while taking no regard to the relevance of the description or of factors external to the company that could affect its share value (for example, a long-term downtown in the value of commodities).

Extensive consideration of the number of factors that could contrive to make the description favourable would be a diversion; but a brief consideration is worthwhile. For example, the description could:

- contain language that appeals to the person's emotions
- fit with existing beliefs about the specific *type* of business
- bolster an existing belief about the value of the company.

Much of the literature reviewed on heuristics and biases concentrates on describing the *nature* of the heuristics and biases, as is expected, and often points to the consequences. Occasionally, there is mention of the involvement of beliefs. However, here, it is proposed that the heuristics and biases may *frequently* engage existing beliefs in the formation of new beliefs. The question of how the heuristics and biases themselves are established becomes both more pressing and intriguing. Morton's suggestion that a belief fails as knowledge if it 'falls short of some expectation that we might reasonably have of it' (2003, p. 90) suggests a possible corollary as one influencing factor: that if a judgment fits the expectation derived from an existing belief (or set of beliefs), then it is likely to be *believed* or taken as knowledge. Similarly, a judgment that supports an emotional goal, for example, may well have a greater chance of being believed or taken as knowledge. The discussion of the complex interactions between beliefs and affect systems and differentiated neural pathways is reserved for Chapters 3 and 4.

Suggestions of this of this complexity occur in Sloman (2002) in which he argues an empirical case for two systems of reasoning, drawing on his own research as well as that of many others. The two systems are:

1 an *associative* system that 'encodes and processes statistical regularities of its environment, frequencies and correlations amongst the various features of the world' (ibid., pp. 380–381)

2 a *rule-based* system.

Sloman (ibid.) refers to Evans and Over (1996)[30] to help distinguish the two systems.

> The associative system is generally useful for achieving one's goals; the rule-based system is more adept at ensuring that one's conclusions are sanctioned by a normative theory.
>
> (ibid., p. 382)

Perhaps even more pertinent in elucidating how a favourable description can lead to reliance on the representativeness heuristic is the *affect* heuristic proposed by Slovic et al. (2002). They suggest that:

> *Affective responses* occur rapidly and automatically – note how quickly you sense the feelings associated with the stimulus words *treasure* and *hate*. We argue that reliance on such feelings can be characterised as the *affect heuristic.*
>
> (ibid., p. 397)

30. Evans, J St B T & Over, D E (1996), *Rationality and Reasoning*, Psychology Press, Hove, England. Not available for viewing at the time of writing.

Slovic et al. look to Zajonc (1980), an early proponent of the role of affect in decision making, and take up his argument that:

> … affective reactions to stimuli are often the very first reactions, occurring automatically and subsequently guiding information processing and judgment.
>
> (Slovic *et al.* 2002, p.398)

Zajonc proposes that:

> … perhaps all perceptions contain some affect. We do not just see 'a house', we see 'a *handsome* house', 'an *ugly* house', or 'a *pretentious* house'. We do not just read an article on attitude change, on cognitive dissonance, or on herbicides. We read an 'exciting' article on attitude change, an 'important' article on cognitive dissonance, or a 'trivial' article on herbicides.
>
> (Zajonc 1980, pp. 153–154)

These suggestions offer some insight into how it is that a favourable description could so easily and quickly lead to a reliance on representativeness. Slovic et al. (op. cit.) also look to the work of Damasio (1994) whose somatic marker hypothesis gives greater weight to the argument in favour of the existence and effect of an affect heuristic. They appear to suggest that the affect heuristic may operate separately from some of the other heuristics:

> Just as imaginability, memorability, and similarity serve as cues for probability judgments (e.g., the availability and representativeness heuristics), affect may serve as a cue for many important judgments.
>
> (Slovic *et al.* 2002, p. 400)

Further investigation of the role of affect is reserved for Chapter 3 since it is not clear, at this stage, that the affect heuristic as proposed can ever operate independently of the other heuristics. At this point, the implication of an affect component, especially in the associative system proposed by Sloman (op. cit.), points to possible answers as to how it is that a favourable description can cause the ignoring of information both relevant to and important for sound judgment.

The illusion of validity

The illusion of validity is the 'unwarranted confidence which is produced by a good fit between the predicted outcome and the input information'. Tversky and Kahneman (1982) point out that this illusion persists even when people are aware of factors which limit the accuracy of their predictions. As an example they offer the predicting of final marks for two students—one with all Bs in their first year, and the other with many As and Cs. People are more confident predicting the mark of the student with Bs. The implications of such predictions could be dire.

The observed confidence, Tversky and Kahneman propose, arises from the selecting an outcome which is most representative of the input. For example, even if the supplied description of a person is a greatly outdated one of a librarian, and is scant, and unreliable, people will tend to confidently predict the person is a librarian.

Other expressions of confidence that might easily be encountered are easily recalled or imagined: *'No, he's only ever had Passes and Credits, he won't make a good postgraduate student'; 'Their investments have outperformed 95% of their competitors for the last 5 years, they've got to be a sure thing*. As with earlier examples, the word 'believe' is not used, but considering the proposition true is implied, actions are being guided by the proposition, and desires are evident (at least in the second example). Thus, the three proposed criteria for belief are met. Given the statements as presented, there is no indication of the range of factors that could affect the prediction being taken properly into account. There are students who have unimpressive undergraduate records who go on to excellence at postgraduate levels; there are companies that perform well but which suddenly fall over because of the change in just one key manager, for example.

As an example of the importance of *not* allowing a current trend in marks to affect judgment, the author has seen a child in a 'special education' class who was presumed to not have great academic potential manage to progress directly to the top class for the next year. Had the history of academic performance been allowed to inform prediction, the expectations may well have affected how the child was taught and, consequently, have restricted the child's progress.

Again, Tversky and Kahneman claim that even psychologists in possession of knowledge of a vast literature attesting to the fallibility of selection interviews still continue to rely on such interviews. This claim was made in 1974, so it deserves to be examined in the context of developments in the last 30 years. Nevertheless, the central tendency to rely on criteria which engage representativeness remains a concern where suitability of selection (for jobs) can be affected by inaccurate prediction.

In the examples above, the input information is readily available and, therefore, there may seem little apparent need to expend cognitive effort imagining other possible outcomes. For the student who has only achieved mediocre results in undergraduate years, for example, there is no apparent evidence for a significantly different performance level at postgraduate level. Even if a person takes time to consider such factors as a different attitude arising from developed maturity and from being able to focus on areas of genuine interest or passion, the possible outcomes are *uncertain* and, therefore, likely to seem far less probable. More pertinently, if we take the affect heuristic into account, the feelings of uncertainty are far less comfortable than those of confidence (for most people, at least). Hence, it is not surprising that there is there a reliance on the input information—it is

readily available and it produces better feelings. In this, there are two forms of perceptual evidence now at work: the easily accessible input information and the *feelings* of confidence. It is the latter of these two that, ultimately, is the more significant (as discussed in Chapter 4).

Misconceptions of regression

Regression towards a mean value occurs frequently in life (ibid.): for example, heights of parents and their children, the performance of students in consecutive examinations. Despite this, people either fail to recognise regression when it occurs, or 'they often invent spurious causal explanations for it' (ibid., p. 10) when they do recognise it.

In one illuminating example involving flight training, Tversky and Kahneman (ibid.) elucidate how failing to understand the effect of regression leads to overestimation of the effectiveness of punishment and underestimation of the effectiveness of reward. This can lead to disastrous effects in training and educating, and in social interactions.

> Consequently, ... one is most often rewarded for punishing others and most often punished for rewarding them. ... In fact, the elusive role of regression in determining the apparent consequences of reward and punishment seems to have escaped the notice of students of this area.
>
> (ibid.)

Whether the unfortunate pattern of reward and punishment has persisted since Tversky and Kahneman's early research has not been established. However, the principal focus, again, is the lack of understanding of regression and how it can easily lead to false beliefs. For example, suppose a child of two tall parents only achieves a height that is considered short. In the spirit of the 'spurious causal explanations' observed by Tversky and Kahneman (op. cit.) it is reasonable to suppose possible suggestions such as: *'Obviously, the parents didn't feed the child properly in the important growing years'; 'Looks to me like the milkman was on extra duties'; 'Perhaps the child is retarded'.*

Despite being somewhat extreme, these suggestions remain realistic. With respect to their formation, they invite consideration of the involvement of other heuristics. For example, one judge has observed in the past two tall parents with a child who is shorter than they and who was known to have a slight intellectual impairment. Now, when seeing this other child who is also shorter than both parents, a representation that is immediately *available* is that of the first, intellectually impaired, child. Hence, with no consideration of regression to the mean, the representativeness heuristic is applied.

It is proposed, again, that contributing factors are those discussed above (*The illusion of validity*). Perceptual information is readily available and this is highlighted in the next

section (*Availability*). As well, if this perceptual information is used as the sole representational input, it could activate existing beliefs[31] and evoke the desirable feelings of certainty.

2.2.1.2 Availability

The second judgmental heuristic, *availability*, occurs when 'people assess the frequency of a class or the probability of an event by the ease with which instances or occurrences can be brought to mind' (ibid., p. 11).

While availability is useful in assessing frequency or probability, it is affected by factors other than frequency and probability. Relying on availability leads to biases based on:

- retrievability of instances
- effectiveness of the search set
- imaginability
- illusory correlation (ibid.)

Retrievability of instances, effectiveness of a search set

If people are asked to judge the size of a class, the ease with which they can *retrieve* examples affects the judgment. Where examples are more easily retrieved, that class will be judged as more numerous than one with as many members but whose examples are less easily retrieved. For example (ibid.), subjects were asked to judge whether a list of men's and women's names contained more names of men or women. Subjects judged that the sex in the list that had more relatively famous names also was the more numerous in the list.

Also, where it is easier to access or think of items for a search set, the frequency of occurrence of such items will be judged higher (ibid.). For example, judging whether abstract words such as 'love' occur more or less frequently than concrete ones such as 'door'. Since abstract words seem to be easier for people to think of (for contexts in which they occur), they judge them as occurring more frequently and in more contexts.

Judging the relative proportion of abstract words to concrete words may not be a significant exercise for success in life. However, as shown in the example above (in *Misconceptions of regression*), availability can have serious consequences for how others are viewed.

Recent events also affect subjective probability. Witnessing a car accident, for example, temporarily increases the subjective probability of road accidents (ibid.).

31. A factor also discussed in Chapter 4.

Schwartz and Vaughan (2002) point to the significance of the ease of recall in relation to personal relevance. They suggest that people rely on the ease of recall when the matter is of low personal relevance, but on the content which is recalled when the matter is of high personal relevance. One of the problems, they suggest is:

> In most situations, these two sources of information are naturally confounded and the experience of ease of recall goes along with a greater amount of recall.
>
> (ibid., p. 118)

Although they consider that the available evidence 'supports the original formulation of the availability heuristic'—namely that 'individuals estimate the frequency of an event, the likelihood of its occurrence, and its typicality "by the ease with which instances or associations come to mind"' (ibid., p. 118)—they do raise issues of complexity discussed earlier (see *Insensitivity to predictability*). They suggest that further research is needed on the 'interplay' (ibid., p. 119) between processing motivation and the ease or recall and accessibility of both experiential information and declarative information.

Imaginability

When people have to assess the frequency of some occurrence which is not stored in memory, their judgment is affected by the ease with which they can *imagine* the event. In an arresting example, Tversky and Kahneman suggest that dangers which might be associated with an expedition can appear very great if they are easily imagined even though the vividness of the imagined event is unrelated to the actual likelihood of the event occurring. Conversely, dangers can be grossly underestimated when they are difficult to imagine (Tversky & Kahneman 1982). A skiing enthusiast going from Australia to the French Alps could easily underestimate the likelihood of avalanches in the snow, for example. Dangerous and large-scale snow avalanches are virtually non-existent in Australia. The Australian skier with this background and who has not recently heard of any avalanches in the French Alps could easily be misled into believing high areas of powder snow to be much safer than they are.

Sherman et al. (2002) show this effect with imagining the likelihood of contracting a disease. When subjects actively imagined experiencing a disease with easily imagined symptoms, they judged themselves more likely to contract the disease. Conversely, if the symptoms are difficult to imagine, people can construe this as signifying a low probability of catching a particular disease or developing a particular syndrome. It is proposed here that it then a small step from estimating little likelihood of developing a condition (e.g. hypertension) to having a consolidated belief such as 'It won't happen to me, there's nothing wrong with my diet and exercise'. It is also probably safe to assume that such a development is all the more likely if the person has had no immediate and recent contact

with friends or relatives with the condition; hence, the availability—or, more precisely, the *lack* of availability—biases the judgment.

Illusory correlation

Experience of co-occurrence of events can lead to overestimating the frequency of such co-occurrence: for example (Tversky & Kahneman 1982), 'suspiciousness' and 'peculiar eyes'. They found that such illusory correlation, as they have named it, is so strong that it persists even in the face of contradictory data showing a negative correlation. Conversely, it can be so strong that people do not see a relationship that actually *does* exist.

When a strong associative bond has been established, people will conclude that the events are frequently paired. So, for example, in our society, particular groups of people—based on hat-wearing, ethnic origin, or type of vehicle—are judged to demonstrate such characteristics as a lesser ability for driving safely, or for being more aggressive.

While these findings continue to highlight the roles of perceptions in making judgments and forming beliefs, they also point increasingly to the role of two systems at work (Kahneman 2011; Sloman 2002) and, looking to the affect heuristic (Slovic *et al.* 2002), the proposed role of the feelings of certainty. The faster associative system has established itself through evolutionary advantage (Kahneman & Frederick 2002; Kahneman 2011); however, it is proposed here that a key problem with the speed of the heuristics involved is that it does not allow for consideration of the appropriateness of the heuristic to the context. Once a feeling of certainty or 'rightness' is established, we can reasonably expect that engaging the slower rule-based system is less likely to occur, especially if doing so arouses the discomfort that many feel when challenging existing belief systems that may be supporting the heuristics involved.

2.2.1.3 Adjustment and Anchoring

In this context, anchoring is the phenomenon of starting with an initial value which is then adjusted in order to reach an estimate. Different starting points yield estimates which are biased towards those starting points.

Tversky and Kahneman (1982) point out that anchoring affects subjective probability estimates in sophisticated subjects as well as naive ones, at least in part.

Insufficient adjustment

This phenomenon occurs both when the subject is given the starting point or when the starting point is the result of some incomplete calculation. An example of the second instance (ibid., p. 14) is the estimation of the product of a sequence of numbers. When the sequence is given in descending order ($8 \times 7 \times 6 \times 5 \times \ldots$), the estimated product is

higher than when the sequence is in ascending order (1 × 2 × 3 × 4 × ...). This occurs when subjects perform the first few steps of the calculation and extrapolate or adjust to get their final answer. Their predictions are affected by whether they are adjusting from a higher number (when using a descending sequence) or from a lower number (ascending sequence).

A context where this heuristic can be exploited is in sales. Suppose a real estate agent shows a prospective buyer a house, but before showing the house for sale mentions that similar houses in the street have recently sold for around $1.2 m. With that anchoring, it may now be much easier for the prospective buyer to accept that $980,000 is a good price, even it is not. Chapman and Johnson (2002) report research showing that both real estate agents and amateurs provided valuations of a house that were positively related to the supplied anchor, the listing price of the house.

Evaluation of conjuctive and disjunctive events

The findings here show that as a result of anchoring, people will often prefer to bet on less likely conjunctive events than more likely disjunctive events.

An example of a conjunctive event (Tversky & Kahneman 1982, p. 15) is the successful completion of a project comprising a series of events. While the probability of success of each event is quite high, the overall probability of success can be 'quite low' if the number of events in the series is large. Belief in the efficacy of complex project plans frequently leads to disappointment and confusion in the business world.

For a disjunctive event, an example is the likelihood of failure of a nuclear reactor. The probability of failure of any one of its many components is low, but the overall probability of failure is high because it can take only one component to fail for the whole complex system to fail.

Chapman and Johnson (2002) declare anchoring to be both prevalent and robust. In summarising others' research, they include the following in the judgment tasks as being influenced by anchoring: answering factual knowledge questions, estimating risks and uncertainty, making statistical inferences, evaluating monetary lotteries, judging self-efficacy, judging spousal preferences, and predicting future performance. They note the effect persists, even when subjects are instructed to ignore it. A simple but powerful example is the finding that evidence suggests that purchasers of canned tomatoes, for example, are more influenced to buy by a sign saying *Limit of 12 per customer* than by one saying *No limit per customer* (ibid.). As an example of the influence of perceived *scarcity* this would be no surprise to Cialdini (1984).

Because of its effects on judgment, Chapman and Johnson remark on the importance of finding the causes of anchoring and investigating what those causes can reveal about the efficacy of debiasing techniques. Their discussion is extensive and the discussion of both possible explanations of and applications of anchoring is seductive. Here, it is sufficient to note that they point to the involvement of other processes. For example:

> The evidence is mounting, however, that anchoring involves a constructive process of priming or memory retrieval that influences judgments of preference … and belief.
>
> (Chapman & Johnson 2002, p. 133)

2.2.2 Heuristics and Biases – Summary

Heuristics and biases were discussed following perception and perceptual mechanisms because of some of the more obvious links to the use of sense-data—or, more accurately, the perceptions derived from sense-data—as a source of data for justifying a belief or a claim to knowledge.

Some of the ways that beliefs can be formed are directly related to perceptual errors such as in the example given on page 66. If some landmark looks closer due to the lighting conditions, it is almost trivially obvious that a person could form the belief that they will be able to walk that distance in less time than is really needed. It also seems that this belief is dependent principally on the perception of distance and the judgment made as a result of that perception, and does not appear to rely on pre-existing beliefs, for example.

However, with much belief formation involving heuristics and biases it has been shown from the examples given that more is involved than a simple inference on the basis of a perceptual error alone. With representativeness, for example, a person may judge that Jim is a librarian on the basis of a description being representative of the stereotype of a librarian. In that, we can say that the heuristic has led to a specific belief about Jim's occupation. Other questions are immediately provoked, however:

- How did the person come to accept the stereotype in the first place? Was it through direct observation of a large sample of the population? Was it from another person offering a description of a 'typical' librarian?
- If the acceptance of the stereotype was based on another's description, how is that they accepted that description as being true?

The second point above invites enquiry of influences as apparently straightforward as liking someone. At a more general level, it is evident that the patterns of influence described by Cialdini (1984) can play a role; but, this is still more a description of *what* occurs and offers little insight into some of the more complex mechanisms underlying the judging processes and subsequent belief formation.

The work of Zajonc (1980) and Slovic et al. (2002) strongly suggests an important role for affect. Indeed, Kahneman and Frederick (2002) now argue a case for replacing the anchoring heuristic with the affect heuristic. Slovic et al. point to the value of Damasio's Somatic Marker Hypothesis which is examined in Section 4.1.2.1. Sloman's (2002) proposal of two systems for reasoning does two things: first, it invites further enquiry into work such as that of Eichenbaum and Bodkin (2000) in which they propose different systems for memory and for knowledge; and, second, it suggests investigation of that view may give insight into how people manage to delude themselves. As well, Chapman and Johnson (2002) suggest that the roles of priming and memory retrieval contribute.

In short, it is shown that the heuristics and biases discussed in the literature contribute to the formation of false beliefs and false claims to knowledge. The longer, and more intriguing, investigation involves examining how pre-existing beliefs, memory systems and affect systems contribute to the formation of false beliefs.

2.3 Other Factors Affecting Perceptions

This section briefly addresses effects on perceptions from influences such as emotion. The principal discussion of these effects occurs in the subsequent chapters. Since the focus of this chapter is on perceptions and how they contribute to belief formation, it is useful at least to note that apart from heuristics and biases there are other factors affecting perceptions which, in turn, can contribute to false beliefs and false claims to knowledge.

Emotion-induced Perceptual Filtering

Styles (2005) briefly reports on an experiment in which rich and poor children were asked 'to estimate the size of coins' and it was found that 'poor children reliably over-estimated the size of coins in comparison to the rich children' (ibid., p. 65). This, she says, indicates that perceptions are influenced by emotional factors. Bruner and Goodman (1947) did the work she quotes, and they noted at the time:

> Throughout the history of modern psychology, until very recent times, perception has been treated as though the perceiver were a passive recording instrument of rather complex design … Such psychology, practised as it were *in vitro*, has fallen short of clarifying the nature of *perception* in everyday life much as did the old nerve-muscle psychophysiology fall short of explaining *behavior* in everyday life.
>
> (ibid., p. 33)

Bruner and Goodman's early work signalled a more recent interest in this area. It seems intuitively obvious that emotions and mood can affect the sense-data that we will be responsive to. For example, if someone is depressed it seems likely that they either will see *less* of their surroundings because their gaze is often directed downwards, or will see

differently (perhaps the environment really does look 'dull' to them). Intuition does not a thesis prove, but recent work suggests that emotion not only affects attention but also affects *how* we see (Anderson & Phelps 2001; Phelps *et al.* 2006; Vuilleumier & Schwartz 2001; Öhman *et al.* 2001). Phelps et al. conclude that their results show:

> … that emotion facilitates early visual processing. The mere presence of a fearful face enhances contrast sensitivity (Experiment 1), and this effect is magnified with transient covert attention (Experiment 2).
>
> (Phelps *et al.* 2006, p. 7)

In their conclusion, Phelps et al. add:

> Given that the emotional salience of a stimulus is an indication of its value or importance, one would expect that emotion may influence even the most basic perceptual abilities. Here we have demonstrated for the first time that emotion facilitates early vision: People actually see better in the presence of emotional stimuli.
>
> (ibid., p. 8)

This investigation is taken up again in Chapter 3 which is devoted to the examination of the nature of emotion and its influence on beliefs.

Belief-induced Perceptual Filtering

Popular self-help literature frequently expounds the view that desired states such as 'success' are achievable if only we look at things the right way. But, looking at things the right way really means we have to believe in the possibility of achieving *personal* success to be able to perceive all the clues and opportunities that we are currently blind to. Is there any substance in this view?

Much of the literature on heuristics and biases reviewed so far implies that there may be. If a person believes that a run of Heads from a fair coin 'just has to' yield to a Tails throw on the next throw, the belief in the *specific* instance might well arise from the application of the representativeness heuristic discussed in the subsection *Misconceptions of chance*. As suggested earlier, it appears that the underlying assumption—or, indeed, belief—is along the lines that there cannot be an unbroken run of heads that goes for too long. Gilovich (1991) captures the popular view of the effect of belief on perception with the quote attributed to the psychologist Thane Pittman:

> I'll see it when I believe it.
>
> (ibid., p. 49)

However, based on the studies of heuristics and biases, the above quote has more substance than it superficially appears. Gilovich continues by pointing out that an

heuristic such as representativeness has advantages in allowing judgments to be made easily and successfully, but that the very same processes come with the price of the systematic errors discussed earlier.

> No feature of human judgment and reasoning illustrates this trade-off of advantage and disadvantage better than the tendency for our expectations, preconceptions, and prior beliefs to influence our interpretation of new information. When examining evidence relevant to a given belief, people are inclined to see what they expect to see, and conclude what they expect to conclude.
>
> (ibid., p. 50)

He then continues to offer numerous examples, many of which are similar to those already discussed. Some notable ones, for the purpose here, include: the perception that someone is 'bad' or 'evil' simply because they are wearing black and have a particular complexion—a device, Gilovich observes, frequently employed by film makers; and, the selective remembering of events that support belief in some superstition or conspiracy theory.

Macknik and Martinez-Conde (2010) investigate how people are deceived by magic illusion. Succinctly, they observe that the human brain:

> is constantly comparing incoming information to what it already knows, expects, or believes. Every experience is measured up against prior beliefs and a priori assumptions.
>
> (ibid., p. 142)

In the literature on confabulation Wheatley takes a similar view:

> We assume that what we see and feel is an objective read-out of the world, unvarnished by personal biases, contexts, and assumptions … Psychological evidence suggests that perception is a constructive process that relies upon assumptions. When those assumptions are inappropriate, errors occur. Thus, errors in perception lead to errors of explanation.
>
> (Wheatley 2009, p. 208)

Further examples are redundant at this stage as the complex interrelationship between beliefs, heuristics, memory, attention and affect is taken up in the following chapters.

The Effect of Language on Perception

Kahneman and Frederick (2002) note the role of language in evoking representations and at the same time allude to the complexity involved in judgment and decision making and, by inference, belief formation:

> The *perception* of a stranger as menacing is inseparable from a *prediction* of future harm. Intuitive thinking extends perception-like processing from current sensations

> to judgment objects that are not currently present, including mental representation that are evoked by language. However, the representations on which intuitive judgments operate retain some features of percepts: They are concrete and specific, and they carry causal propensities and an affective.
>
> (ibid., p. 30)

From a folk psychology perspective, it is relatively easy to propose that the choice of language can affect perception. For example, the wine connoisseur asks the naive wine taster *'Can you taste the delicate raspberry tones that linger at the back of the palate?'* The naive party, previously unaware of any such possibility ('I though wine was made from grapes, not raspberries!'), nods in perplexed agreement as they convince themselves that what they are tasting is delicate raspberry tones. Or, a friend, A, tells B that the latest action film is 'massively exciting'. In response, B imagines scenes that they perceive as exciting. This, of course, is the source of potential disappointment if the projected movie does not satisfactorily compare with the imagined one.

In the marketing world and political world, choice of words occupies almost all attention because of the supposed effects. Of course, it is difficult to separate any effect on perception from the appeal to values, for example, and the associated affective response.

While the literature appears to contain few specific references to 'language affecting perceptions', the significant body of work relating to framing and decision making—for example, Kahneman and Tversky (2000)—makes it clear that how the problem is framed has a clear effect on decision making. If Kahneman and Frederick (op. cit.) are accurate, describing the price of insurance as 'money thrown away' will evoke different representations from 'an investment in peace of mind'; different representations with connections to different emotions and, most probably, beliefs.

Barrett at al.'s (2007) review of literature related to the language-as-context hypothesis leads them to suggest that the perception of emotion in others does not occur independently of cognitive processes such as language. They observe, for example:

> The act of providing an emotional label to caricatured emotional faces (as opposed to a gender label) increases neural activity in right IFG and produces a corresponding decrease in amgydala response … This reduction in amgydala response can be thought of as reflecting a reduced ambiguity in the meaning of the structural information from the face.[32]
>
> ....
>
> In addition, emotion words cause a perceptual shift in the way that faces are seen. Morphed faces depicting an equal blend of happiness and anger are encoded as

32. IFG = inferior frontal gyrus

> angrier when those faces are paired with the word 'angry', and they are encoded as even angrier when participants are asked to explain why those faces are angry.
>
> (ibid., p. 329)

In building their case, Barrett et al. also reference research relating to the effect of language on time perception, spatial reasoning and colour perception. Boroditsky (2001) reports on research examining the difference in time perception between Mandarin and English speakers in which Mandarin speakers have metaphor of time running vertically while English speakers have time running horizontally. They acknowledge that 'the strong Whorfian view … has long been abandoned' but suggest that 'many weaker but still interesting formulations can be entertained' (ibid., p. 2). Boroditsky concludes that:

> (1) language is a powerful tool in shaping thought about abstract domains and (2) one's native language plays an important role in shaping habitual thought (e.g., how one tends to think about time) but does not entirely determine one's thinking in the strong Whorfian sense.
>
> (ibid., p. 1)

Along broader lines, Levinson et al. (2002) report experimental results which they claim support their position that 'the non-linguistic representation systems used in memory and inference are systematically influenced by the language spoken' (ibid., p. 186). While acknowledging vigorous debate regarding this conclusion, the reports of their own and others' research appear to strengthen their case as do observations such as:

> Finally, to understand how language could have an effect on cognition, no outlandish mechanisms need be supposed. To drive a car, you need to acquire new motoric and cognitive skills. To speak Tzeltal, you'll need to be able to do base-20 math in the head, since it has a vintegesimal number system, and more relevantly, you'll constantly need to maintain a mental compass, since 'downhill' denotes a quadrant based on c. N 345°, for without that notion you can't describe where anything is. Further, a central mechanism responsible for the cognitive efficacy of language is provided by one of the corner-stones of cognitive psychology, namely Miller's coding theory of short-term memory limitations (Cowan, 2001; Miller, 1956).
>
> (ibid., p. 185)

Gilbert et al. (2006) acknowledge the controversy associated with a Whorfian view. However, they also observe that the issue 'has been debated largely on the basis of cross-language data, without considering the functional organisation of the brain' (ibid., p. 489). Their findings indicate that a person's native language affects perception of what they see in the right half of their visual field. The mechanism is not clear: they suggest that 'language affects perception directly', or 'language affects postperceptual processes' (ibid., p. 493), or possibly that both mechanisms play a role. In short, they are clear that there is

an effect on perception from language. Drivonikou et al. (2007) report replication of Gilbert et al.'s findings but also that there are 'significant category effects' (ibid., p. 1097) in the left visual field. They, too, argue that 'considerations of brain organisation … may be relevant to' (ibid., p. 1100) the debate over the Whorfian hypothesis. Following this line of enquiry Tan et al. (2008) also suggest there is 'neurophysiological evidence' for the Whorfian hypothesis and they conclude that 'Language appears to affect neural activity patterns activated in the course of color perception' (ibid., p. 4007). Mo et al. (2011) also claim electroencephalographic evidence for this lateralised Whorfian effect.

This is not the forum for resolving the extensive and controversial debate over the Whorfian hypothesis. Here, the instances above serve to indicate that there is some concrete evidence supporting the intuition that language can affect perception. Therefore, since perception is used as a basis for justifying belief and claiming knowledge, it appears that one more agent that can contribute to false belief needs to be considered.

2.4 Summary

It was observed that perception appears to be a natural starting point for examining how people justify beliefs and claim knowledge. For the layperson in particular, attempts to settle disputes about what is a true or a more accurate account of events generally leads to relying on sense-data.

Then, it was proposed that variation in function between people's sensory receptors and apparatuses could lead to different perceptions and beliefs. This, it is suggested, is especially likely if there is damage such as high frequency hearing loss of which the person is *not* aware.

Recent research showing the extent of individual genetic variation in response to sense-data was discussed. Because people are unaware of many of these differences, examples were offered that supported the argument that these differing responses can lead to differing opinions of what was experienced—for example, the taste of a some particular food. In effect, the 'opinions' are either beliefs or can become consolidated as beliefs. There is some argument still about whether some of these beliefs can be considered false if they are based on an individual's experience that differs only because of genetic coding and not from some other factor such as inattentional blindness. However, it was argued that the development from these beliefs can be significant (for example, *'Jean has no taste; you should see the colours she chooses*') in the way in which they can affect relationships and future action.

Another frequent source of false perceptions is illusions. While the majority of the literature focuses on visual illusion, it was noted that illusions do occur in other sense

modalities. There is no doubt about the existence of illusions and that they affect perception. Attention was given in this section to extended examples that illustrate how an illusion can lead to a false belief. What emerged from these examples is the significant role played by the context and by attention.

There is a growing body of literature dealing with inattentional blindness and change blindness. What is clear is that people, generally, have a belief that they notice more than they really do or, in some cases, can. Examples of how inattentional blindness or change blindness can lead to false beliefs were offered.

Associated with basing beliefs on perception is a whole body of work on heuristics and biases. One of the key points related to perception (see page 64) is that different inferences are drawn from the same sense-data by different individuals. The heuristics and biases research both demonstrates and explains to some extent how this can occur.

Perhaps the most succinct summary of the nature of heuristics and biases findings is that noted earlier: the 'relative neglect of [relevant] other considerations' (Tversky & Kahneman 2002, p. 20) when making judgments and decisions. The section summarises sufficient research, with examples, to establish that the heuristics and biases to which humans are subject can easily lead to false beliefs and false claims to knowledge. So far, reference has been made to 'inferences' and to how heuristics and biases can 'lead to' false beliefs. However, it is proposed that many of the propositions that have previously been labelled a 'judgment' or a 'decision' *are beliefs*. If the person judges that they can reach a distant landmark in a particular time, based on their observations and subsequent perceptions, that proposition *is* a belief: they hold it to be true, it is related to their desire to reach the landmark, and it informs their actions. In this, then, heuristics and biases directly cause many beliefs.

The other highly significant consideration arising from the heuristics and biases project is the complexity of the interactions between perceptions, heuristics and biases, existing beliefs, attention, memory, and affect. An enquiry into this complexity was opened in the section *Other Factors Affecting Perceptions* in preparation for Chapters 3 and 4. Research indicates that perceptions can be influenced by emotions, beliefs and even language.

Having now established that perceptions can be used to justify false beliefs and claims to knowledge via a range of processes from simple illusions to complex heuristics that entail emotion and beliefs, the task now progresses to a closer enquiry into the roles of emotions in generating and maintaining false beliefs.

Chapter 3: Emotions, Feelings and Belief

While considerable debate continues over details of the mechanisms involved in the role of emotions in affecting beliefs and memory, there is general agreement that:

1 emotions exert effects on beliefs

2 emotions are influenced by beliefs

3 emotions affect memories

4 emotions are influenced by memories.

In an earlier submission (Thompson 2007), the challenges of researching *The Cognitive Science of Belief* were presented. These challenges included:

- a web-like cause-effect interrelation among emotions, feelings, memory, perception, attention, metacognition
- the lack of clear definitions of terms commonly used in the discussion. At the time it was noted that 'One of the striking features of much of this literature is the paucity of clear distinctions when terms such as *emotion*, *memory*, and even *cognition* are used' (ibid., p. 3).

The literature agrees either explicitly or by implication that there is a complex relationship between neurological processes, physiological processes, body states and mental states. Continued reading has only served to emphasise the weight of the role of emotions in this complex set of interactions. What has also emerged is:

- growing agreement that *feelings* are distinct from *emotions*
- the implication for emotion–cognition interaction and the growing debate about the degree to which conscious awareness is necessary for feeling.

3.1 Emotion and Belief

Frijda et al. (2000b) regard beliefs as 'one of the major determinants of emotion' (ibid., p. 1) and consequently consider that emotion falls properly into cognitive psychology research. However, they note that the reverse influence—that of emotion on cognition—has received little attention. That emotions have a role in beliefs is suggested by the resistance of beliefs to change:

> That the role of emotions in judgments and beliefs is nevertheless plausible is brought home by considering a few pertinent facts. One of these is the sheer resistance of beliefs to being modified by information.
>
> (ibid., pp. 2–3)

It is this resistance of beliefs to modification that is a key concern of this thesis. Frijda et al. (2000a) offer contributions from a number of researchers who not only affirm that emotion affects belief but also are beginning the inquiry into how that may be so (Clark & Brissette 2000; Clore & Gasper 2000; Eysenck 2000; Fiedler & Bless 2000; Frijda *et al.* 2000a; Frijda & Mesquita 2000; Harmon-Jones 2000; Oatley 2000).

A sampling of the views follows.

> These effects [of emotions on beliefs] are in fact central to the place of both emotions and beliefs in human functioning. It can be argued that they are in no way restricted to belief distortions. They are at the heart of what beliefs are about.
>
> (Frijda *et al.* 2000b, p. 3)

> ... affect may influence beliefs because it provides experiential information or feedback about one's appraisal of objects to which the feelings appear to be a reaction.
>
> (Clore & Gasper 2000, p. 12)

> Emotions influence beliefs in basically two ways. They may give rise to beliefs where none existed, or change existing beliefs; and they may enhance or decrease the strength with which a belief is held.
>
> (Frijda & Mesquita 2000, p. 45)

> ... beliefs are highly personal, subjective and idiosyncratic creations ... It is hardly surprising then that fluctuations in daily mood can have such a delicate, subtle and ongoing influence on the kind of beliefs people formulate, access and use in charting their daily behaviors.
>
> (Forgas 2000b, p. 139)

> Emotion is a particularly interesting factor that can affect the extent to which other features are bound together during encoding or during subsequent thinking or talking about events.
>
> (Johnson & Raye 2000, p. 44)

Frijda and Mesquita (2000, p. 52) see the connection between emotion and belief as being more than an interrelatedness:

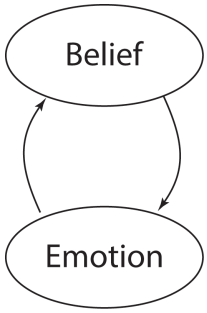
> Beliefs are part of emotions. They are part of emotional experiences, as they constitute the meaning attached to events. The beliefs that are part of emotions, usually called cognitive appraisal (e.g., Lazarus, 1991), stem from the relevance of the eliciting event to one or several of the individual's concerns.[33]

Beyond the collected papers in Frijda et al. (2000a), Bennett and Hacker (2003, p. 218) acknowledge that emotions 'are linked in complex ways to what the agent knows or

33. Citing Lazarus, RS 1991, *Emotion and adaption*, Oxford University Press, New York.

believes'. They suggest that the emotion must have a proper object and offer examples such as 'If [the agent] fears A or A's action, he must know or believe that A or A's action is a threat. If an agent feels pity or compassion for another, he must know or believe that person to have suffered a misfortune'. Although they seem to be more focused on the effect of belief on emotion (rather than the reverse), they do point to a direct interrelationship between emotion and belief (as suggested in **Figure 1: Belief-emotion feedback loop**), as do the other researchers cited above.

Figure 1: Belief-emotion feedback loop



It has been suggested that understanding this interrelationship is important from a psychology perspective as it helps in understanding action (Frijda *et al.* 2000b, p. 4). More particularly, it is proposed here that such an understanding will be helpful in understanding the intractability of many false beliefs. In the field of cognitive behaviour therapy, James and Barton (2004) provide a glimpse of potential application for a better understanding of the relationship between emotion and belief when they stress the importance of attending to 'high levels of the emotional content of the material' (ibid., pp. 441–442) when working with self-referent core beliefs.

Here, we recall that the focus of this thesis is gaining insight into *how* people can form and maintain beliefs that are demonstrably in error or can be shown to be only as plausible as other equally plausible beliefs. The characteristics of the beliefs in focus have been outlined above (see page 9) and in an earlier submission (Thompson 2007, p. 4). In seeking to understand *how* such beliefs can be sustained, Frijda and Mesquita's (2000) paper provides claims about beliefs being 'part of emotions' which, in turn, indicate useful directions for enquiry. Similarly, other chapters in Frijda et al. (2000a) affirm that emotions and feelings have effects on beliefs and they, too, provide direction for more detailed investigation below. However, in order to effectively gain insight into how beliefs and emotions may be so intricately related, it is necessary to examine, first, the nature of emotion.

3.2 The Nature of Emotion

It has already been shown (Chapter 1) that there is extensive debate about what is meant by the term 'belief'. However, there is also extensive debate about the nature of *emotion*. Any investigation of the effect of emotion on belief cannot escape addressing the debate over the nature of and definitions for emotion. Solomon (2008) notes that although the question *What is an emotion?* was asked by William James, the question has been

concerning philosophers since before Socrates. Despite the long history of philosophical interest, as well as continuing interest, Solomon notes:

> … it is now clear that philosophers cannot ignore or neglect the rich neurophysiological literature on emotions … [They] may continue to argue that Aristotle knew all about emotions even though he did not know anything about the brain, but they do so at their peril—and in the face of the obvious fact that among the factors that have altered the history of philosophy and its concepts most radically have been new advances in previously unknown or undeveloped sciences.
>
> (ibid., p. 11)

Similarly, LeDoux (2000) argues for the value of using 'psychologically well-defined' (ibid., p. 177) aspects of emotion in order to devise experiments that avoid the 'vague and poorly defined aspects of emotion' (ibid., p. 177). He complains that:

> Neuroscience meetings these days have numerous papers on the role of the brain in emotion, affect, hedonic tone, and the like. Unless these vague concepts can be operationalised, as was done in the work on fear, they are likely to impede, if not recede, the progress.
>
> (ibid., p. 177)

3.2.1 A Preliminary View of Emotion

Frijda et al. (2000b, p. 4) treat both emotions and beliefs as mental states and note that, therefore, defining the terms *emotion* and *belief* is 'not unproblematic'. They attempt to distinguish the two as follows:

> Emotions can be defined as states that comprise feelings, physiological changes, expressive behavior, and inclinations to to act. Beliefs can be defined as states that link a person or group or object or concept with one or more attributes, and this is held by the believer to be true. The general proposal thus is that emotions can awaken, intrude into, and shape beliefs, by creating them, by amplifying or altering them, and by making them resistant to change.
>
> (ibid., p. 5)

In referring to the definition for beliefs proposed on page 22, the proposition that beliefs are related to desires and wants, as well as being a preparation for action, fits with the suggestion from Frijda et al. (ibid.).

While potentially being useful in pointing to the relationship between emotion and belief, Frijda et al.'s definition of emotion provides an example of the unfortunate conflation of meanings of terms such as *feelings* with *emotions*. Forgas (2000b, p. 110) comments that 'there is still little general agreement about how best to *define* terms such as affect, feelings, emotions or mood'. He applies *affect* as a generic label for both moods and emotions, then

distinguishes emotions from moods by being 'more intense, short-lived' than moods and usually having 'a definite cause and clear cognitive content (e.g. anger or fear)'. Ortony et al. (1988, p. 1) propose that:

> Emotions have many facets. They involve feelings and experience, they involve physiology and behavior, and they involve cognitions and conceptualizations.

This is highly suggestive of the complexity of the research being undertaken, but is not helpful either as a definition or as an explication of the nature of emotions. Solomon (2004, p. 13) defines emotions in terms of five 'aspects' which are common to 'every emotion':

- behavioural (includes facial expressions, verbal expressions, report, plans for action)
- physiological (hormonal, neurological, neuromuscular)
- phenomenological ('everything from "physical" sensations to ways of seeing and describing the "objects" of one's emotions and "meta-emotions"')
- cognitive (including appraisals, perceptions, thoughts, and reflections about one's emotions)
- social content (from the immediacy of interpersonal interactions to pervasive cultural considerations).

He notes that these aspects are 'interwoven' and that they are not 'competing conceptions of emotions' (ibid., p. 13). In a somewhat similar vein, Elster (2004, p. 38) defines emotions in terms of *six* features:

- physiological arousal – any departure from the physiological baseline
- physiological expressions – expressions perceived as forerunners to action
- valence – the pain and pleasure accompanying an emotion
- cognitive antecedent – for Elster these are beliefs (rational or irrational) that trigger emotional reactions
- intentional object – the object that the emotion is directed towards or is about
- action tendency – emotions instigate action (e.g. ignoring a person for whom contempt is felt).

The above two definitions are offered to indicate some of the more specific aspects that are frequently agreed on as contributing to emotion or as having some direct interrelationship with the phenomenon.

Elster (ibid.) discusses rational choice and emotions, so his project is different from that of this thesis. However, the relevance of his work lies, initially, in two aspects:

1 He presents a 'rational-choice explanation' of behaviour that is explicit about the effect of 'desire' on belief and, *by implication*, of emotion on belief.

2 His summary of the 'modern theory of rational choice' (p. 36) highlights both the terminological problems involved as well as the problems with sorting out the interrelationships.

The summary he offers (ibid., p. 36) is:

Figure 2: Rational-choice Explanation of Behaviour

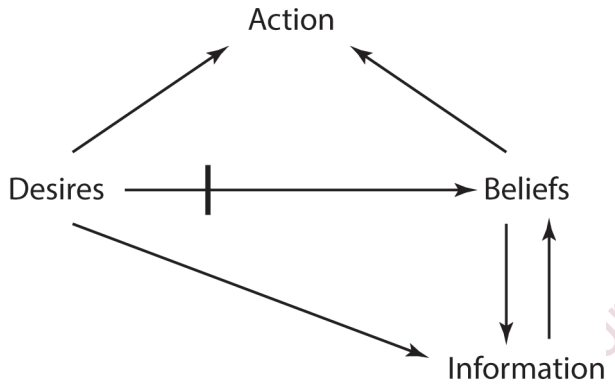


Figure 3.1 Rational-choice explanation of behavior

Here, he shows interactions between the depicted aspects of the model. He offers no discernibly clear definition of 'desire', although he later draws a distinction between *desires* and *wishes*. The principal example he gives in support of this distinction is that 'One may wish that a certain state obtain, or one may desire to bring it about' (ibid., p. 40); but it is not at all clear what the distinction is. This is not the place to fully critique either the model that Elster offers or the conclusions he draws with respect to rational choice. What is of interest is his way with terms such as *desires*, *emotions* and *beliefs*.

Elster's reader is left to infer the specifics of the relationship between emotion and 'desire' in the model. The use of 'desire' in the model presumably arises from the view that 'emotions have intentional objects'; so, presumably one has an object of desire. But, given that he then distinguishes between desires and wishes, we have to wonder whether the model might be better served by explicitly including 'emotion'.

The emotions that Elster considers, he claims, are 'indeed triggered by beliefs' (ibid., p. 39). However, this relationship is not apparent in the model. Surely, for example, a belief can instigate a 'desire': so why is there no arrow going directly from 'belief' to 'desire' or, alternatively, via 'action' to 'desire'? (The assumption is being made here that there is no misprint in the arrow directions in the diagram.)

Elster is not alone in offering arguments where there is lack of attention to definitional precision and what we might call 'high level' models of emotional and cognitive interrelationships.[34] For the moment, it suffices to consider a 'high level' model as one such as Elster's which gives little explanation of the mechanisms underlying such a model. That is not to say that such models are not useful as a starting point. They are not sufficient as an endpoint, however, for the reason that a folk psychologist could look at Elster's model and say *'So, beliefs instigate actions. Tell me something I didn't know!'*

Therefore, in order to get closer to understanding mechanisms that allow people to maintain false beliefs, it is necessary to get something richer than these high level models. Much of the literature in Manstead et al. (2004), Frijda et al. (2000a), Uttl et al. (2006a) and Barrett et al. (2005a) serves as worthy summaries of large bodies of research and provides a good base from which to start. However, to get the best from that material it is worth starting with the bold, and still somewhat controversial, work of Antonio Damasio (1994, 1998, 1999, 2000b, 2000a, 2000b, 2003b, 2004; Damasio *et al.* 2000).

3.2.2 Why Damasio's Hypothesis?

Two terms that appear regularly in the literature relating to emotions and cognition are 'emotion' and 'feeling'. Frequently, they are conflated. However, an increasing number of researchers (some of whom are discussed later) are making a distinction between *emotion* and *feeling*. This distinction appears important in elements of the debate relating to:

- conscious and unconscious experience of emotion
- the role of awareness in the experience of emotion, and
- (ultimately) the connection with cognitive processes that create, maintain, alter and delete beliefs.

Clore and Gasper (2000, p. 10) hint at the potential importance of 'feelings' in guiding cognition processes:

> Two important aspects of emotional feelings are that they provide information and guide attention.

Here the term 'emotional feelings' belongs to the set of conflated terminology common to this area of discussion. Winston and Dolan (2004, p. 204) highlight the problem of the imprecise use of words and the value of separating them:

> ... for the best part of several decades in the middle of the twentieth century, neuroscientists treated 'emotion' and 'feeling' as interchangeable terms. This approach

34. More on definitions in section 3.2.2

> to the language of emotion, coupled with the behaviorist movement ... meant that the neuroscientific study of emotion was grossly neglected, often seen as conceptually ragged and its practitioners as pursuing 'soft science' ... Among the reasons that emotion has assumed center stage in neuroscience is a rediscovery of the importance of a dissociation between emotion and feeling ... which potentially offers a more tractable approach to the neuroscientific study of emotion.

Damasio has an interest that lies at the core of this thesis:

> … I must confess that I worry about the dividing line between the beliefs of sane individuals and the irrational beliefs of neurological and psychiatric patients. I suspect it is not as wide as one might wish. Many beliefs are by nature 'arational', and many are downright irrational.
>
> (Damasio 2000b, pp. 328–329)

Damasio's hope is that the 'understanding of the cognitive neuroscience of belief will make considerable progress as we advance our understanding of the underpinnings of memory, of emotion and feeling, and of the biological nature of the self' (ibid., p. 332). In pursuit of this hope, Damasio and his colleagues hypothesise a view of emotions and feelings that provides a strong starting point for proposing how emotions may be so intricately interrelated with beliefs.

Damasio's definitions of emotions and feelings are, at first, counter-intuitive. His requirement that *emotion is a precursor to feeling* seems singularly foreign given the frequent conflating of the two terms in vernacular conversation. However, given that most of this thesis is examining the relationships between emotion, feeling, memory and belief, it would be helpful if such concepts had clearly distinguishing attributes. Hence, Damasio's hypothesis on emotions and feelings (1994, 1999, 2000b, 2000a, 2003b, 2004) is put forward as a valuable starting point for clarifying the interrelationships between emotion, feeling, memory and belief.

Damasio acknowledges that he is using 'working hypotheses' (2004, p. 50)—yet his view on the difference between emotions and feelings provides a very plausible position for interpreting the neuroscience findings. Winston and Dolan (2004, p. 205) attest to the value of Damasio's work in reviewing neuroimaging work:

> Though not universally accepted, this distinction between automatically elicited emotional states and experiential (and necessarily subsequent) feeling states is a useful heuristic in reviewing work in the relatively new field of functional neuroimaging of emotion.

Further, Damasio's view is useful in sorting how, for example, the processes involved in emotion and feeling relate to belief systems (discussed in Chapter 4). Daniel Dennett's

review (1995) of *Descartes' Error* (Damasio 1994) mainly celebrates Damasio's assault on the separation of mind and body, but his comments with respect to the neuroscience are pertinent:

> Damasio's vision is in itself not entirely novel—its major elements can be discerned in Aristotle, in Nietzsche, and most recently in Nicholas Humphrey's *A History of the Mind*—but under Damasio's boldly constructed umbrella of neuroanatomical details, these elements join to become not just compelling, but retrospectively obvious, provoking the theorist's ultimate accolade: 'Now why didn't I think of that?'
>
> (Dennett 1995, p. 1)

Dennett (ibid., p. 4) continues by putting a strong view for the value of such a model as Damasio's in processes of advancing our understanding.

> This is a time for bold attempts, not worrying overmuch about proving each conjecture along the way, but rather getting enough of a story clearly out on the table to suggest tests, and the inevitable further refinements. It is time for tentative model-building, in other words, without undue anxiety about overstating the case at the outset. There are hundreds (or thousands) of bits lying around that need to be incorporated into a 'working model' of a whole human agent, and just finding one plausible (and ultimately testable) model of a whole agent is an extremely demanding task. It requires a mixture of global vision and detailed knowledge, and lots of imagination. Damasio has bravely provided a pretty good—maybe a very good—model ...

3.2.3 Separating Emotions and Feelings

In putting forward separate definitions of *emotion* and *feeling*, Damasio acknowledges arguments that would treat them as synonymous, or would avoid the term *feeling* on the grounds of its being too vague. He suggests that from a *neurological* perspective:

> … emotions and feeling are not the same thing as objects of study, notwithstanding the fact that, in real time and real life, emotion and feeling seemingly occur as if they were merely two sides of the same coin, affect pure and simple.
>
> (2004, p. 49)

Damasio's suggestion is that the two terms 'allow for a clearer communication regarding the profile' (ibid., p. 50) of physiological and mental events in affective processes and of related theories and hypotheses. His simple (more recent) definition of emotions is the following:

> ... emotions are *bioregulatory reactions* that aim at promoting, directly or indirectly, the sort of physiological states that secure not just survival but survival regulated into

> the range we, conscious and thinking creatures, identify with *well-being*.
>
> (ibid., p.50)

The reactions are 'a patterned collection of chemical and neural responses' (ibid., p. 50) which is produced when the brain detects 'an emotionally competent stimulus' (ECS) (ibid., p. 50). The ECS is an 'object or situation actually perceived or recalled from memory' (ibid., p. 50). The inclusion of 'well-being' is especially pertinent as it points to the overall survival of the 'self' and its domains of interest that emerge in the discussion in Chapter 4.

As noted above, the aspect of Damasio's hypothesis that is most germane to this thesis is the separation of *feelings* from *emotions*. Damasio's working definition of feelings is:

> ... *the mental representation of the physiologic changes that occur during an emotion*. The essence of feelings of emotion is the mapping of the emotional state in the appropriate body-sensing regions of the brain. But *feeling an emotion also includes the mapping of changes that occur in the cognitive processing style, as well as the evocation of thoughts that are congruent with the feeling state.*
>
> (ibid., p. 52)

He acknowledges that the definitions offered are provisional, that they are different from those he has used when starting his work, and that because they are *working definitions* they are likely to be different in the future (ibid., p. 50). Much fuller definitions are offered in *Looking for Spinoza* (Damasio 2003b) along with a lengthy account of what is meant by *emotion* and *feeling*. The account therein summarises and adds to his hypothesis which has clearly been in development since the publication of *Descartes' Error* (Damasio 1994).

3.2.3.1 Emotions in Detail

Damasio relates emotions and feelings to homeostatic mechanisms, and uses a tree structure as an analogy for the various 'levels' of homeostatic responses he proposes (2003b, pp. 30–37). At the base, Damasio places the most fundamental homeostatic mechanisms and survival responses: metabolic regulation, basic reflexes and immune responses. Above that, he places behavioural responses to pain and pleasure. For example, for 'pain behaviours' his suggestions include removing the whole body from the source of pain, protecting wounded parts of the body, expressions of alarm and suffering, and internal responses such as increases in the number of some white blood cell types. 'Pleasure behaviours' include those which facilitate the approach to others, express 'confidence and well-being', and promote the secretion of such compounds as endorphins. Damasio's key point here appears to be that the *experience* of pain or pleasure is not the

same as the cause of the pain or pleasure. From this, he points out that the behaviours associated with the experience are *not actually necessary*. (Although, we could say, for example, that removing oneself from the jaws of a pain-inflicting sabre-toothed tiger would be extremely useful!)

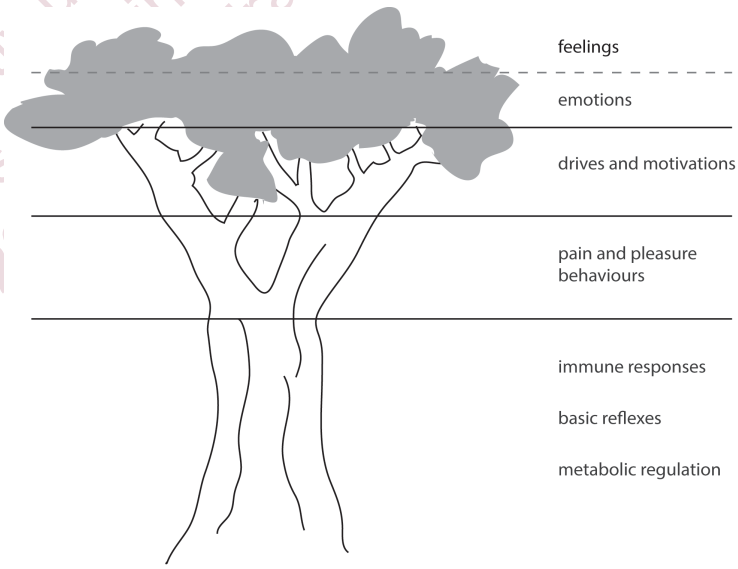
On his next level, Damasio places 'drives and motivations' which include hunger, thirst, curiosity, exploration, play and sex. The key distinction here is that the organism has 'conscious feelings' of the drives and motivations and of taking action to satisfy them or not. At the top, Damasio places 'emotions proper'—for example, joy, fear, sorrow, pride, sympathy—and feelings (at the very top).

Damasio summarises the role of all these processes as follows:

> The entire collection of homeostatic processes governs life moment by moment in every cell of our bodies ... First, something changes in the environment of an individual organism, internally or externally. Second, the changes have the potential to alter the course of the life of the organism (they can constitute a threat to its integrity, or an opportunity for its improvement). Third, the organism detects the change and acts accordingly, in a manner designed to create the most beneficial situation for its own self-preservation and efficient functioning. All reactions operate under this arrangement and are thus a means to *appraise* the internal and external circumstances of an organism and act accordingly. They detect trouble or detect opportunity and solve, by means of action, the problem of getting rid of the trouble or reaching out for the opportunity.

(ibid., pp. 35–36)

Figure 3: Damasio's view of emotions



Feelings support yet another level of homeostatic regulation. Feelings are a mental expression of all other levels of homeostatic regulation.

(After Damasio 2003, Figure 2.2, p. 37)

Damasio (ibid., pp. 43–46) offers three categories of emotions: background, primary, and social. **Background emotions** appear to be the consequence of combinations of the simpler homeostatic mechanisms: he essentially includes the first two levels of his tree analogy in these mechanisms. They are equated, rather loosely, with an individual's 'state of being'.

For the **primary emotions**, he uses those identified by Ekman (1992) as being those which are readily identifiable across cultures: fear, anger, disgust, surprise, sadness, and happiness. (Damasio credits the study of these as giving rise to much of what we know about the neurobiology of emotions.)

His **social emotions** include sympathy, embarrassment, shame, guilt, pride, jealousy, envy, and more. How the brain produces these emotions is barely understood yet. Damasio draws on observations of social interaction such as those which appear related to the learning of these emotions: for example, that a young monkey's fear of snakes does not develop until it witnesses its mother's fear of snakes (2003b, p. 47). He also speculates on the role of these social emotions in developing more complex cultural interactions.

In accounting for the different 'levels' of emotion he proposes, Damasio offers a nesting principle (ibid., pp. 37–38) in which aspects of the simpler regulatory mechanisms are preserved in the more complex mechanisms. In part it says:

> It consists of having parts of simpler reactions incorporated as components of more elaborate ones, a nesting of the simple within the complex. *Some* of the machinery of the immune system and of metabolic regulation is incorporated in the machinery of pain and pleasure behaviours. *Some* of the latter is incorporated in the machinery of drives and motivations (most of which revolve around metabolic corrections and all of which involve pain or pleasure). *Some* of the machinery from all the prior levels—reflexes, immune responses, metabolic balancing, pain or pleasure behaviours, drives—is incorporated in the machinery of the emotions-proper.
>
> (ibid., p. 37)

Damasio puts a case for the nesting principle applying to social emotions (ibid., pp. 43–46) in which he accounts for social emotions in the distinctions he draws between them, 'emotions-proper', and background emotions (ibid., pp. 43–46). Damasio's full definition of 'emotion-proper' is:

1. An emotion-proper, such as happiness, sadness, embarrassment, or sympathy, is a complex collection of chemical and neural responses forming a distinctive pattern.
2. The responses are produced by the normal brain when it detects an emotionally competent stimulus (an ECS), the object or event whose presence, actual or in mental recall, triggers the emotion. The responses are automatic.

3 The brain is prepared by evolution to respond to certain ECSs with specific repertoires of actions. However, the list of ECSs is not confined to those prescribed by evolution. It includes many others learned in a lifetime of experience.

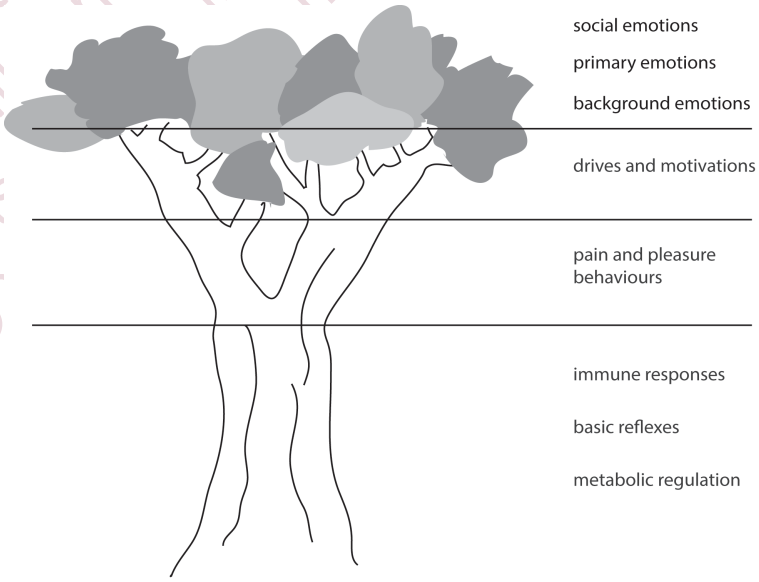
4 The immediate result of these responses is a temporary change in the state of the body proper, and in the state of the brain structures that map the body and support thinking.

5 The ultimate result of the responses, directly or indirectly, is the placement of the organism in circumstances conducive to survival and well-being.

(ibid. , p. 53)

To defend Damasio from a possible charge of begging the question here, it must be remembered that his view of emotions comes from an understanding of the neurological and physiological changes occurring in response to an emotionally competent stimulus. Based on empirical work, Damasio's nesting principle is coherent with his neurophysiological view of the basis of emotions. Damasio attempts to represent the nesting in **Figure 4: Kinds of Emotion** (below). This is not necessarily the best way to make his point—Venn diagrams may be more apt, for example. Nor does it reflect the dynamics of the interactions that are suggested by Damasio's definition. However, it does suggest a form of inclusion of some of each of the levels below any given 'level' of emotion as defined by Damasio.

Figure 4: Kinds of Emotion



There are at least three kinds of emotion-proper: background emotions, primary emotions, and social emotions. The nesting principle applies here, too. For example, social emotions incorporate responses that are part of primary and background emotions.

(After ibid., Figure 2.3 p. 45)

Solomon (2004, p. 13) suggests that 'no single claim or analysis will suit all emotions' because of the vast range of emotions. Damasio's suggestion of 'nesting' would help account for the features that both Solomon (ibid.) and Elster (2004) consider to be part of emotions. Some may apply, or be evident, to different degrees depending on the emotion being experienced. Similarly, there may be difficulty with particular features at certain times: for example, Elster's requirement that emotions have a cognitive antecedent seems harder to apply in the case of Damasio's background emotions unless we allow for unconsciously held beliefs. However, as a general observation, Damasio's model appears to explain the presence of such features as those proposed by Elster (ibid.) and by Solomon (2004).

In commencing his discussion on whether we have control of our emotions, Solomon (ibid., p. 12) says that 'emotions *seem* to happen to us, quite apart from our preference or intention'. If emotion is the result of homeostatic rebalancing in response to the environment—something which it is useful to *not* have conscious control of, for the most part—then this is no surprise. When the emotions are what Damasio calls emotions-proper—disgust, fear, happiness, sadness, sympathy, shame (Damasio 2003b, p. 39)—there is still engagement of the limbic system for example, giving rise to the emotion$_D$ before conscious thought can override. Indeed, thought—as in cognitive process—can not regard the emotion$_D$ occurring unless the emotion$_D$ is already present.[35]

Intentionally or not, Solomon argues for Damasio's view of feeling and also, therefore, for thought following emotion when he claims that thoughts are 'telltale symptoms of emotion' (Solomon 2004, p. 21). When he goes on to argue that 'an emotion *is* its expression', Solomon is arguing for our being responsible for emotion '*as* its expression' and calls on us to 'examine the reasons and motives' for how we behave (ibid., p. 28). Indeed, Damasio's invoking of conscious awareness of body state and of *mind state* when feeling$_D$ (1994, 1999, 2003b) provides a platform for discussing the conscious modification of the *expression* of an emotion.[36]

Solomon's paper is just one proceeding from the Amsterdam Symposium on feelings and emotions (Manstead *et al.* 2004). Some common themes arising in the papers, and to which Damasio's hypothesis can provide a perspective are discussed below (See section 3.2.3.3).

35. The term 'emotion$_D$' is used to refer to *emotion* as defined by Damasio.

36. The term 'feeling$_D$' is used to refer to *feeling* as defined by Damasio.

3.2.3.2 Feelings Defined

One of the challenges in working with Damasio's hypotheses is his rather extended development of his explanation of it. In *The Feeling of What Happens* (1999) and, later, in *Looking for Spinoza* (2003b), he continues to develop and expound his thoughts on feelings. The most succinct definition found is the following, which Damasio offers as a provisional hypothesis.

> ... *a feeling is the perception of a certain state of the body along with the perception of a certain mode of thinking and of thoughts with certain themes.* Feelings emerge when the sheer accumulation of mapped details reaches a certain stage.
>
> (ibid., p. 86)

To simplify what Damasio says, he points out that the brain has information arriving in it about all parts of the body and about all the systems in the body. From that, a 'map' of the body can be formed: that is, what is happening where and, importantly for his argument, if there is a disturbance from the normal (or desired) homeostatic conditions. Hence, emotions are a disturbance in homeostatic conditions. Once we are *aware* of the body map, once we perceive it, then, he says, we are feeling$_D$.

He suggests (ibid., p. 92) that '*all* feelings are feelings of some of the basic regulatory reactions ... or of appetites, or of emotions-proper, from straight pain to beatitude'. Even when the word 'feel' is used imprecisely in daily speech, Damasio suggests that its use still indicates an emotion:

> ... as in 'I feel I am right about this' or 'I feel I cannot agree with you'—we are referring, at least vaguely, to the feeling that accompanies the idea of believing a certain fact or endorsing a certain view. This is because believing and endorsing *cause* a certain emotion to happen.
>
> (ibid., p. 93)

It is in this quote that the relevance of Damasio's work to the question of belief becomes more obvious. When someone is saying *'I feel I am right about this'* it is reasonable to propose that, in many cases if not the majority, the person is expressing a belief that they are holding a true and verifiable view about some issue. Accepting this type of statement as a statement of believing was the subject matter of Section 1.1 and fits with the working definition established on page 27. The problem for the person is that they are taking their *feelings* about a situation as *evidence* for a belief. Assuming Damasio's view of emotions and feelings withstands scrutiny, it becomes a most useful insight into a mechanism for the formation and maintenance of at least one form of irrational beliefs—those formed on the basis of a strong emotional reaction and in which the associated feelings are taken as evidence for the belief.

3.2.3.3 Where Damasio's Hypothesis Can Pertain

Elster, in defending 'the traditional view that emotions interfere with and subvert instrumental rationality' (2004, p. 46), does not cite Damasio's work. Despite that, there are several aspects of his discussion to which Damasio's hypothesis could provide insight. As discussed earlier, Elster claims six features of occurrent emotions, drawing on Frijda's work (ibid., p. 38): physiological arousal, physiological expressions, valence, cognitive antecedent, intentional object, action tendency.[37] Generally, these fit comfortably with Damasio's hypothesis; however, one clear difference is the cognitive antecedent. While emotion_D can arise from 'cognitive' activity, such activity is not a necessary precursor in Damasio's hypothesis. Moreover, Elster is considering emotions that are 'triggered by beliefs' (ibid., p. 39) and is working from a basis that 'emotions have intentional objects' (ibid., p. 40). It is natural with this view that he will seek a cognitive antecedent. However that does not take into account the work of Damasio (already cited) and LeDoux (1998, 2000) and the growing support from others (Frijda *et al.* 2004; Frijda 2004; Panksepp 2004; Scherer 2004; Winston & Dolan 2004; Öhman & Wiens 2004). It is noted that Damasio claims that the nonconscious and automatic triggering of emotions does not preclude an 'appraisal' phase preceding emotions.[38]

> The process by which, at a given moment, an object or situation *becomes* an emotionally competent stimulus often includes a conscious, cognitive appraising of the circumstances.
>
> (Damasio 2004, p. 51)

Öhman and Wiens (2004) recognise and acknowledge the importance of automatic processes:

> … recent developments in physiological, cognitive, and social psychology … document the importance of automatic processes in many psychological contexts.
>
> (ibid., p. 58)

They also recognise three types of fears: social (e.g. threatening faces), animal (e.g. predatory or dangerous species), and nature (e.g. lightning, thunder). While these do not completely correlate with Damasio's (2003) model (see **Figure 4: Kinds of Emotion**), the distinctions open the way to discussion of *learned* fears and how beliefs can instigate emotional experiences.

The direct or indirect reference to appraisal mechanisms recalls the vigorous debate that occurred between Zajonc (1980, 1984) and Lazarus (1982, 1984). Lazarus (1984, p. 124),

37. He cites Frijda, N 1986, *The Emotions*, Cambridge University Press, Cambridge.

38. Additional development of this point follows later.

for example, insists that 'cognitive activity is a necessary precondition of emotion because to experience an emotion, people must comprehend … that their well-being is implicated in a transaction, for better or worse'. Zajonc (1984) counters that Lazarus's definition of emotion as *requiring* cognition establishes a circular argument and that Lazarus 'has broadened the definition of cognitive appraisal to include even the most primitive forms of sensory excitation, thus obliterating all distinction between cognition, sensation, and perception' (p. 117). Later, Leventhal and Scherer attributed the 'confrontational' (1987, p. 4) nature of this debate as an unproductive debate over definitions. They suggest that:

> … these seemingly contradictory positions can only be reconciled, and the focus moved from debate to theoretically guided research, by first distinguishing between the mechanisms and processes underlying emotion and cognition and the various covert and overt reactions (or events) used to record emotional and cognitive behaviour.
>
> (ibid., p. 4)

In examining the definitions of *emotion* and *feeling*, Leventhal and Scherer propose a componential model (see page 109) which they suggest 'makes clear that a number of innate, sensory motor mechanisms are likely to be responsible for the very earliest behaviours that are labelled emotional' (ibid., pp. 11–12). They see emotion and cognition as *separate processes* but suggest they become 'inseparably interrelated' (ibid., p. 3) as 'complex cognitive-emotional patterns … result from the participation of at least two distinct levels of memory and information processing, a schematic and a conceptual level' (ibid., p. 3).

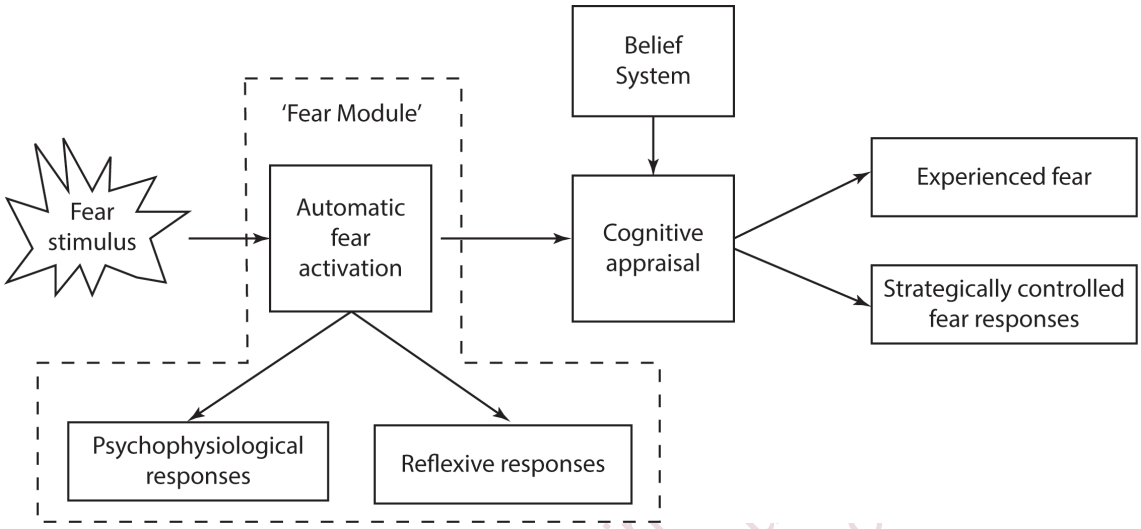
LeDoux (1993) also engaged in similar debate with Parrot and Schulkin (1993). He adds to Leventhal and Scherer's (1987) voice in highlighting the issue of semantics in the debate. In a similar view to Zajonc (1980, 1984), LeDoux suggests that 'If cognition is defined broadly as information processing, then emotion must be dependent upon cognition' (1993). The more pertinent observation that LeDoux makes here is that:

> We are now capable of and have begun discussing the problem in terms of what is processed where and how the different processing modules are related in terms of anatomical processing sequences.
>
> (ibid., p. 63)

Earlier, LeDoux's (2000) comments on the need for addressing vague definitions were noted (see page 94). His call for using an experimental approach with 'psychologically well-defined aspect[s] of emotion' (p. 177) echoes Zajonc's (1980, 1984) exhortation to give priority to experimental results in determining the primacy of emotion and cognition.

Öhman and Wiens, too, come to a conclusion that 'appears to correspond closely to Zajonc's (1980) influential thesis that "preferences need no inferences"' (2004, p.73) and prefer to reserve the term *cognitive* to 'postperceptual processes' (ibid., p.74). They suggest a fear module between the fear stimulus and the associated cognitive processing. In this they appear to be in accord with the general suggestions of Damasio (cited above).

Figure 5: Fear module and cognitive appraisal



(After ibid., Fig 5.3(b), p. 75)

Scherer, too, suggests the *feelings* and *emotions* are 'different and distinguishable' (2004, p. 136) and urges the clarification of these terms:

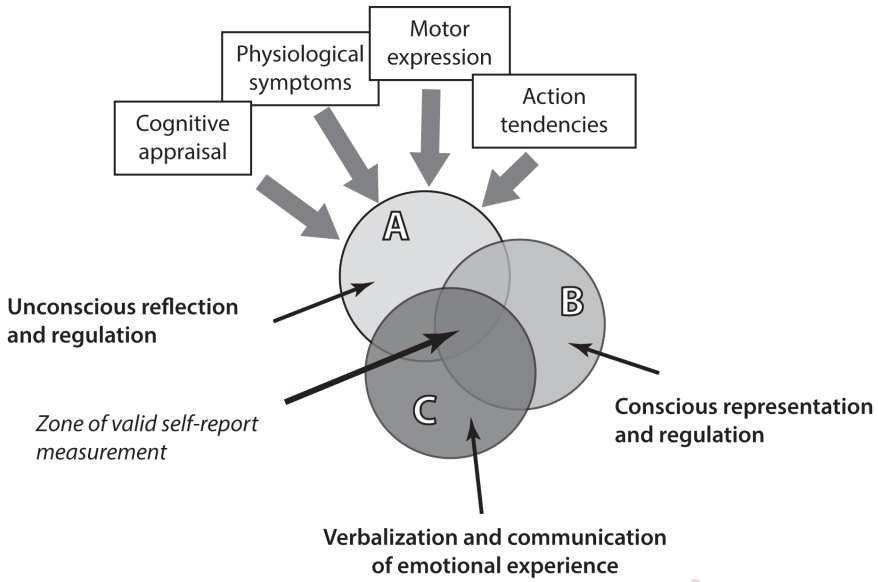
> Further progress in modeling the process of emotion can be achieved only if the central concepts are clearly defined and differentiated from each other, especially when they represent different but interrelated aspects of the underlying process, as is the case with feeling and emotion.
>
> (ibid., p.136)

He views *feeling* as a component of the emotion system that acts as a:

> … monitoring system that consists of a central representation of the response organisation, including the underlying cognitive processes in an emotion episode.
>
> (ibid., p. 137)

Feeling has a special integrative and regulatory role in which 'it needs to integrate *all* information about the continuous patterns of change in all other components, as well as their coherence' (ibid., p. 138). He represents the feelings as this separate component as the circle B in the following scheme:

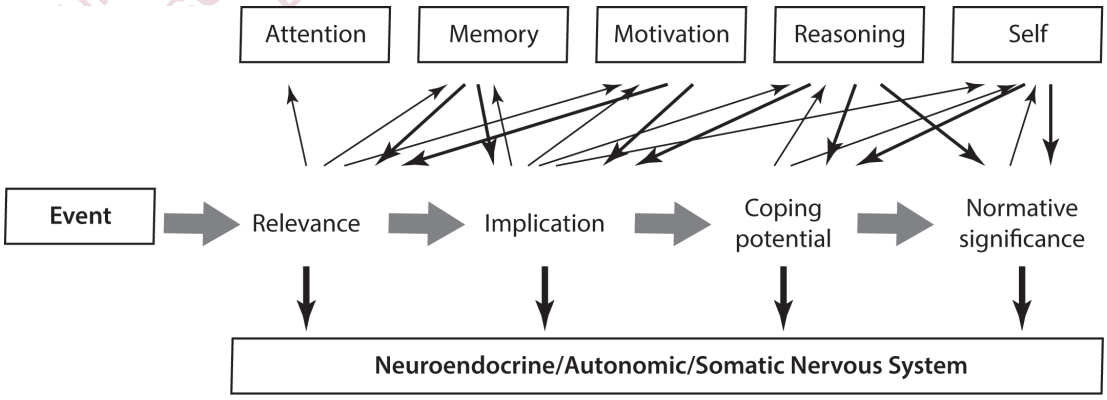
Figure 6: Three hypothetical types of central representation of component processes



(After ibid., Fig. 9.1, p. 139)

Scherer sees feeling as a 'reflection or representation of *all* the component changes produced by the appraisal' (ibid., p. 147). The model shown in **Figure 7: Central assumptions of the component process model** provides an indication of how Scherer can view *feeling* as 'subjective experience' (ibid.) that can, and should, be decoupled for specific research. Here again, there is a view that is aligned with Damasio's in separating *feeling* from *emotion.* Whereas Damasio suggested that the process whereby an 'object or situation becomes an emotionally competent stimulus often includes a conscious, cognitive appraising' (cited above), the model offered by Scherer shows promise in suggesting how that might be the case.

Figure 7: Central assumptions of the component process model



(After ibid., Fig. 9.2, p. 144)

At this juncture, fuller investigation of the mechanisms of Scherer's model is not yet warranted. What it offers is a strengthening of the case for clearly defining *emotion* and *feeling* as separate, but interrelated, phenomena. Such separation opens the pathway for better targeted empirical research and, as shown in Scherer's model, immediately provides scope for suggesting mechanisms whereby emotions can affect memory, motivation, reasoning, the sense of self, and belief. It also suggests mechanisms whereby beliefs and feeling, as proposed by Scherer, can alter the course of an emotional experience: for example, when the same stimulus produces very different responses in different people. In this model, we now see emerging the significant role of the 'self' construct which is discussed in Chapter 4.

Frijda (2004), along with Scherer (2004), includes as *appraisal* 'all processes, automatic or cognitive, that provide objects and events with emotional value or meaning' (2004, p. 163). While vague on the specifics of how appraisal might work, he, too, is in accord with Damasio's acknowledgement that there is often a cognitive appraisal component in emotional response.

Panksepp considers that cognitive skills 'emerge in the service of various basic needs' (2004, p. 188) and that 'basic affective feelings reflect characteristic neurodynamics that generate ancient emotional-instinctual tendencies of the body' (ibid., p. 189). In his view, the higher neocortical areas (anterior cingulate, peri-amygdaloid, insular, perihinal, orbito-frontal, frontal cortical areas) are likely to contribute 'comparatively little to the *generation* of affect' (ibid., pp. 180–181) but rather are instruments for emotional regulation and for learning about 'affective nuances of world events' (ibid., p. 181).

In line with Damasio's view of feelings$_D$ and Panksepp's suggestions, Winston and Dolan consider the empirical evidence implicates 'widespread cortical areas' (2004, p. 216) as providing the neural basis for feeling.

Frijda et al. (Frijda *et al.* 2004) summarise the findings from The Amsterdam Symposium (Manstead *et al.* 2004) as follows:

> Researchers are nearly unanimous in distinguishing emotions from feelings. This reflects the common contemporary view that emotional processes exist independently of awareness and may occur without awareness, and that feelings have a special place in emotional reactions.
>
> (Frijda *et al.* 2004, p. 458)

Explicit or implicit in many of the contributions to The Amsterdam Symposium is the call for greater clarity with the terms *emotion* and *feeling*. The very nature of Damasio's hypothesis, in giving distinct definitions of these terms, leaves no doubt as to its contribution to this discussion. His definitions of emotion$_D$ and feeling$_D$ and his suggested

relation to consciousness and cognition give possible explanations of, or insight into, mechanisms for:

- appraisal
- motivation and activation
- the neurobiological nature of emotions
- rationality and/or unconscious emotion
- affect and cognition.

3.3 Emotion, Attention and Memory

It has been established that there is broad agreement that there is a strong link—indeed, a feedback loop—between emotion and belief. However, the literature largely focuses on the direction of belief to emotion: that is, how existing beliefs can induce and affect emotion. Little appears to be proposed with respect to *mechanisms* whereby emotions can generate or strengthen beliefs, although there are significant hints that such mechanisms exist, as can be inferred from the research discussed in the previous sections (*Emotion and Belief* and *The Nature of Emotion*).

In section *2.1.3.2 Attention, Inattention and Change Blindness* it was established that false beliefs can be established and maintained on the basis of what a person perceives. Implicit in the process is the belief that perceptions are accurate; and, the more strongly this belief is held, the greater is the reliance on the presupposed validity of the the perceptions. The effect of such meta-beliefs, and of belief systems, is discussed in Chapter 4.

The role of emotion in affecting perceptual filtering was also introduced in *2.3 Other Factors Affecting Perceptions* on page 84. It is this effect of emotion on perceptual filtering that points to one way by which emotions can affect or even induce beliefs and warrants the following review of relevant literature.

3.3.1 Emotion and Attention

Rowe et al. note the effect of mood on perceptual resources and suggest that 'viewing the world "through rose-coloured glasses" may be less proverb and more an empirical fact' (2007, p. 383). Vuilleumier and Huang (Vuilleumier & Huang 2009) also pick up the rose-coloured glasses metaphor and make the observation that is generally familiar to most people:

> Some see the world with gloomy shades, while others see it through rose-colored glasses. But how is it possible that emotion may govern our senses? Past philosophers and scientists (such as Descartes or Zajonc) have often considered that perception and cognition are separate from emotion processing. However, increasing evidence

> from both psychology and neuroscience now clearly indicates that emotion and cognition do not operate entirely separately but reciprocally influence each other. Although it may come as no surprise that what we look at can determine what emotions we will experience, a much more fascinating finding from recent research is that, conversely, emotional reactions may influence what we perceive from the external world.
>
> (ibid., p. 148)

Anderson and Phelps claim to show a 'neural substrate for affective influences on perception' (2001, p. 305). Their results, they suggest, show that

> the amygdala-mediated neural modifications may … enhance perceptual sensitivity to events of importance to the organism, making them less dependent on attentional resources to achieve awareness.
>
> (ibid., p. 308)

This view is supported in the later work of Phelps et al. (2006) (also cited on page 85) in which they declare their expectation that emotionally salient events would influence 'even the most basic perceptual processes' (ibid., p. 298), and they present research showing the effect of emotion on early visual processing in which perception is facilitated and attention potentiated. Their experiment concentrated on fear: the presence of a 'fearful face' (ibid., p. 297) enhanced contrast sensitivity and magnified transient covert attention. They acknowledge that effects could vary depending on the emotion involved and the type of the attention being investigated.

Additionally, Phelps (2006) proposes that the amygdala is 'ideally situated to influence perception with emotion' (ibid., p. 39) and that this, in turn, could facilitate attention. It is proposed that emotion could capture attention by impairing processing of the non-emotional aspects of an event.

Phelps observes that studies on the capture of attention 'have primarily used negative, fearful, or threatening stimuli' (ibid., p. 40); however, she draws attention to work examining other aspects of attention capture. Anderson (2005) presents studies demonstrating that 'arousal reduces the attentional prerequisites for perceptual consolidation, promoting enhanced entry into awareness' (ibid., p. 273). Öhman et al. (2001) propose a system that 'automatically evaluates the significance of the output from a feature detection perceptual system' (ibid., p. 467). In their experiments they found:

> Participants specifically fearful of snakes but not spiders (or vice versa) showed facilitated search for the feared objects but did not differ from controls in search for nonfeared fear-relevant or fear-irrelevant, targets.
>
> (ibid., p. 466)

Phelps (2006) suggests that studies such as these support earlier discussion (for example, Whalen 1998) and provide evidence of emotion's influence on attention and perception and that findings are consistent with the interpretation that emotion facilitates preferential detection of stimuli that signal importance or threat. Tibboel et al. (2011) raised the question of whether the apparent effect of arousal on attention—as reported by Anderson (2005), for example—could be caused by report bias. Their findings, however, supported theories (for example, ibid.; Öhman *et al.* 2001) that 'emotionally arousing stimuli are preferentially attended to and processed more efficiently than neutral information' (Tibboel *et al.* 2011, p. 1181).

The effects of emotion go beyond capturing attention in response to signals of importance or threat. Dreisbach and Groschke (2004, p. 351) suggest that 'adaptive action requires a dynamic, context-dependent balance between maintenance and switching of goals, cognitive sets, and behavioural dispositions' and present data that provide support for the role of affect in modulating cognitive control. This control 'enhances cognitive flexibility but also incurs a complementary cost in the form of increased distractibility' (ibid., p. 351). Plausibly, the increased distractibility could occur, they suggest, because the positive affect signals an absence of danger, for example, thereby allowing for more 'explorative' thinking.

Gray (2004) tested the selective effects of emotions on cognitive control. The work was carried out in the view that emotion and cognitive control are integrated and at times even work in harmony.

> The main finding was that some brain areas showed a selective effect: The influence of a particular emotion (amusement or anxiety, induced by the same videos as before) depended on the type of cognitive-control task (verbal, non-verbal), so the results met strict criteria for demonstrating integration.
>
> (ibid., p. 48)

Mitchell and Phillips, in a review of the psychological and biological bases of the relationship between *normal* mood and executive functions, note the 'relative lack of research on this topic' (2007, p. 627). Nevertheless, they conclude that:

> … minor fluctuations in normal mood can affect executive functioning, such that positive mood impairs many aspects of executive function, possibly due to increased reliance on heuristic processing. Negative mood appears to have less effect on executive functions, although few relevant studies have been carried out. Dopamine may be responsible for mediating the cognitive effects of positive mood, and serotonin may be involved in fluctuations in negative mood.
>
> (ibid., p. 627)

Both Mitchell and Phillips (ibid.) and Lindström and Bohlin (2011) distinguish moods as being longer lasting and milder than emotions and without the same intensity of physiological and cognitive reactions as emotions. Beedie et al. (2005) discuss the difficulty in distinguishing emotion from mood and suggest that they 'are closely related but distinct phenomena' (ibid., p. 847) but do acknowledge that any distinctions between them are 'clouded' (ibid., p. 847). Engaging in pursuing any distinctions between mood and emotion is beyond the scope of this inquiry. In the general use of the words it would seem that there has to be a close interaction between mood and emotion if, indeed, they are distinguishable. Any chance of distinguishing mood from emotion is dependent on agreed definitions. For the purposes here, it is taken that there is a close connection between emotion and mood, at least on occasions. For example, an event that produces an energetically angry response can result in a 'bad' mood that lasts for hours or even days. The sustaining of such a mood may be assisted by cognitive rumination—that is, regular review of the events and reminding oneself of the wrongs perceived. However, at other times mood might be considered as what happens *between* emotions (Damasio 1994). Beedie et al. pose the question as to whether they are 'actually the same construct in different guises' (2005, p. 874).

Returning to the findings of Mitchell and Phillips (2007), it is notable that even *minor* fluctuations in mood affect executive functions. Intuitively, this is to be expected in the case of emotions which are felt and are possibly intense. However, it is important to recall that emotions—especially as defined by Damasio (1994, 1999, 2003b)—can be experienced unconsciously. Hence, it is reasonable to hypothesise that 'mood' may well emerge as a result of unconscious emotions or emotional experiences of low intensity, at least in part.

Clearly, mood affects cognitive function as does what we may distinguish as emotion. For the purposes of this inquiry, given the lack of clarity regarding any differences between 'mood' and 'emotion', mood is taken to be a state that is derived from low intensity emotion or is a result of emotional and cognitive interactivity.

One particularly interesting effect of emotions is that on the perception of time. Wittman (2009) provides a review of literature on these effects. Some relevant effects Wittman notes are:

- The more experiences for a given time duration that are stored and retrieved, the longer the duration is perceived. Conversely, routine activity leads to a perception of shorter time duration. (ibid., pp. 1956–1957)
- When in an unpleasant or uneventful situation people will overestimate the time duration. Conversely, being engaged in a rewarding activity leads people to underestimate the duration of the event. (ibid.p. 1960)

- People prone to boredom estimate duration to be longer than the actual duration. Stress and fear also lead to an overestimation of duration. (ibid.p. 1960)

Other papers examined (e.g., Craig 2008; Droit-Volet & Gil 2009; Droit-Volet & Meck 2007) that relate to the effects of emotion on time perception have been examined by Wittman (2009) in his review. With the findings in the significant body of research cited by Wittman (op. cit.) and Hammond (2012), there is little doubt that emotions affect the perception of time. Of course, this is intuitively obvious, and would find little argument in folk psychology. The acceptance of the perceptual shift is signified by the investigation and discussion arising that relates to the mechanisms by which such a perceptual effect can be caused. Droit-Volet and Gil succinctly summarise the findings when they say that a series of studies showed that:

> "the representation of a particular duration is highly context dependent. It depends on both intrinsic context, such as the emotional state at the onset of time processing, and extrinsic context, such as others' activity rhythm
>
> (Droit-Volet & Gil 2009, p. 1950)

While the concentration in the literature reviewed is focused on understanding the mechanism for such effects rather than how they might lead to false beliefs, it is straightforward to propose plausible examples. Under the influence of fear, for example, a person who is present during a bank robbery could easily overestimate the timing of specific events when giving witness statements. People involved in accidents such as rolling a vehicle report experiencing significant time distortion: it is not unusual to hear people say something like *'Everything slowed right down, it seemed to take ages'*. Such a perception could later have a profound influence on how the autobiographical memory is established and later recounted. Someone who is bored on a holiday might report that a friend (or partner or parents) *'made me wait **all** morning for them'* when, in fact, the time had been only one hour; yet, years later, the report of 'all morning' is what they hold as truth. Conversely, someone engaged in a rewarding task might energetically argue that they had completed a report in significantly less time than it had really taken if they made no reference to a clock.

In an example of fear relevance, Öhman et al. observe that 'it is often the spider phobic individual in the company who discovers the spider inconspicuously moving across the ceiling' (2001, p 467). While Öhman et al. do not discuss the extrapolation of this effect of fear on how attention is directed, it is easy to imagine how the spider phobic person leaves a place with a firm belief that 'the place was crawling with spiders' on the basis of having noticed three or four spiders that passed unnoticed by others.

These few simple examples suggest the possibility of beliefs being formed from the direction, or misdirection, of attention by emotions. It is clear that emotion can affect attention, and it is clear that the *interpretation* of what is then perceived can easily develop into a belief. The mechanisms by which this might occur are shown by the literature to be complex and, at this stage, little understood. However, this highlights the desirability of further inquiry into these mechanisms as the consequences of a false belief cannot be dispelled simply by an appeal to logic when the belief is born of unconsciously driven evolutionary responses that evoke such strong emotions that the person's record of events (based on the altered perceptions) is taken to be real.

In a link back to the discussion of Chapter 2, Öhman and Mineka (2001) draw on illusory correlation studies, among others, in describing an evolved module for fear learning and fear elicitation. So we can imagine a spider phobic child being taken on holidays to visit relatives in a country setting where there are large but non-venomous spiders. The child sees one or two of these creatures, although others do not notice them, and forms the belief that 'it is dangerous in that part of the country'. Years later, their partner suggests that they holiday in that part of the country because of the beautiful nature walks; the spider phobic person insists they avoid that part of the country because of the danger, and an argument and marital stress ensues.

Given that that there is little debate over the view that emotions can affect perceptions, the conclusions from the heuristics and biases section become even more significant because of the suggested complexity of the interaction between memory systems, affect systems and attentional systems. The need to address the complexity is aptly captured by Phelps, and this is particularly pertinent to the research discussed in Chapter 4:

> *Examining cognitive functions without an appreciation for the social, emotional, and motivational context will result in an understanding that may be limited in its applicability outside of the research laboratory*. The traditional research domains of psychology, such as cognitive, social, and clinical, may help create unified areas of research, but may also diminish our appreciation of the complexity of human behaviour by discouraging discussion of their interactions. Adding the complexity of emotion to the study of cognition can be daunting, but investigations of the neural mechanisms underlying these behaviours can help clarify the structure and mechanisms.
>
> (Phelps 2006, p. 47, emphasis added)

In a look to material more relevant to Chapter 4, Wittman points to the role of the perception of time in autobiographical memory—intricately interrelated with beliefs—and asks:

> What if mood processes and the representation of body sensations themselves function as a timekeeper? Since emotions and physiological states seem so fundamental to the experience of time, it is tempting to assign a pivotal role to these processes related to a core time-keeping system.
>
> (Wittmann 2009, p. 1961)

3.3.2 Emotion and Memory

Uttl et al. (2006b) draw attention to William James' awareness of the effect of emotion on memory. In James' words:

> An impression may be so exciting emotionally as almost to leave a *scar* on the cerebral tissues; and thus originates a pathological delusion.
>
> (James 1890/1950b, p. 670)

Interestingly, Uttl et al. (op. cit.), when quoting James, do not include the final clause of James's sentence; yet this is particularly pertinent to this thesis. James goes on to give examples:

> A woman attacked by robbers takes all the men whom she sees, even her own son, for brigands bent on killing her. Another woman sees her child run over by a horse; no amount of reasoning, not even the sight of the living child, will persuade her that he is not killed.
>
> (ibid., p. 670)

James is presenting this in the context of what makes memories more easily recalled and consisting of more detail and more vivid detail. In a forecast of the work discussed above (*3.3.1 Emotion and Attention*), James observes:

> The *attention* which we lend to an experience is proportional to its vivid or interesting character; and it is a notorious fact that what interests us most vividly at the time is, other things equal, what we remember best.
>
> (ibid., p. 670)

As has been demonstrated in Section 3.3.1, there is agreement that emotions affect attention and perception. Much of the current literature is investigating mechanisms by which such effects might occur. Some of that discussion is beyond the scope of this thesis: for example, the complexities of the neuroscience findings. However, the more generally agreed aspects are relevant as is some of the discussion from appraisal theories.

Reisberg (2006) invokes the need to include cognitive influences as well as emotional arousal to explain memory narrowing in which emotion 'improves memory for an event's gist, but undermines memory for more peripheral elements within the event' (ibid., p.

15). In line with James' observations and from recent research, some of the effects of emotion that Reisberg highlights include:

- emotional events tend to be remembered well
- the memories of emotional events seem to be longer lasting—emotion appears to slow the process of forgetting
- when comparing events that are similar but in circumstances where one is emotional and the other is not, the emotional memories tend to be more accurate

These kinds of observations can suggest that emotional memories are indelible, but Levine and Pizarro (2004), in their review of various claims from research findings, report that emotional memories are not free from error. They do suggest, as has been implied earlier (Section 3.3.1), that 'what matters most may be *that* the event occurred, the gist of what occurred, and the implications it had for the individual' (ibid., p. 533). Further, they conclude that neither the memory for an emotional event nor the memory for the emotional feeling is indelible.

Reisberg (2006) generally agrees with Levine and Pizarro but notes that they draw their conclusions predominantly from a social psyschology perspective while Reisberg comes from a cognitive psychology perspective. His view is that while we understand *some* of the biological processes—particularly the role of the amygdala—responsible for the emotion's effects on memory, and acknowledge that these processes may be necessary, they are not sufficient for the effects. In particular, he questions whether the memory-narrowing effect is solely the result of emotional arousal or whether it arises from the presence of visual stimuli that are salient in the moment. An example is 'weapon focus' that might occur if one is involved in bank robbery scenario: what the weapon is and its likely impact on one's person is highly salient in such a situation. It's no wonder then, suggests Reisberg, that a person will zoom in perceptually on the weapon to assess the degree of immediate personal threat.

In a study in which visually salient elements were excluded and the emotions were aroused auditorally via a narrative, the memory narrowing effect was not evident. He acknowledges that this does not diminish the importance of arousal-based mechanisms or of memory narrowing 'for the simple reason that many emotional events … do contain salient visual stimuli' (ibid., p. 27); and he continues to suggest that memory narrowing must be understood in order to understand how different types of emotional events are remembered.

An interesting challenge is raised by Reisberg, particularly for this study, with respect to how events are **mis**remembered when he observes that this aspect has been 'largely unexplored' (ibid., p. 29). An earlier element in addressing this question may be in

Reisberg's own paper. He notes that memory narrowing occurs in both positive and negative events and proposes the narrowing occurs if there is a suitable 'attention magnet' (ibid., p. 31) in the event. Levine and Pizarro (2004) draw on a highlighter metaphor to suggest how this might work:

> If happiness serves as a highlighter, it appears to be a broad and inclusive one that increases the salience of a wide swath of information, some from general knowledge, some from the environment. The composite representations that result tend to be vivid and creative but not very discriminating. Negative emotions, on the other hand, may work like fine-tip highlighters that increase the salience of a narrow range of information in the service of either preventing, fixing, or adjusting to goal failure. Given the differing motivations associated with specific negative emotions such as fear, anger, and sadness, the types of information likely to be encoded and retrieved in these states should differ.
>
> (ibid., p. 547)

If Levine and Pizarro are correct, and combined with Reisberg's (op. cit.) observations, then Reisberg's suggestion that cognitive aspects of emotion need to be included in the account of emotions' effects on memory is valid. Levine and Pizarro (2006) argue the case for drawing on appraisal theories of emotion for a complete understanding of emotions' effects on memory. They claim that since elements such as motivations and problem-solving strategies are likely to vary with different emotions, they should be taken into account. Findings that 'people process and remember information differently depending on whether they are feeling happy, fearful, angry, or sad' (ibid., p. 37) are examined in support of this view. They reiterate the general consensus the 'emotional arousal has a powerful effect on memory' (ibid., p. 38) and note that the evidence comes from a number of types of study including autobiographical memory studies, laboratory studies on both animals and humans, and brain imaging work. What is not accounted for by the arousal-based models, they claim, is the effect on judgment in particular. Their review of the literature shows that judgments vary depending on whether the person experiences positive or negative emotions.

> … happiness appears to promote a schema-based, top-down information processing strategy in which people draw freely and flexibly on prior knowledge … When people are feeling negative emotions, they tend to process information in an analytic, data-driven manner, and are more conservative in their judgments …
>
> (ibid., p. 41)

The suggestion from Levine and Pizarro is that because motivations differ depending on whether positive or negative emotions are being experienced, the information processing is

different and these differences affect memory. Drawing on the work of Herbert Simon[39], they suggest that discrete emotions serve as a solution to the monitoring of the environment that alerts people to signals that are relevant to their multiple goals. The proposal from appraisal theory is that the discrete emotions signal that a response is required as there is an event occurring that is relevant to at least one of the person's goal. That the monitoring can be unconscious as well as conscious makes the role of emotions as a signal for response all the more significant. This view of emotions being linked to goals may help make sense of Elster's use of 'desire' and 'intentional objects' (see discussion starting on page 95).

Emotions, then, serve to reprioritise goals and reorganise behaviour, perception, judgment and memory (ibid.). Levine and Pizarro (ibid.) pick up the suggestion of Lerner and Keltner that emotions tend to 'activate appraisal tendencies' (Lerner & Keltner 2000, p. 489). Damasio's definition of emotions (see *3.2.3 Separating Emotions and Feelings*) lends weight to this view as emotions are signalling concerns for the living system from basic homeostatic needs (for example, water) through to social needs.

While the mechanisms underlying the processes whereby emotions affect memory need more research, there are the beginnings of possible explanations in the literature looking to appraisal theories. Aspects of appraisal theories were discussed in Section 3.2.3.3 and are examined more closely next.

In Section 1.1.1 it was noted that Schacter and Scarry (2000a) propose that memory is a form of belief. Additional support for this view is discussed further in Chapter 4; however, if the view is sustainable—and it is suggested here that it is—then by affecting memory, emotions have a direct affect on at least some beliefs. This is saying *that* emotion affects belief but does not yet offer insight into *how* that might occur.

3.4 Towards an understanding of the influence of affect

Even at the time of writing,[40] the collection *Emotions and Beliefs* (Frijda *et al.* 2000a) appears to be the *only* collection of work specifically and overtly devoted to the investigation of the nexus between emotions and beliefs and, more specifically, on the role of emotions in creating beliefs, maintaining beliefs, changing the strengths of beliefs (that is, how well they resist challenge), and changing beliefs. Of course, there are numerous articles on which this collection relies for the various arguments presented and many more have been published since the collection's publication. However, some considerable

39. *Models of Thought*, 1967 (not available to sight at the time of writing)

40. 2013

attention is given to the articles contained in the collection because they manage to bring together many of the threads of inquiry that have bearing on understanding the role of emotions in influencing beliefs.

3.4.1 Emotional engagement of existing beliefs

Fielder and Bless (2000) note (at the time of writing) that while there had been considerable research on the effects of emotion on memory, thinking and social judgment, there had been little research on the effects of emotions on beliefs[41]. They argue that 'beliefs should be particularly sensitive to affective influences' because they can be considered to be 'located at the very interface of emotion and cognition' (ibid., p. 144). In proposing four defining features of beliefs[42] they suggest that two types of adaptive regulation—assimilation and accommodation—are applicable in examining the proposed interface of emotion and cognition. This dichomoty is drawn from Piaget's work[43] and they describe *assimilation* as a 'top-down process in which the individual imposes her cognitive structures and schemata that have been successfully employed to new problems and affordances' while *accommodation* is 'stimulus-driven and charaterised by the organism's readiness to react reliably to external demands or threats [which means] sticking to the stimulus facts and updating one's internal structures as a function of external requirements' (ibid., p. 145).

Fiedler and Bless suggest that many of the findings relating to affect and cognition can be considered as referring to *beliefs* where the beliefs are:

> conceived as cognitions that involve inference under uncertainty, genuine involvement, active elaboration, and cooperation.
>
> (ibid., p. 146)

Using their view of assimilation and accommodation, Fielder and Bless (op. cit.) offer the following theoretical assumption:

> Positive emotional states facilitate active generation, whereas negative emotional states support the conservation of input data.
>
> (ibid., p. 147)

41. Writing in 2000. Of course, there has been more since then, but in the literature reviewed the focus remains predominantly on the effects on attention and perception.
42. (1) There is a distinction between believing and knowing. (2) There is a distinction between believing and saying. (3) Believing involves *not* distrusting. (4) The 'semantic surplus meaning of believing …is not applicable to passively held or tolerated information' (ibid., p. 144).
43. *The Origins of Intelligence in Children,* 1952.

While Fieldler and Bless observe many similarities with their model and others (for example, Clore & Gasper 2000; Forgas 2000b), they admit to significant differences with respect to the interpretation of the empirical results when positive affect is studied. Fiedler and Bless cite some researchers who suggest there is reduced cognitive effort, for example, whereas they propose that positive emotions 'elicit active cognitive processes' (Fiedler & Bless 2000, p. 150). The focus of their work was to find where their theory both converges with and diverges from other theories. They claim their experiments allow the concepts assimilation and accommodation—which could be considered vague—to be 'tied down to relatively "hard" paradigms' (ibid., p. 164). Beliefs, in their view, become 'agents of assimilation' (ibid., p. 164). Overall, they propose that:

> Positive mood states, due to their assimilative power, increase the reliance on beliefs systems, especially positive beliefs. However, thanks to the accomodative function of negative affective states, beliefs systems can be updated in the light of new significant data, thus preventing the individual from losing reality.
>
> (ibid., p. 164)

In proposing the active role of beliefs when the person is experiencing positive emotions Fieldler and Bless differ from others who suggest less cognitive activity or greater distractibility. If their model withstands scrutiny, this particular aspect may offer a key to how beliefs may be both strengthened or created by affect.

Forgas (2013) looks to Fiedler and Bless's (op. cit.) proposal for assimilative and accommodative processes as being the most plausible explanation for the differences in how affective states influence the *process* of cognition, this being the '*how* people construct and use beliefs' (ibid., p. 6). However, the effects on *processing* are just one of two types of effects that Forgas draws attention to: the other is *content* effects in which affect influences *what* we think as well. He suggests that these content effects have been explained by two complementary theories: memory-based accounts (for example, affect priming as discussed in Fiedler & Bless 2000; Forgas 2002) and inferential models in which, for example, affect is taken as information (Clore & Gasper 2000).

In proposing the Affect Infusion Model (AIM), Forgas and others (Eich & Macauley 2000; Forgas 1995, 2002, 2013) suggested that there were limitations to affect priming models and affect-as-information models:

> Despite some impressive evidence supporting each of these two models, neither can fully explain the growing variety of empirical findings … It is argued here that these mechanisms represent complementary rather than conflicting avenues of affect infusion, operating under different (substantive vs. heuristic) processing conditions,

> and that they are empirically distinguishable through the analysis of processing latency, judgmental latency, memory, and other cognitive variables.
>
> (Forgas 1995, p. 40)

Fiedler and Bless suggest their approach using assimilation and accommodation provides an *integrated* approach for interpreting the empirical results and observe that Forgas's account converges 'in many respects' (2000, p. 146) with their account. Forgas (2000b) views the AIM as being more comprehensive than Fielder and Bless's approach but acknowledges the complementary nature of the approaches. In reviewing an experiment in which positive mood increased memory distortions while negative mood improved accuracy of eyewitness accounts, Forgas suggests:

> These results add an important dimension to Fiedler and Bless's conclusion that positive moods support the use of exising beliefs, while negative moods weaken belief systems.
>
> (ibid., p. 131)

Forgas's Affect Infusion Model (1995, 2000b) aims to answer a number of issues, including:

- accounting for apparent context sensitivity of many moods on beliefs
- identifying the types of information processing strategies that are likely to bring affect to bear on beliefs
- identifying what types of beliefs are least likely to be open to distortion from affective influences
- determining the role of affect in processing complex or problematic information
- determining whether prolonged, systematic processing is more or less likely to produce affect infusion.

Affect infusion is defined by Forgas as a process 'whereby affectively loaded information exerts an influence on, and becomes incorporated into a person's cognitive processes, entering into their constructive deliberations and eventually colouring the ourcome in a mood-congruent direction' (Forgas 2000b, p. 110). An important aspect of the AIM is that it:

> … assumes that affective states, although distinct from cognitive processes, do interact with and inform cognition and judgments by influencing the availability of cognitive constructs used in the constructive processing of information.
>
> (Forgas 1995, p. 41)

The AIM (ibid.; Forgas 2000b) identifies four processing strategies:

1 *direct access* of pre-existing information

2 *motivated processing* in the service of a pre-existing goal (and, it is suggested here, pre-existing beliefs)

3 a simplified *heuristic*

4 a *substantive heuristic* involving an elaborate processing strategy.

The direct access strategy is considered to be resistant to affective distortion because it involves 'little or no constructive elaboration' (Forgas 1995, p. 46). With motivated processing it is considered that the person will use highly selective and targeted information search strategies to support a pre-existing motivational objective. As with the first strategy, this too is thought to be a 'low-infustion judgmental strategy' (ibid., p. 47).

In the heuristic strategies (three and four), Forgas (ibid., 2000b) considers that the requirement for elaborate and constructive processing allows either a direct or indirect influence by affect. Forgas (1995) sees that this model assists a 'comprehensive theory of judgments' (p. 47) because the:

> different processing strategies imply different kinds of mood effects on judgments, including the possibility of no mood effects or mood incongruence when direct access or motivated processing is used.
>
> (ibid., p. 47)

It is noted here that Forgas (2000b) uses the term *affect* as a generic label referring to both *moods* and *emotions*. His observations on the problems with defining these and other terms were noted earlier (see page 94). Despite the challenge he suggests that:

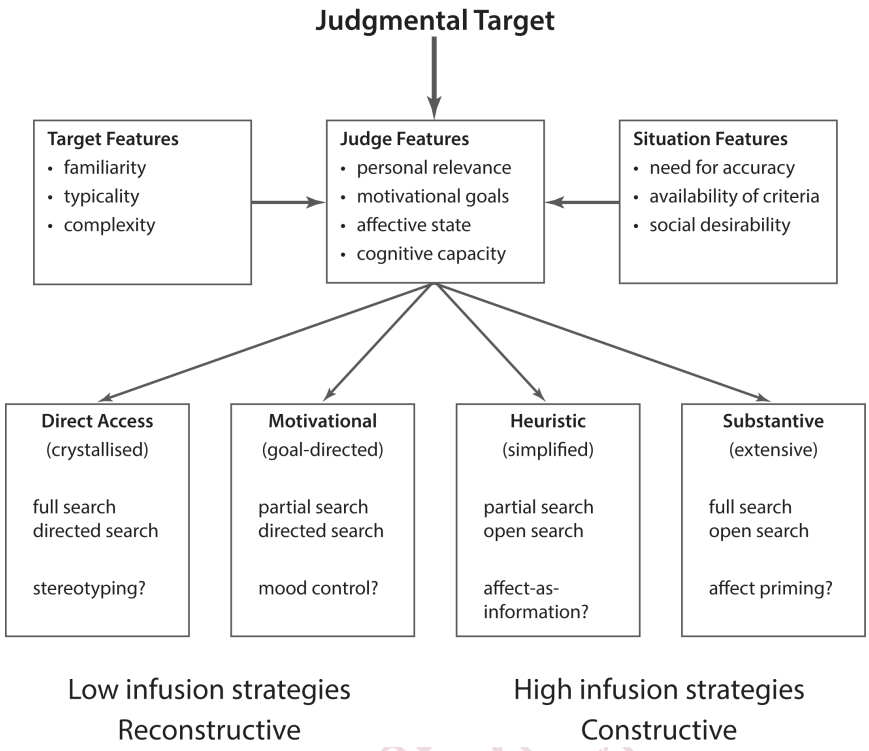
> *Moods* in turn **could** [emphasis added] be described as 'low-intensity diffuse and relatively enduring affective states without a salient antecedent cause and therefore little cognitive content (e.g. feeling good or feeling bad)', where *emotions* 'are more intense, short-lived and usually have a definite cause and 'clear cognitive content'
>
> (ibid., p. 110)

Forgas's tentative definitions are repeated here[44] as he draws attention to research suggesting that 'subtle, non-specific moods' (ibid., p. 10) may have a more enduring effect on beliefs than emotions (as defined by him). However, Forgas's definition does not obviously take into account the possibility of a mood existing because of the cognitive rumination on an earlier antecedent cause, as suggested on page 114.

A diagrammatic representation of the AIM is shown below.

44. Previously mentioned on page 94

Figure 8: Affect Infusion Model



(After ibid., p. 116)[45]

Where Forgas distinguishes the AIM from other multiprocess frameworks with which it has some affinity, is in its focusing specifically on the 'role of affect in processing and judgments' (1995, p. 47) and also in its enabling of testable predictions. One of its most important predictions—one which Forgas suggests is counter-intuitive—is that affect infusion into beliefs should be increased by prolonged and more substantive processing.

As mentioned above, it is in the heuristic and substantive strategies—the constructive strategies—that affect infusion occurs, according to this theory. With *heuristic* processing Forgas suggests that there are 'no precomputed beliefs' (Forgas 2000b, p. 112) at play, nor is there any strong motivation directing the effect of those beliefs. Such heuristics come into play when a response that requires minimal effort is needed. He further suggests that social beliefs are formed when the heuristic cues are used and that they are based on such elements as stereotypes, environmental variables that are not relevant, and misinterpretation of a current emotion. A more obvious indication of how affect can infuse a belief is the reliance on the 'how do I feel about it?' heuristic.

45. Appendix 1 contains a flowchart illustrating the hierarchy of factors that determine processing choices and the consequences of affect on judgments (Figure 17).

However, affect infusion is much more likely to occur with what Forgas calls *substantive* processing occurs (Forgas 1995, 2000b). Forgas seems to combine all the elements of this dictionary definition of 'substantive' into his application of the word:

> Having a firm or solid basis; important, significant; of substantial extent or amount …
>
> (Oxford University Press 2007)

In Forgas's words, substantive processing is a demanding processing mechanism that:

> … requires judges to select, learn, and interpret novel information about a target and relate this information to preexisting knowledge structures. {It] is more likely when the target is complex or atypical and the judge has no specific motivation to pursue, has adequate cognitive capacity, and is motivated to be accurate, possibly because of explicit or implicit situational demands.
>
> (Forgas 1995 , p. 47)

Forgas cites the role of affect in engaging connected memories and how this, in turn, has an influence on the 'selection, learning, interpretation and assimilation of novel information into a pre-existing knowledge base' (Forgas 2000b, p. 113). Results show that as the processing required becomes more elaborate, it becomes more likely that affectively primed information will infuse beliefs and representations—as predicted by the theory.

Forgas (ibid.) presents a strong review of empirical evidence supporting the AIM and its ability to predict the influence of affect depending on whether the person is experiencing positive or negative moods. It accounts for both affect-priming mechanisms and for affect-as-information mechanisms. Additionally, it accounts for the resistance to affect infusion when direct access and motivating strategies are engaged. An area for further development, Forgas suggests, is developing the AIM to more comprehensively account for individual differences arising from personality factors such as trait anxiety.

From the literature reviewed, one of the strengths of the AIM is its ability to address the different types of responses to affect and by addressing different processing strategies and factors affecting their selection, it provides insight which more nebulous appraisal theories fail to do. However, while the theory appears to have strong predictive ability about the conditions under which it is more likely that emotion will affect beliefs, it still does not present significant specific insight into any detailed *mechanism* by which that might occur. We do see in Figure 17 a clear link to appraisal type mechanisms, and Forgas does suggest that emotions activate memories (citing Bower 1981; Bower 1991; Atkinson & Adolphs 2005). As well, we are presented with evidence that positive moods engage a form of top-down processing that relies on existing belief systems while negative moods can bring about the questioning and weakening of beliefs. With respect to this latter effect from

negative moods, there is the complicating factor that negative emotions *can* lead to the strengthening of a belief as demonstrated in work on cognitive dissonance (Harmon-Jones 2000).

What becomes obvious at this point is that a false belief formed under the influence of an emotion is not just a result of the attentional shift or perceptual distortions caused by the emotion. Even if there were a temptation to attribute a belief formed in a strongly emotional state—either positive or negative—to attentional or perceptual effects solely, there is still the necessity to consider that the believer is relying on a pre-existing belief that they can trust their perceptions, even if that belief is held unconsciously.

Then if memories are beliefs (Schacter & Scarry 2000a)[46]—some memories, if not all—we begin to glimpse how the emotional engagement of memories that may be beliefs as well as the reference to other pre-existing beliefs provides an elaborate (to use Forgas's word) cognitive computation that can strengthen an existing belief or generate a new one. This could be summarised by saying that Forgas, along with many others, is able to show *that* emotions maintain or change beliefs or create new beliefs, but gives only a very high level indication of *how* this might happen.

3.4.2 Emotions as part of beliefs

Frijda and Mesquita also acknowledge the effect of emotions as 'being accepted as a fact on the basis of massive informal evidence' (2000, p. 46) and consider that emotions affect beliefs by either creating beliefs where none previously existed or by altering the strength of the belief.

Their reference to belief strength is one of the key issues for this thesis. They note that a belief can be held sufficiently strongly that its possessor will remain convinced of its truth regardless of the strength of evidence against it.

> A belief may persist in the face of evidence that contradicts it. The belief renders that evidence powerless. It is ignored or dismissed; arguing appears useless. Furthermore, a person holding a belief may accept information as credible that would appear doubtful to other people or, by contrast, is considered as not worth believing when it appears convincing to others.
>
> (ibid., pp. 45–46)

It is plausible, argue Frijda and Mesquita, that belief strength can be increased by emotional meaning. Presumably, the authors are using 'meaning' to refer to the meaning that we attach to the emotion(s) being experienced. This would fit with the proposal that

46. A proposal discussed briefly in Section 1.1.1

affect that acts as information not only informs actions but also shapes interpretation. It also plausibly fits with Damasio's proposed definition that *feeling* is the conscious awareness of the underlying emotion which is the somatic response to events. If this is accepted, then $feelings_D$ become a principal instrument in the shaping of the meaning that is associated with the emotion experienced.

Frijda and Mesquita (ibid.) acknowledge that emotions and beliefs can be considered as different types of psychological events and remind the reader that Hume, indeed, thought so. However, the challenge for this notion is the intimate relationship between belief and emotion: so much so that Frijda and Mesquita boldly state:

> Beliefs are part of emotions.
>
> (ibid., p. 52)

They then make a direct link to appraisal theory and state clearly their view that beliefs that are part of emotion *are* the cognitive appraisals in being the 'perception or evaluation of the eliciting object as beneficial or harmful for one's person or for [the individual's] concerns' (ibid., p. 52). To some degree, however, they perpetuate an ambiguity in appraisal theory when, later, they refer to beliefs as being the *product* of appraisal processes. This ambiguity was noted by Frijda (1993) previously when he suggested the 'appraisal' refers to two phenomena. The first is its use to refer to the *content* of emotional experience; and the second is its use to refer to the emotion as the *result* of an appraisal process. 'Joy', for example, can be the verbal label for the 'experience of an event appraised in a joyful way' (ibid., p. 358) while it can also be '*elicited by* the appraisal of an event as beneficial and within reach' (ibid., p. 359). Frijda points out the confusion that can arise from this dual use and argues that the process of appraisal precedes the components (for example, pleasant or unpleasant feeling, action readiness, physical arousal) that determine what an emotion is called. So, he suggests that 'how events that are appraised *during* emotions appears often to result from a cognitive elaboration of the appraisal processes *eliciting* the emotion' (ibid., p. 371). This view provides support in the first instance for development of Damasio's argument that feeling is the perception of the body state and the associated thinking (see page 105). It also helps imagine more specifically how existing beliefs could become entrained in the appraisal processes as suggested by Forgas (2000b) and Frijda and Mesquita (2000).

Suppose, for example, a person, our believer, sees another, Person X, approaching them who they have come to believe is argumentative and aggressive on the basis of a description from a work colleague whom they trust. Hence, they've formed a belief that this Person X is 'trouble' and can be intimidating. Moreover Person X's physiognomy is ambiguous in that they could be interpreted as being angry, aggressive, or simply enduring

some form of chronic physical discomfort. But, our believer already has a belief that Person X is intimidatory; so when they see them approaching, the look on Person X is immediately interpreted as anger partly because this existing belief is now entrained in the appraisal, albeit unconsciously, and perhaps partly because of the types of responses that occur automatically in the 'emotional' systems in response to any (consciously or unconsciously) perceived threat (Damasio 1994; LeDoux 1993, 1998, 2000, 2012). So now our believer is experiencing some form of anxiety. But what happens if the person speaks in a friendly manner and proves to be quite affable? It is reasonable to suppose that the anticipation of the need to withstand an intimidatory attack, and the resulting action readiness, will contribute to both the perceptions of the person and the interpretation of the perceptions. At the very least our believer might become confused as a result of the cognitive dissonance occurring. They may even progress to experiencing guilt as a result of the cognitive contributions to the appraisal process: how could they have so strongly prejudged someone on the basis of hearsay?. Alternatively, their action readiness might have them respond in a manner that finally evokes an antagonistic response from Person X. In this latter instance, they will now have clear perceptual evidence supporting the belief that Person X is intimidatory and is indeed 'trouble'.

Frijda examines a number of scenarios to propose that 'even emotions like anger, guilt, and shame, that have cognitively complex definitions, can result from rather elementary stimulus constellations, and through rather elementary appraisal processes' (1993, p. 374). An additional, and important, point he makes is that people generally monitor their emotions in order to be able to regulate them and to adapt or, in some instances, justify them. In the example above, we can plausibly argue that confusion and guilt arise from such monitoring: the discrepancy between Person X's expected behaviour and the observed behaviour calls to consciousness the belief that one should not prejudge people on the basis of unsubstantiated hearsay, for example. Now the believer realises that they have violated one of their more strongly held beliefs (despite its not being in conscious awareness initially) and this gives rise to the emotion they label as guilt. Hence, we see an example of how the appraisal of events during the experience of the emotion can lead to the kind of cognitive elaboration that Frijda proposes elicits the end-result emotion. Examples such as these illustrate why Frijda and Mesquita would say:

> These beliefs are the product of appraisal processes that include activation of pre-existing beliefs relevant to the event, or create novel beliefs about its meaning.
>
> (Frijda & Mesquita 2000, p. 52)

The above example allows us to postulate that the believer could arrive at a new belief about Person X; for example, 'They're not intimidatory, they're just in some sort of chronic pain—that explains their expression'. This product is the result of the discrepancy between

X's behaviour and that anticipated, the resulting dissonance, the calling to consciousness of an existing belief about not prejudging people, and the generation of a new, more charitable belief. Of course, there is no inherent necessity for the newly created belief to be any more founded in truth than the original: they are likely to be more often developed on the basis of existing experience and other pre-existing beliefs (e.g. people in chronic pain always look grumpy).

Related to this is the proposal from Frijda and Mesquita that people project their beliefs on to situations and people. For example, if experiencing enjoyment in the presence of others, they become wonderful people; a resort in which one experiences relaxation become a 'great place to go'. So, we see again, the interpretation of the emotion while experiencing the emotion not only influencing the end-result emotion but also the role of beliefs in shaping new beliefs. People become wonderful people, for example, because there is a pre-existing belief that when people make you feel that good in their presence, they *are* wonderful people. Frijda and Mesquita describe this phenomenon thus:

> The beliefs constructed in emotions may go further than turning an emotional impression ... into the belief about a true property. Such information may serve as a starting point for elaboration by the evoked emotion.
>
> (ibid., p. 53)

Frijda and Mesquita explain this with an example of how startle can turn to anger. Beyond such developments they also suggest that beliefs can be created to add to the 'picture' (ibid., p. 53) of an emotional situation. 'Picture' is a vague term in the way they have used it; but here it is taken to mean that additional interpretation is generated (for example, 'Ah, so that expression must be from chronic pain) or even altering of representation (for example, 'Yes, I can see how those lines on his face are from pain and not anger'). Beliefs can also be used to justify emotions: for example, 'I had every reason to be anxious at first because everyone was saying he intimidates people'. This elaboration of beliefs, Frijda and Mesquita suggest, is evidence that emotions involve processes over time and also involve 'cyclical processes in which information generates emotional responses that generate new information and so on' (ibid., p. 54).

Before examining how emotions help maintain beliefs in the face of demonstrably contrary evidence, Frijda and Mesquita make some valuable observations about emotion strength, sentiment, and features of emotional beliefs. Starting with the features of emotional beliefs, they suggest that:

- they tend to involve generalisations
- they are about things that last, and
- while they are held, they are strong beliefs in that they appear to have a high

likelihood of being true or are felt to be true.

Sentiments are relevant because they involve concerns which, as discussed shortly, have a significant role in belief strength. Frijda and Mesquita see a *sentiment* as 'an appraisal structure that includes concern relevance of its object [and] thus are dispositional emotions' (ibid., p. 55). They suggest that temporary beliefs involved in emotions can become generalised long-term beliefs and when this happens, the emotion becomes a sentiment. More straightforwardly, they characterise sentiments as attitudes towards what may affect or impinge on one's concerns.

In examining how emotions relate to belief strength, Frijda and Mesquita (ibid.) propose the following factors are involved:

- concern relevance
- emotional intensity
- the intervention of virtual emotions.

Frijda and Mesquita acknowledge that 'concern' is a broad concept. It includes 'motives, major goals, personal sensitivities, attachments, and supra-personal values' (ibid., p. 61). When one perceives uncertainty about being able to protect or to satisfy one's concern, then emotions arise.

Emotional intensity is an aspect of belief strength according to Frijda and Mesquita. It, in turn, is considered a function of elements such as the strength of concerns, the number of concerns, the urgency for action, perceived possibility of coping, and the attribution of causal agency.

When beliefs belong to what Frijda and Mesquita consider sentiments, they suggest the emotion anticipations are likely to be instrumental in maintaining or strengthening those beliefs. They define emotion anticipations as follows:

> Emotion anticipations have the structure of emotions. They consist of appraisals of the anticipated situation as well as anticipations of the ensuing affect and action readiness or, in other conditions, of imaginings of how someone else may appraise, feel, or become ready for … They are felt or imagined, rather than thought. It is likely that this involves low-level activation of some of the emotional response components that come to the fore more manifestly in actual emotions.
>
> (ibid., p. 58)

In this description, there appears to be strong alignment with the notion of the 'as if body loop' that Damasio proposes (1994, 1999, 2003b). This is a subject of further investigation in Chapter 4.

Frijda and Mesquita propose that it may be this emotional anticipation that gives beliefs great strength when those beliefs are challenged. Because beliefs guide judgments, decisions, and action, they suggest that challenging one belief can challenge a whole belief system. Changing this one belief could require changing the whole system. In effect, it is the person's world view that is being challenged and that equates to challenging their sense of security.[47]

However, as with others in previous discussions, Frijda and Mesquita so far have managed to demonstrate *that* emotions, sentiments and emotion anticipation influence belief, but not yet specifically indicated *how* this occurs. In attempting to do this, they point out that emotions influence thinking in general, not just beliefs. They propose that four features of 'emotion-steered thinking' (Frijda & Mesquita 2000, p. 64) are involved: instrumentality, motivational force, control of the scope of thought, and motivated bias.

With instrumentality, the suggestion is that the thinking is aimed at helping achieve our emotional goals. In doing so, this thinking will distribute attention such as being sensitised to danger signals when frightened; examine options for action including assessing one's sense of being able to cope with the situation; and, explore the nature of the emotional situation which can include anticipating the affect resulting from various courses of action. Such thinking, Frijda and Mesquita suggest, may create beliefs that help make a situation seem more intelligible; may simply make one feel better by, for example, justifying one's own behaviour in a situation; or may strengthen or change a belief to eliminate or reduce cognitive dissonance.

Motivational force is predominantly about eliminating discomfort or achieving pleasure and harmony. Frijda and Mesquita propose that this greatly helps in explaining resistance to changes. If a belief is challenged, the thought of abandoning it may be assessed as too risky, especially if such abandoning undermines the readiness to act. Further, as noted earlier, challenging a belief is rarely a comfortable experience for most people in day-to-day life; and, it is becoming clearer from the literature that feelings play a most significant role.

As shown in earlier sections, emotional arousal affects perceptions and attention. The proposal from Frijda and Mesquita is that, when emotional, a person focuses on information that is directly relevant and has neither the motivational nor the attentional resources to explore beyond that which seems most relevant. And the key point with motivational bias is that a biased view may simply refuse to countenance an alternative view. A throw-away phrase from Frijda and Mesquita about the person possibly being not interested in alternative views, may point to a much more profound matter. The question

47. This is a major focus in Chapter 4.

can be asked as to why the person is not interested in entertaining alternative views and possible answers range from ones that refer to self perceptions ('I don't engage in illogical nonsense') to perceived pragmatism ('I've always been right about this in the past, I don't see how this time is any different ... it's bound to work'). These possible answers take us back to the acknowledgement of having to change a whole belief system or world view. To the extent that this becomes a challenge to the security of one's sense of self,[48] it is easy to accept that resistance to changing a belief would be a natural line of defence.

In a section of their paper titled *The process of belief formation*, Frijda and Mesquita manage to present additional influences that are instrumental in creating beliefs. Starting with the definition that a belief is a 'proposition considered to be true', they make a link to the discussion in Chapter 2 in stating that 'truth judgments are largely based upon probability, credibility, and plausibility estimates' (ibid., p. 68). While this may be the case, it demonstrates the complex nature of influences on emotion and belief because judgments and estimates of probability fall foul of perceptual influences, attentional influences, existing knowledge and memories. And, as we have seen amply, these are under the influences of emotion. Once again, as with many others, we are presented with a catalogue of factors that influence emotion and consequently beliefs: viz. apparent reality; emotional meaning; time discounting in which events that are remote in time have little emotional impact; confirmation bias; goal relevance. Frijda and Mesquita make an important point that also supports the earlier suggestion about judgments being *believed* if they fit expectations (see page 75):

> A belief is plausible when it agrees with mental models and mechanisms, and with expectations.
>
> (ibid., p. 69)

Frijda and Mesquita also refer to *disbelieving*. They use this in the sense that one might not believe what another says because one considers them untrustworthy, for example.

So Frijda and Mesquita largely leave us with plausible high-level mechanisms by which emotions can create, maintain or change beliefs. But their proposals are still closer to the *that* end of the explanatory spectrum than the real *how* end.

Their discussion is highly valuable nevertheless as they have:

- given a plausible account of why beliefs can be considered an integral part of the emotional experience

48. An investigation taken up in Chapter 4.

- proposed features of emotional beliefs (for example, that they are usually generalisations)
- highlighted the importance of emotional strength
- proposed four factors contributing to emotional thinking.

This takes us several steps further than, for example, Forgas's (1995, 2000b, 2006) acknowledgement that pre-existing beliefs are engaged in elaborate strategies needed for making judgments. There is potential, too, for Frijda and Mesquita's work to suggest a next level of detail in the type of hierarchy that Forgas (1995) proposes. A final point from Frijda and Mesquita's conclusion should be noted as it, too, points to a later, important discussion:

> At the centre of these emotional influences on beliefs is concern relevance.
>
> (Frijda & Mesquita 2000, p. 73)

It is proposed here that recognising something that is of concern to oneself will necessarily entail feelings in the sense that Damasio sees them—as the conscious experience and awareness of emotion in Damasio's terms (that is, the body state). It is easy to acknowledge this with the feelings that arise from homeostatic imbalances—for example, hunger, or being too cold or too hot. The more important issue is how these feelings help identify matters of great concern such as one's sense of security and, in so doing, engage existing belief systems and emotions to the end that irrational beliefs are created, maintained or even strengthened.

3.4.3 Feelings as Feedback

Clore and Gasper's paper (2000) is second in the collection (Frijda *et al.* 2000a) introduced earlier; but, a choice has been made to examine it last (of that collection) because of its focus on the role of feelings. Their focus is on the role of feelings in providing information and guiding attention.

The role of *emotions* in influencing attention and perception has been examined earlier in Section 3.3.1. Clore and Gasper's focus on *feeling* once again raises the issue around definitions of emotion and feeling. The proposal here is that Damasio's way of defining emotion and feeling is the most useful in clarifying how these two phenomena can influence beliefs as well as potentially providing a clearer understanding of appraisal processes. Clore and Gasper's work, in providing a focus on feeling, develops that line of inquiry. They note:

> There is insufficient research on emotions and belief to derive well-supported principles [on how emotions affect beliefs].[49]
>
> (Clore & Gasper 2000, p. 10)

Nevertheless, they propose generalised hypotheses about how emotion influences the creating and maintaining of beliefs via effects on attention and information processing. In examining affective feelings and *mood*, they propose that feelings 'serve an important feedback function' (ibid., p. 10). Because they see mood as *not* having a particular object of focus, they suggest that in the case of general moods, 'the object about which feelings appear to give feedback depends in part on what is most salient to the experiencer at the time' (ibid., pp. 10–11). Moods, then, allow attributions to be manipulated making the study of the influence of affect easier.

They suggest seven principles by which feelings of moods affect beliefs.

1) **The experience principle**

   The feelings that are experienced with a mood or emotion lead to cognitive consequences. They argue that even in the case of unconsciously experienced affect, there are cognitive effects even if feelings are not involved. From a Damasian[50] view, this would be attributed to the somatic state considered to be the *emotion*, but not to *feeling* as that requires the perception of the body state and related thinking.

   Nevertheless, Clore and Gasper's point is still important in highlighting that physiological changes—albeit not consciously detected—can mediate cognitive processes. This is to be expected on the basis of other models examined (e.g. ibid.; Forgas 1995; Frijda & Mesquita 2000; Scherer 2004).

2) **The information principle**

   Here, Clore and Gasper refer to 'emotional feelings' and in so doing seem to align with the Damasian view. These feelings are consciously experienced and considered to provide 'experiential feedback from appraisal processes that are largely nonconscious' (2000, p. 14). Hence, they indicate how significant the situation is to the person. This significance is supported by Damasio's work (1994) showing that the lack of affective feedback rendered brain-damaged patients unable to make everyday decisions.

49. Later writing by Clore and Ortony (2008) indicates this is still the case.

50. If I may coin an adjective.

3) **The attribution principle**

The information that a person derives from feelings depends on how they attribute the experience of affect. A typical experiment cited shows that people's beliefs about their satisfaction with life are affected by mood that they are experiencing as a result of the weather. However, if their attention is drawn to the weather before being asked about their satisfaction in life, they attribute their mood to the weather. While the mood remains, they no longer make a judgment on their life satisfaction based on their mood.

The problem with attribution is that it is 'usually completely implicit, rather than being explicit, deliberative, or effortful' (Clore & Gasper 2000, p. 16). We can see here that unless there is a conscious element in cognitive deliberation to correctly attribute the stimuli inducing the mood, the mood could easily be misattributed and, from that, (false) beliefs created. For example, a bad weather mood causes a person to attribute their mood to a brusque conversation with a work colleague (who is time-pressured and anxious about a deadline) and consequently decide that the colleague is 'not a nice person'.

4) **The immediacy principle**

Feelings provide immediate feedback on the current body state. Consequently, they are generally experienced as 'reactions to current mental content' (ibid., p. 16). The contrast is drawn between pain sensations which provide a signal of injury and the affective feelings that result from psychological causes. Like the pain that is related to the immediate body state, affective feelings tend to be attributed to the current conditions. However, they may be the result of a bad mood that is being maintained by the rumination over an argument that occurred some days previously. Clore and Gasper note that:

> Mood and mood-like states, such as depression, are notorious for colouring people's beliefs about whatever they focus on. We propose that this is a natural outcome of this more or less automatic tendency for attributions to be guided by the immediacy principle.
>
> (ibid., p. 16)

5) **The attributional constraint principle**

Clore and Gasper consider the specificity of the object of the emotion or mood to be a key element for distinguishing between these phenomena. While moods tend to have no particular focus or perhaps a broad focus (e.g. 'It's just everything that's going on.'), emotions are more likely to have a specific object. In this case, where there is an

object that is related to a belief, the feelings are likely to be attributed to that object. Clore and Gasper's example is that of attributing feelings of depression to perceived personal shortcomings. With such an attribution it could be easy, for example, to strengthen a belief about that the chances of finding employment are minimal.

6) **The elaboration principle**

Clore and Gasper draw on research indicating that a person's conception of an object—their related schema, or cognitive organisation—affects the potential for elaborating the meaning of the feelings being experienced. Where there are many related and interconnected aspects, the feelings can affect the whole cognitive structure. An example offered is that of a person's self-concept, an aspect that emerges as being particularly significant (see Chapter 4).

In addition to the nature of the object, Clore and Gasper suggest that the influence of affect depends 'also on how one parses or punctuates [an] object domain' (ibid., p. 20). They provide the example of a student attributing a poor exam grade to a lack of effort; but the lack of effort can be linked to fatigue, or laziness, or interference from family members. We can see in this example how the engaging of a network of related aspects and beliefs could generate beliefs that extend beyond the immediate object. For example: 'If only my parents didn't make me pay board, then I wouldn't have to work and would've been able to study properly for the exam—they're so selfish and uncaring'.

7) **The processing principle**

When someone is performing a task or solving a problem, they may experience their feelings as feedback about their performance or about the value of accessible information (ibid.). Clore and Gasper predict from this principle that positive moods should lead to reliance on available information (habits, beliefs, expectations) while negative moods should lead to a focus on new information. They cite a number of studies that support this view: for example, voters in positive moods relying more on party identification while those in negative moods more likely relying on what the candidates say (Markus & Mackuen 1993).

An interesting finding is that anger, like happiness, can lead a person to validate their own beliefs and perspective. Clore and Gasper conclude that the importance lies not in the feelings themselves, 'but in the information they convey' (Clore & Gasper 2000, p. 22).

As with the feelings of mood, Clore and Gasper take up the view of emotions as information. They suggest that confronted with indirectly presented knowledge (propositional knowledge) or direct experience (experiential knowledge), the experiential knowledge will often have priority.

In presenting their *feelings-as-evidence hypothesis*, they suggest that 'people need to *feel* that the case against their position is compelling before they change their minds' (ibid., p. 25, emphasis added). They liken this to a person adjusting their belief to account for external, perceptual evidence, except in this case they are using feelings as internal evidence.

They view emotions as having more 'cognitive structure' than moods, and this would appear to put them in agreement with the idea that beliefs are part of emotions (Frijda & Mesquita 2000) and that they are necessarily entailed in appraisal processes as described earlier (Fiedler & Bless 2000; Forgas 1995, 2000a). Because of this, Clore and Gasper propose the 'information value of emotions is more constrained than that of moods' (2000, p. 26). For example, they can indicate whether an outcome is desirable or not, relative to a given goal; they can indicate whether a person's actions meet desired standards; or, they can indicate if some object or circumstance is appealing or not.

In their proposed *emotion categorisation hypothesis*, Clore and Gasper suggest that because of the reliance on internal evidence from emotions, similar emotional reactions arising from different situations can cause a person to include those situations into a single belief system.

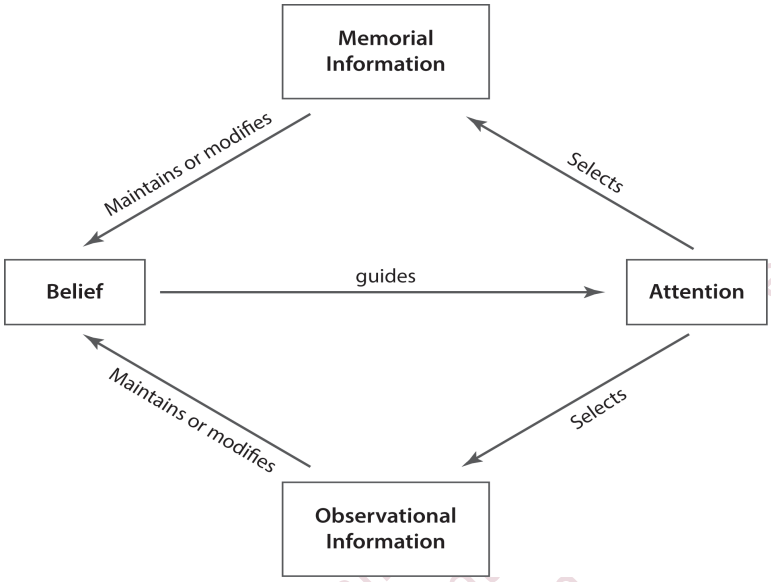
> In other words, to the extent that subjective similarity is taken to imply objective similarity, then beliefs relevant to one situation may be applied to another situation that elicits a similar emotion.
>
> (ibid., p. 28)

Such conflation of beliefs can have emotional consequences which Clore and Gasper address in their *emotion misattribution hypothesis*. In this, they suggest that if multiple situations are categorised together then it becomes difficult to accurately ascribe the emotions experienced to any one of them. Consequently, the beliefs that might be separately relevant to the different situations can also become transferrable from one to the other. They provide an extended example of an academic whom they name Professor Bitterman who conflates his emotional response to the waning popularity of his area of specialty with his emotions relating to feeling personally under-appreciated. The resulting anger and resentment had him 'entertain an image of himself as valiant figure engaged in a high-minded, if doomed struggle in defence of [his area of specialty]' (ibid., p. 28).

As with others (see Section 3.3.1), Clore and Gasper note that emotions guide attention, governing the breadth or narrowness of focus. They suggest that this is not so much from

the feelings, but from what the feelings signify. The influence of emotion on attention and perception needs little additional coverage here[51]; however, they do offer a model that helps elucidate likely interactions between attention, memories, beliefs, and observations.

Figure 9: Alternative Perceptual Cycles



(ibid., p. 32)

Feelings become relevant in this scheme, Clore and Gasper suggest, because they can serve as *evidence* that leads to commitment to a belief or is experienced as a commitment. By this mechanism, belief can guide attention. As they put it:

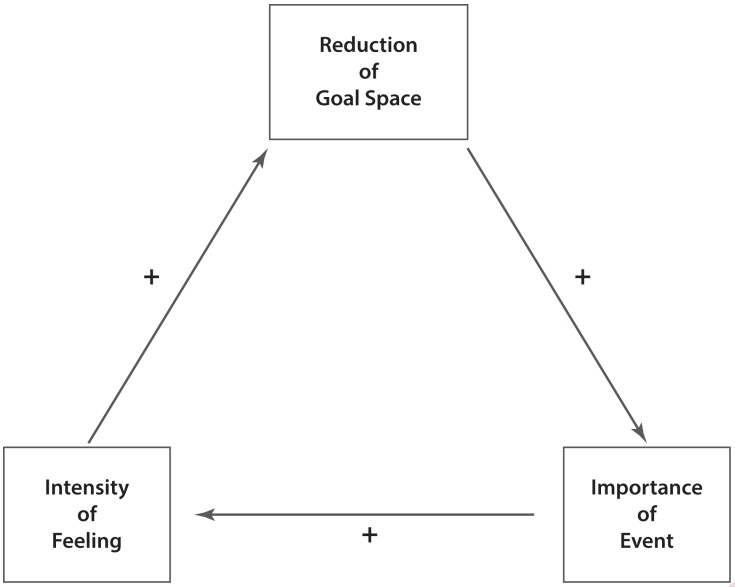
> … the emotional person acts like a prosecutor or a defence lawyer seeking by any means to find evidence for the belief.

(ibid., p. 33)

Picking up a proposal from Simon (1967) that a primary function of emotion is to ensure that urgent matters are those that are attended to first, Clore and Gasper suggest that the intensity of an emotion creates an 'attentional funnel'. As emotional events take precedence, there is more focus attention on the relevant goal, or goals, which in turn increases the perceived importance of goal-relevant events. Since there is a positive feedback loop, this increased importance increases the intensity of the emotion.

51. Having been covered in Section 3.3.1

Figure 10: Attentional Funnel



(Clore & Gasper 2000, p. 35)

Clore and Gasper cite evidence from studies of sports fans that support this view. More generally, it is easy to propose everyday situations that fit with this model. For example, the person leaving for an important meeting cannot find their house keys; as their anxiety about being late for the meeting grows, their search for the keys becomes ever more feverish; and, as time passes the importance of the event is emphasised and grows in perception, producing greater anxiety, and even more frantic searching.

They note that the study with sports fans showed that the success of their teams had little relative importance when compared to their life overall.[52] What this highlights in Clore and Gasper's *goal importance hypothesis* is that what they refer to as the 'momentary importance' (ibid., p. 36) of events and emotions related to goals will narrow the focus to the goal deemed more important. It may be better to consider momentary importance as *contextual* importance, at least where a change of context is clearly identifiable: for example, a sports game can seem intensely important while there, but relatively unimportant three days later. In their conclusion, Clore and Gasper highlight an important point in language that makes the role of feelings more obvious.

> One can argue with logic, but not with feeling.
>
> (ibid., p. 39)

52. However, it would be interesting to research if similar results vary across cultures. The studies were US-based. This writer suspects some European, English and Australian sports fans may invest rather more of their life in sport.

This emphasises the primacy of feelings as directly experienced evidence. They continue to suggest that for a given emotion type there is an inherent model of the situation. Fear, for example, involves the assumption that there is a threat being posed. That ties in well with the previously discussed work of Fiedler and Bless (2000) and Forgas (1995, 2000a), in which they say:

> Thus, one of the most obvious and important implications of emotion for belief is that the occurrence of an emotion already means that the emotional person is committed to particular beliefs about that situation.
>
> (Clore & Gasper 2000, p. 40)

3.5 The Case for Damasio's Definitions

There is clearly agreement between researchers that emotions influence beliefs and even lead to the creation of beliefs. As demonstrated, the literature reviewed largely falls to saying *that* emotions affect beliefs. The models for appraisal and the various principles proposed do provide valuable directions for inquiry and plausible high-level processes. Nevertheless, attempts to explain *how* emotions influence beliefs remain somewhat nebulous. This is exacerbated by the general lack of agreement over relevant terms including mood, emotion, feeling—a point noted by LeDoux (2012) recently.

Thus far, the definitions of emotion and feeling from Damasio have been offered as those most promising for investigating the actual mechanisms by which emotions influence beliefs. While they do not have universal acceptance, they do have a strong following that appears over time to be increasing in the literature; and, they do have a significant influence among those researchers in this area. Frijda et al. (2004, p. 458) observe that researchers 'are nearly unanimous in distinguishing emotions from feelings'.

The area of appraisal theory is too large to cover in detail in this thesis; however, the appraisal approach has shown itself to be unavoidable in this inquiry. The few, key models presented here are sufficient to highlight a major challenge with this approach. Taking just the model from Scherer (2004) (Figure 7 on page 109), he shows an arrow going from 'Event' to 'Relevance', followed by a complicated network of signalling between various appraisal factors. It is possible to imagine that some of those arrows carry representations or language; but the unavoidable question is 'How, at any stage, does the person know that something is of significance?'. Scherer partly addresses that with the model shown in Figure 6 in which feeling is shown as having a monitoring role.

It is argued here that the value of Damasio's definitions is that they provide a very clear distinction between the phenomena of emotion and feeling. In doing so they provide an understanding, for example, of how a person could have a physiological response to an event even though it is not consciously noticed. So, given a choice in a gambling game

they might provide a story about their cogitations that led to their choice, but the researcher can identify clear signals that can be influencing that choice.

The case for feeling$_D$ grows with the support of work discussed here (notably, Clore & Gasper 2000; Scherer 2004). Once a person is aware of their body state (emotion$_D$), they are experiencing feeling in Damasian terms; and not just the feeling but the strength of it. Such a view of feeling fits well with the earlier arguments that feelings are feedback about core concerns (Frijda & Mesquita 2000). Again, Damasio's view provides greater focus for modelling and for research and it also provides a model that is far more testable (that is, falsifiable) than others. His link to feelings being an awareness of the body state highlights the importance of including the body in appraisal approaches. This theme is taken up by Northoff (Northoff 2008) in his investigation of emotions:

> The 'lived body' must thus be considered an integral and central part of appraisal rather than remaining either [sic] epiphenomenal as a mere byproduct.
>
> (ibid., p. 70)

> Feeling and appraisal must then be considered integral aspects (rather than components) of emotions which are intrinsically linked to each other.
>
> (ibid., p. 89)

Frijda et al. (2004, p. 458) claim that 'feeling has regained respectability'. Indeed, it is difficult to imagine how it would be possible to fully investigate the effect of emotion on belief without the acknowledgement of the importance of the body state and a conceptual separation of emotion and feeling that permits more refined and testable research.

Damasio's view of feeling also extends more recently (and perhaps controversially) to the concept of the self (1999, 2010). Even if this examination of feelings and the self is left aside, the role of feelings (Damasian) in influencing autobiographical beliefs and memories takes this inquiry to one of its most important aspects which is a major theme in Chapter 4.

3.6 Summary

The chapter opened by recording that there is general agreement that emotions affect beliefs (and vice versa), and emotions also affect memories (and vice versa). A quote from Frijda et al. (Frijda *et al.* 2000b) pointed out that the resistance of beliefs to being changed indicated the need to account for the influence of emotions.

The strong interrelationship between emotions and beliefs was established and represented in the simple but significant Figure 1.

Extensive examination of the range of definitions of *emotion* and *feeling* followed. A number of factors are agreed on by researchers. For example, emotions involve physiological arousal, cognitive activity, and action tendency. However, while there is agreement on a number of factors, there is no one, clear definition of emotion or of feeling on which the researchers agree. In colloquial terms, one could describe the definitional aspects of this field as being messy.

Hence, the definitions provided by Damasio were offered as being worthy not only of consideration but also of employment. While they do not have universal acceptance; they do have broad acceptance by researchers in the field of emotions and they have been, and continue to be, highly influential. More importantly, in separating emotion and feeling distinctly, it is argued that this provides a way forward for more focused research and modelling. Additionally, the casting of emotion as the body state and feeling as the representation of that body state (and the accompanying awareness of it) allows Damasio's hypothesis to be tested in ways that theories employing vague terminology can not be.

One of the challenges in this field is the proliferation of models from appraisal theorists. This challenge is all the greater because of the lack of agreement on precise definitions for mood, emotion and feeling. It was noted, however, that there is almost universal agreement that emotion and feeling are separate phenomena.

To get a better insight on how Damasio's hypothesis can assist this inquiry, the research on the effects of emotion on attention, perception and memory was examined. Since the study of emotion has become respectable again, there has been an almost exponential growth in emotion-related research. Despite this, only one major *collection* on the topic of emotions and their effects on belief appears to be available. There is a strong reliance on this collection for the argument being developed in this thesis. It is felt that this is justified as the collection reviews the significant literature that preceded it and subsequent literature (reviewed to date) has developed the lines of inquiry without any significant refutations of the earlier findings.

In short, there is agreement that emotion affects attention and perception. Whether the emotion is positive or negative does matter, with a broad pattern that positive emotion allows the individual to utilise peripheral information more than they would if experiencing negative emotion, for example. Similarly, memory is affected by emotions: for example, emotional events tend to be remembered well and for longer. One of the challenges posed by one researcher relates to how events are misremembered and it was noted that there is very little research on this matter. A key suggestion was that positive and negative emotions engage different forms of information processing and thereby can produce differences in memories.

This was followed with an examination of key literature relating to the effects of emotion on belief specifically. What was established is that there is agreement that beliefs are involved in processes occurring when a person experiences an emotion. Fiedler and Bless (2000) supported an argument for assimilative and accommodative processes and proposed that positive affect increases the reliance on existing belief systems while negative affect can lead to updates of beliefs as a result of significant new information. Of course, this latter effect does not always occur: as cognitive dissonance studies indicate, negative affect can result in even stronger entrenchment in an existing belief. Then, Forgas's (1995) more extensive Affect Infusion Model was examined. The principal proposal in this is that as the cognitive strategies required for a judgment become more elaborate, affect has more influence on beliefs.

Other researchers (Clore & Gasper 2000; Frijda & Mesquita 2000) provided a strong case for including beliefs as an integral part of experiencing and appraising emotion. Frijda and Mesquita made the significant observation that there is an unfortunate ambiguity with the term 'appraisal' as it refers to both the content of the emotional experience and the emotion that is the result of an appraisal process. As well, they highlighted the significance of beliefs affecting the appraisal.

Finally, there was a focus on the significance of feelings and the acknowledgement that as feedback about the significance of events and related emotions they are a principal aspect of emotional experience and of an emotion's influence on belief. This returned the inquiry to the Damasian view of feelings. Since feelings can be considered as the conscious awareness of emotion, it seems that they cannot avoid engaging cognitive elements including memories and beliefs. This opens the way for investigating the effect of feelings on autobiographical memories and beliefs—and the related inquiries—in Chapter 4.

Chapter 4: Beliefs Sustain the Self

> Frank: "Ah! The intensity of your feelings is *not* the test of reality!"
>
> (Bandler & Grinder 1982, p. 35)

The findings of Lepper et al. (1986) point directly to the type of belief that a person may hold of themself which is:

- either irrational in that it is refuted or called into question by clear evidence; or, possible but implausible because there are strong arguments that call the belief into question; and
- actually or potentially limiting to the person's self efficacy in attaining desired and worthwhile goals.

In their experiment where some students had been deliberately given poor instruction to set them up for failure, the subsequence information-discounting procedures proved to be 'utterly ineffective and perhaps counter-productive' (ibid., p. 489). They note that additional efforts to help students understand what influenced their results 'may have heightened rather than attenuated subjects' feelings of success or failure' (ibid., p489). The reference to *feelings of success or failure* is seemingly incidental to the presentation of the results; yet, given the findings discussed in Chapter 3, the phrase, it seems, is pointing to the central element in forming such beliefs. This is highlighted in the following:

> Simply demonstrating to a child, even in a clear and concrete fashion, that his or her poor performance may well have been the consequence of an inept or biased teacher, a substandard school, or even prior social, cultural, or economic disadvantages may have little impact on his or her feelings of personal competence or potential.
>
> (ibid., p. 490)

Similar findings are reported by Forgas (2013) in which students 'in a negative mood blamed themselves more when failing, and took less credit for their successes' while 'those in a positive mood claimed credit for success but refused to accept responsibility for their failures' (2002, 2013, p. 7). Forgas notes that beliefs 'about the self represent a particularly complex, elaborate and problematic domain' and that they 'are also strongly influenced by affect, as positive affect improves, and negative affect impairs the valence of self-beliefs' (2013, p. 7). He suggests, further, that affect may be a 'key organising principle of self-related beliefs' (ibid., p. 7).

The quote from Lepper (1986), above, draws attention, once again, to the word *feeling*. The context and usage here suggests that 'feeling' is being used in a general sense and

possibly synonymously with emotion or affect. However, the pervasiveness of the use of 'feeling' in this sense in everyday language combined with recent attention to both defining 'feelings' and examining the influence of them points to it being worthwhile undertaking a closer examination of how 'feelings' may help form and consolidate beliefs (Bower 1981, 1991; Cromby 2007, 2012; Damasio 1994, 1999, 2003b, 2004; Forgas 2006; Frijda *et al.* 2000a; Harmon-Jones & Mills 1999; Harmon-Jones 2000; Manstead *et al.* 2004; Oatley 2000; Sedikides 1995).

The quotes above and the findings such as those from Forgas (2002, 2013) echo, perhaps with amplification, the previous quote from Clore and Gasper (2000, p. 39) in Section 3.4.3: 'One can argue with logic, but not with feeling'. Since the earlier work cited in Chapter 3, continuing research confirms the influence of emotion, and therefore feeling, on perception, decision making, judgment, creativity, memory, and beliefs (for example: Clore & Huntsinger 2007; Forgas 2001, 2006; Gasper 2004; Gasper & Clore 2002; Storbeck & Clore 2008). In particular the case for feeling-as-information or feelings-as-evidence appears to have strengthened with continuing research (Clore *et al.* 2001; Clore & Storbeck 2006).

At this stage, the conclusion could be drawn that emotion and feeling have been shown to influence processes from perception through to cognitive processes of judgment, and by implication, belief; and, given these influences, emotion and feeling lead to the establishment and maintenance of beliefs. From that, it could be relatively straightforward to draw on research on the effect of positive and negative affect on memory and self-concept and swiftly propose a generalised system for creating false and, in some instances, potentially debilitating autobiographical beliefs or beliefs about the context enclosing the self (e.g. 'the world will end tomorrow'). However, a closer look at Clore and Gasper's (2000) feelings-as-evidence hypothesis through the lens provided by Damasio's view of feelings promises to give greater insight into how feelings and cognitive appraisals may combine to construct and maintain false beliefs and memories.

4.1 Feeling Really is Believing; but Why?

In proposing their feelings-as-evidence hypothesis, Clore and Gasper (ibid.) remind us that experiental knowledge often takes priority over propositional knowledge. As they put it:

> To disbelieve what we experience is uncommon.
>
> (ibid., p. 25)

In an interesting view of 'disbelieving' or, more specifically, 'unbelieving' (Gilbert *et al.* 1990, 1993; Gilbert 1991), it is suggested that before a judgment or belief can be rejected

as false, it must first be accepted. Gilbert (1991, p. 161) proposes that people 'faced with shortages of time, energy, or conclusive evidence, may fail to unaccept the ideas that they involuntarily accept during comprehension'. This view is strengthened when considered in the light Kahneman and Tversky's (and related) research as discussed in Chapter 2. With respect to perceptions, Gilbert suggests: 'As perception construes objects, so cognition construes ideas' (ibid.). If we recall that included in those perceptions are the *feelings* that relate to assessing a judgment or belief as true or false, we can see a possible contributing reason for why disbelieving what we experience is uncommon, as suggested by Clore and Gasper.

Clore and Gasper's hypothesis suggests that just as beliefs may be adjusted on the basis of *external evidence*, the sensations of feelings can be treated similarly to external evidence. The result is that something that is believed (propositionally) *and* felt emotionally 'may seem especially valid' (Clore & Gasper 2000, p. 25). Using Damasio's definition of feeling, it can also be proposed that the *conscious awareness* required by his definition will engage some degree of cognitive appraisal and, especially, attribution of cause and making of meaning.

Clore and Gasper (ibid.) do suggest that it is not only that feelings often take priority over propositions; the *information value* of feelings is also an important influence on whether a belief is strengthened. They suggest that emotions have 'inherent cognitive structure' (ibid., p. 26) as well as a specificity by virtue of having an object; hence, the emotion is a reaction to the model of the situation. Referring to the work of Ortony *et al.* (1988), they see three sources of value (goals, standards, and attitudes) that can generate specific, differentiated emotions depending on the nature of the affective reaction. Thus, the feelings associated with the emotion provide information about desirability, approval or liking and, for example, the relevance of the affect to oneself and/or others.

The role of feelings in guiding attention has already been introduced in looking at Clore and Gasper's (2000) 'attentional funnel' (see p. 140). However, the significant suggestion is that attention is drawn to what the feelings *signify*. Feelings have their impact on attention, it is suggested, by giving 'greater weight' (ibid., p. 32) to a belief (or beliefs) involved in the perceptual cycle (see p. 139). Clore and Gasper suggest that this occurs whether feelings are experienced as commitment or whether they indirectly lead to commitment to a belief. One potential problem that Clore and Gasper suggest is the assumption that there 'should be a strong relationship between the intensity of emotion and the perceived importance of relevant goals' (ibid., p. 36), and research they cite shows that this is not always the case. So, instead of intensity of emotion and feelings being determined by the *absolute* importance of a goal, it may be the *momentary* importance of a goal that drives 'emotional potency'. This accords with Simon's (1967) suggestion that

emotion is an 'interruption mechanism' (ibid., p. 39) that allows realtime response to the most urgent needs.

Linville (1985) produced an influential study examining a model relating the complexity of one's self-representation to emotional and evaluative response to situations. The less complex the self-representation, the more the affect and self-appraisal 'swings' following an experience of failure or success. While there is some question remaining about the buffering effect of self-complexity (Solomon & Haaga 2003), the study is suggestive of an important link between emotion and self-concept and self-representation. Clore and Gasper (2000) draw on this work to suggest that:

> … being emotional creates a sort of simple-mindedness in the sense that as emotions focus attention on relevant goals and information, the accessibility of competing goals and concerns is reduced, simplifying current mental structure.
>
> (ibid., p. 37)

Given that they consider that an emotion carries a model of the situation, and that the emotional response indicates commitment to beliefs about the situation triggering the emotion, the following conclusion with respect to the role of feelings is highly plausible:

> We assume that perception involves a constant interplay of top-down (belief-driven) and bottom-up (data driven) processing in which beliefs about the current situation guide the sampling of available information, which in turn may modify beliefs. If feelings provide their own apparent validation for beliefs, selective data-sampling and resistance to bottom-up modification of the belief may lead to inappropriate belief maintenance.
>
> (ibid., p. 40)

With the invocation of self-concept, self-representation and self-complexity, a reasonable continuation of this enquiry involves looking more closely at Damasio's view of feelings as necessarily engaging awareness and how the awareness of such 'evidence' as feelings provide affects cognitive appraisal, especially with respect to the autobiographical memories and beliefs that serve to maintain that construct of the 'self'.

From important initial writing, Damasio (for example, Damasio 1994, 1999, 2000a) has separated *emotion* from *feelings*. A principal benefit of his definitions (outlined in Chapter 3) derives not only from their having a neurophysiological basis but especially from the ability to devise and test hypotheses. The case for these definitions has been argued in Section 3.5. However, it is suggested here that this 'clean slate' approach is particularly helpful in providing definitions that could be considered as coming from first principles. Damasio's views still do not have full acceptance; however, it has to be asked whether some of the detracting comments arise because his definitions—and, more particularly, his

process for defining 'emotion' and 'feeling'—are not really understood. For example, Ratcliffe (2005), who seems broadly supportive of Damasio's views, still manages to suggest that Damasio:

> … fails to distinguish between a feeling as something that is perceived and a feeling as something that structures the perception of something else …
>
> (ibid., p. 53)

While Damasio might not draw out this distinction as overtly and clearly as we might desire, there is clearly the suggestion that feelings, being the *conscious* awareness of body state *and* the mode of thinking, will engage cognitive appraisal processes that can direct attention. Such direction of attention in an emotional state affects perception as already shown (Section 3.3). Ratcliffe does grant, however, that:

> … Damasio still provides a neurophysiological theory of how certain bodily feelings act as a structuring background for experience and thought.
>
> (ibid., p. 53)

This example is provided as one of several that appear to be attempting to understand Damasio's view of emotions and feelings in terms of existing definitions or, at least to some extent, with fragments of existing definitions (for example, Frijda 2005). From this point forward, the approach taken is to accept Damasio's definitions and examine the additional insight they can offer into *how* feelings in particular become a major contributor to the process of creating and maintaining beliefs and how false and irrational beliefs remain so resistant to change.

4.1.1 Information and Damasian Feelings

In more recent work, Damasio (2010) re-emphasises the difference between emotion_D and feeling_D:

> It makes no difference what words we choose to refer to these distinct processes, provided we acknowledge that the essence of emotion and the essence of feeling *are* different.
>
> (ibid., p. 109)

In employing Damasio's definitions outlined in Section 3.2.3, and working with his more recent development, it is important to recall that he bases his definitions on neurophysiological phenomena. As suggested above, this approach is probably best considered as a 'first principles' approach; and its advantage is that it permits distinctions for both discussion and research that are not provided by some of the poorly defined

terminology that is found in much of the literature. More recent descriptions of emotion$_D$ and feeling$_D$ that capture all the aspects of emotion outlined in Section 3.2.3 are:

> Emotions are complex, largely automated programs of *actions* concocted by evolution. The actions are complemented by a *cognitive* programs that include certain ideas and modes of cognition, but the world of emotions is largely one of actions carried out in our bodies, from facial expressions and posture to changes in viscera and internal milieu.
>
> Feelings of emotion, on the other hand, are composite *perceptions* of what happens in our body and mind when we are emoting.
>
> (ibid.. p. 109)

These feelings, according to Damasio, privilege interoception over exteroception. Damasio also emphasises the difference between feelings$_D$ and feelings and thoughts with a *theme* that may fit with a feeling label:

> … feelings are *functionally distinctive* because their essence consists of the thoughts that *represent the body involved in a reactive process*. Remove that essence and the notion of feeling vanishes.
>
> (Damasio 2003b, p. 86, emphasis added)

The implications of this view may be best encapsulated when he says:

> Feelings open the door for some measure of wilful control of the automated emotions.
>
> (ibid., p. 80)

And, the reason that feelings$_D$ can permit wilful control of emotions is that, in Damasian terms, they only become feelings$_D$ when we are *aware* of them. It is the conscious awareness that defines these perceptions as feelings$_D$.

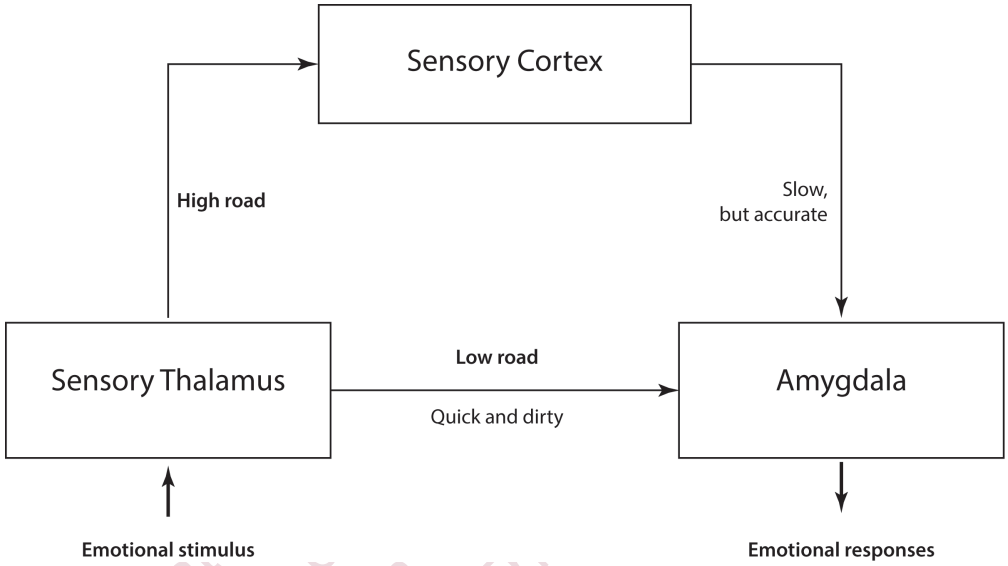
4.1.1.1 A First Level of Information

In keeping with Damasio's model for emotions (Section 3.2.3.1), it is relatively straightforward to appreciate how feelings$_D$ initiate a response aimed at re-establishing homeostatic balance. For example, the awareness of feeling thirsty prompts us to drink; the awareness of hunger prompts us to eat. Similarly, if we feel pain—caused by an external source such as a sharp object—there is an impulse to move away from the source of the pain.

The inevitable question now arises about the timing of when we become conscious of a feeling and how thoughts arising from and associated with that awareness begin to interact in the emotional state. Recalling the previous discussion on Damasio's definitions (see

Section 3.2.3.3), Öhman and Wiens (2004) proposed that *cognitive* should refer to postperceptual processes. Using the (relatively simple) example of the experience of fear, they proposed a fear module that operates in response to the fear stimulus and before cognitive appraisal processes initiate (see Figure 5). This fits well with the descriptions of the fear module from LeDoux (1998, 2000, 2002). In this mechanism which, from an evolutionary perspective, is designed for survival, LeDoux describes the operation in detail in *The Emotional Brain* (1998): the following diagram captures the key elements.

Figure 11: The Low and High Road to Fear



(LeDoux 2002, p. 123)

As the diagram shows there are different paths that both deliver input to the amygdala. LeDoux explains:

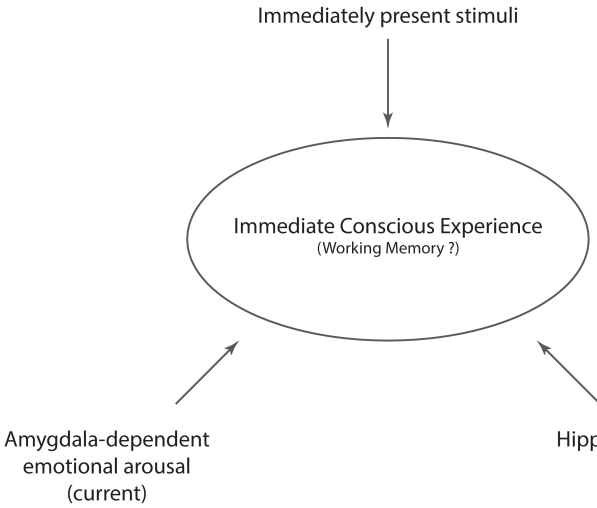
> This means that in addition to jump-starting the system, allowing for rapid initial responses, the thalamic information can also prime lateral amygdal cells to receive the more exacting information from the cortex. As a result, the cells are then more capable of charging ahead if the cortical information confirms the threat, or of putting on the brakes if the cortical information establishes that no danger is present (for example, that the loud crackling sound was from you stepping on a branch rather than from something dangerous, like a bear …)

(ibid., pp. 122–123)

Along with others, LeDoux considers that consciousness is related, at least in part, to working memory (Baddeley 2007; LeDoux 1998, 2002). The main difference between a conscious emotional experience and another kind of conscious experience may lie in the systems that provide input to working memory. In the case of fear there may be some

emotionally competent stimulus,[53] related to long-term memories and amygdala activity (LeDoux 1998), as depicted below.

Figure 12: Inputs into Immediate Conscious Experience



(After LeDoux 2000, p. 176)

In the case of the fright in the example above, even though a physical reaction starts before we may be fully aware of what we are reacting to (for example, we have not yet realised that the noise was not caused by a threatening animal), we do very quickly become aware of the feelings associated with quickened pulse, faster breathing, etc. It is only then that our attention may shift enough to realise the noise was created by a branch that we stepped on, and we soon experience a sense of relief accompanied, perhaps, with expressions such as *'What a fright that gave me!'*.

Thinking about the phenomenology of this set of responses LeDoux (1998) notes what is necessary for the emotion fear:

- Aspects of the experience must be represented in working memory.
- To experience complete fear, there must be activation of the amygdala.
- Arousal systems must be activated to have a *sustained* feeling of fear.
- There must be feedback from the body. As well, there is likely to be the involvement of long-term memories if the incident is not one experienced for the first time.

Now, we can imagine that once aware of the feelings of fright and subsequent relief, the person could produce a great variety of assessments of the situation. For example:

> *Well, I was lucky it was a branch and not a bear, but there may be bears, so let's get out of*

53. To use Damasio's term

here.

How silly, there's nothing to worry about here.

That's it. I hate bush-walking ... far too scary.

In simplest terms, the feelings in these circumstances are clearly information that there is something that needs urgent attention: the sound of a breaking branch could indicate danger. From a survival perspective this is useful information for us, giving one a chance to respond appropriately. How any one individual assesses the situation subsequently will be affected by how and where their attention is directed, what reference experiences they have stored in long-term memory as suggested by LeDoux (op. cit.), and their associated current beliefs. The point here is that even though unconscious processes may be preparing the body for action, the signalling from the feelings becomes evidence that the brain is reacting to *something* and appraisal is needed and action may need to be taken. Considering the range of possible appraisals appears to lend support to a model such as Scherer's depicted in Figure 7. However, that model could also be considered an apt depiction summarising LeDoux's words:

> We don't fully understand how the human brain sizes up a situation, comes up with a set of potential courses of actions, predicts possible outcomes of different actions, assigns priorities to possible actions, and chooses a particular action, but these activities are unquestionably amongst the most sophisticated cognitive functions.
>
> (ibid., p. 177)

LeDoux, like Damasio, looks to what we are learning from the neuroscience and separates emotion from cognition to better understand each as well as understand the interdependence of both. He suggests that emotions did not evolve as *conscious* feelings, and he also separates emotions from the associated feelings albeit, perhaps, not as distinctly and strictly as Damasio does (ibid.).

If we accept Damasio's definitions, then the view of feelings$_D$ as a perception that initiates action leading to re-establishing homeostasis or moving away from a source of injury seems uncontroversial and acceptable. However, as the emotions become removed from *obvious* homestatic concerns and become more abstract—for example, satisfaction, embarrassment—additional concepts seem to be required.

Damasio proposes a number of 'levels' of the processes leading to our conceptualisation of the 'self'. His notion of levels of 'self' helps us integrate his model of emotions and feelings into a coherent model. The contributions from three systems, Damasio proposes, lead to a 'preconscious biological precedent' (1999, p. 153) which he calls the *protoself*

(ibid., 2003b, 2010). These three systems are clearly associated with body functioning and homeostasis:

- internal milieu and visceral signals—principally chemical signals (e.g. hormonal changes, pH, etc.), but also signals from smooth muscles
- vestibular and musculoskeletal signals—proprioceptive signals
- fine touch—principally concerned with sensors in the skin detecting features such as texture, temperature, shape, etc.

Signals arriving in the brain from these three systems provide it with a map of the 'state of the physical structure of the organism' (1999, p. 154) moment by moment. We have no consciousness of this map, and Damasio suggests that it involves no language and that it 'has no powers of perception' and that it 'holds no knowledge' (ibid., p. 154). It is simply a record of the state of the organism. The principal brain structures involved in registering the requisite signals for a protoself are identified (ibid.; Damasio 2010; Parvisi & Damasio 2001).[54] A useful distinction in understanding the protoself is that it is 'not an interpreter; it is a **reference**' (Parvisi & Damasio 2001, p. 138, emphasis added).

It is proposed here that this concept of protoself provides a more coherent reference for facilitating appraisal of a situation than just the inputs that LeDoux suggests (see Figure 12). Even with signals from the amygdala, the emotionally competent stimulus and long-term memory, we can hypothesise that the brain needs some reference point to which all these signals can be related. If there is a danger present, for example, the question arises as to what is the object of the danger: it is reasonable to propose that the brain needs a representation of an organised entity that it directs to run away from the danger, for example. With the protoself we can represent possible results of exposure to this danger. In this there appears to be alignment, to some extent, with Prinz's notion of 'registration' of changes in the body that are signalled by emotions (Prinz 2004a, 2005, 2006).

In reading Damasio, he clearly has a philosophical interest in the concept of 'self' and how it arises from consciousness which, in turn, arises from neurophysiological activity. However, there is the sense that the move into the enquiry of how the self emerges was an inevitable step in explaining the implications of feelings$_D$ being the *conscious awareness* of the body–mind state.

54. This is mentioned to highlight that there is strong empirical research in neurophysiology underlying Damasio's hypothesis. Details of the neural structures are beyond the scope of this thesis.

4.1.1.2 Levels of Self

Damasio's interest is in developing a *neuroanatomical* model for the development of consciousness and for the sense of 'self'. If there is a major challenge in some of Damasio's writing, it is that he seems to be developing his hypothesis while in the act of writing, especially in books compared to journal articles. Consequently, it requires considerable detective work to distil key elements from numerous apparent repetitions. Additionally, he publishes developments over time; and a major effect can be the anticipation of further clarification. For example, one of the early presentations of the concept of a protoself (Damasio 1999) is made significantly clearer in subsequent work (Parvisi & Damasio 2001). Therefore, it is not surprising that he has further developed his views (Damasio 2010) on how the protoself is instrumental in the processes giving rise to consciousness.

As an aside, the aim is not to critically examine Damasio's model for consciousness per se. Rather, it is to enquire how its key elements contribute to our understanding of the creation and maintenance of beliefs. Hence, while there are aspects of his hypothesis that can leave one eager for future clarfication, there remain insights from Damasio's work that contribute significantly to understanding how beliefs, and false and irrational beliefs in particular, may be formed and maintained. One particular example is the introduction of 'primordial feelings' (ibid.).[55] It seems that these are proposed in response to the problem suggested above: if the working memory is to refer to *something*, what is that something. Damasio expresses it thus:

> The issue of how perceptual maps of our body states become bodily feelings—how perceptual maps are *felt* and *experienced*—is not only central to the understanding of the conscious mind, it is integral to that understanding. One cannot fully explain subjectivity without knowing about the origin of feelings and acknowledge the existence of *primordial feelings*, spontaneous reflections of the state of the living body. In my view, primordial feelings result from nothing but the living body and precede any interaction between the machinery of life regulation and any object. Primordial feelings are based on the operation of upper-brain-stem nuclei, which are part and parcel of the life-regulation machinery. Primordial feelings are the primitives for all other feelings.
>
> (ibid., p. 101)

55. This term seems to appear in the book *Self Comes to Mind* but not in journals (at least in literature reviewed). Damasio has a more conversational approach in his books and it is in these that the appearance of writing as a way of developing a hypothesis seems more evident. It is suggested, with respect, that journal editors may require greater precision in terminology. The results of the writing-to-think process may be no less valid but they do require some dedication from the reader – dedication which is rewarded, it should be added.

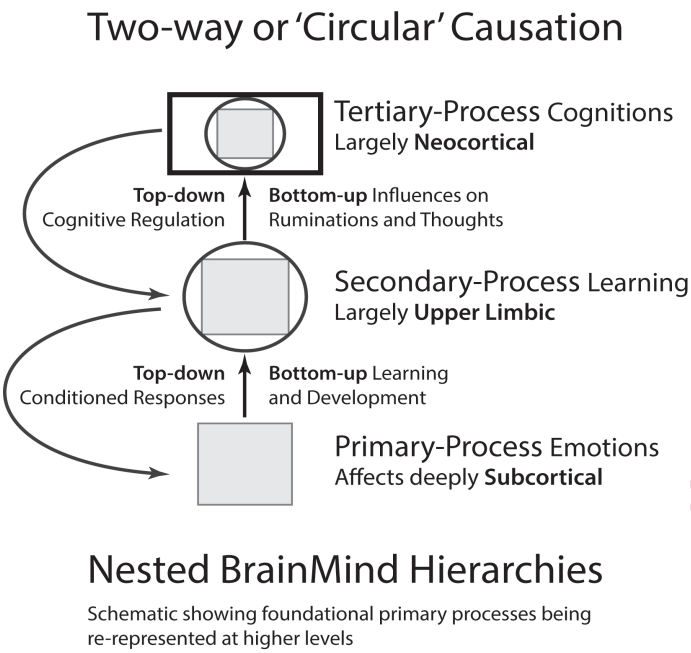
Damasio's (ibid.) recent inclusion of the 'primordial self' helps account for the feeling of existence that is experienced by a 'core self' (explained below). He suggests this process can only work if there is an 'elementary feeling' (ibid., p. 322) generated by the brain-stem component of the protoself. This elementary feeling is hypothesised to be independent 'of any object interacting with the organism and thus modifying the protoself' (ibid., p. 322). The term 'primordial' appears to be chosen because these feelings 'precede all other feelings' (ibid., p. 193).

Others also propose a similar arrangement, most notable amongst them are Panksepp and Northoff (Northoff *et al.* 2006; Panksepp & Northoff 2009; Panksepp 2005a, 2012) and Gallagher (Gallagher 2000). To reiterate, the object here is not to carry out a full critique of Damasio's hypothesis; but, it is important to note that there is a growing acceptance of the notion of levels of consciousness, and that Damasio is not alone in seeking to understand consciousness from a neurophysiological basis. Although Damasio points to differences in details between his hypothesis and Panksepp's (see pp. 322–323 Damasio 2010), Panksepp supplies some explanation which, it is suggested here, illuminates Damasio's work.

Panksepp (2005a, p. 32) sees emotional experience as 'closely linked to internal brain action states, triggered typically by environmental events'. What Panksepp refers to as 'primal emotional feelings' (2012, p. 4) are responsible for informing 'the rest of the mental apparatus about basic survival values'. In a meta-analysis, Northoff et al. (2006, p. 454) highlight the central, and essential, nature of *self-referential* processing which they conclude is 'at the core of what is called the self'.

The proposal from Panksepp (2005a) is that consciousness emerges from a multi-tiered process. There is a primary-process consciousness comprising predominantly the 'raw sensory/perceptual feelings' (ibid., p. 32). The ability to reflect on these experiences and have thoughts about them is considered 'secondary-consciousness', while 'tertiary forms of consciousness' include 'thoughts about thoughts' and 'awareness of awareness'. Panksepp's (2012) representation of the relationship between these levels is depicted thus:

Figure 13: Panksepp's Nested BrainMind Hierarchies



(After ibid., p. 7)

Northoff (2011, p. 187) succinctly describes 'self-related processing' as the 'formal mechanism whereby the relationship between brain and stimulus is realised in the neural activity of the brain'. Returning to Damasio's protoself and its primordial feelings: these, then, comprise the first object that the brain can recognise as a coherent entity, a basic self (for want of more precise terminology) that can be motivated to act and which can be perceived in relationship to some stimulus. Northoff et al. (2006, p. 440) simply suggest that the protoself 'refers to the domain of the body'.

Similar to Panksepp, Damasio proposes three 'levels' of processing in creating consciousness.[56] The first step in developing consciousness is generation of the primordial feelings from the protoself. Then follows the *core self* which is able to 'describe an object engaging the protoself and modifying that protoself, including its primordial feelings' (ibid., pp. 22–23). According to Damasio, these first two stages comprise the 'material me' while the final stage—the *autobiographical self*—is the domain of biographical knowledge, memories, plans for the future, and social persona.

56. As mentioned previously, there are differences with Panksepp's model. Damasio (2010) gives a detailed neuroanatomical explanation of the difference. Although discusssion of those differences is not relevant to this thesis, it is important to record that the differences are acknowledged and explained.

In order for the brain to engage in self-referencing processes it follows that there has to be some representation in the brain to reference. As Damasio puts it, the brain '*informs* itself' (ibid., p. 63); but there has to be some 'thing' about which it informs itself as well as a mechanism for effecting the informing.

Initially, Damasio used the terms *neural pattern* or *map* to describe the neural activity in the brain while using the term *image* 'as a synonym of mental pattern or mental image' (ibid., p. 64).[57][58] In this later writing, he admits to using these terms (neural pattern, map, image) 'almost interchangeably' and his reason for doing so is 'to underscore the fact that the distinction [between brain activity and mind], while valid, can block the view' (2010, p. 65) of what is being explained. Damasio also takes great pains to emphasise that this is not dualism: there is no humunculus, but there is the risk that the idea of levels of consciousness could lead to that misinterpretation in the attempt to explain how the brain instigates self-referencing processing in which one appears to observes one's own state (for example, 'I'm tired from standing so long, my legs are aching, I need to sit down').[59]

The maps that Damasio identifies are:

	Maps	**Source Object**
I	Maps of the organism's internal structure and state (Interoceptive maps)	The functional condition of body tissues such as the degree of contraction/distension of smooth muscles; parameters of internal milieu state
II	Maps of other aspects of the organism (Proprioceptive maps)	Images of specific body components such as joints, striated muscles, some viscera
III	Maps of the world external to the organism (Exteroceptive maps)	Any object or event that engages a sensory probe such as the retina, the cochlea, or the mechano-receptors of the skin

(Adapted from ibid., p. 76)

57. 'Image' might be better replaced by *representation*. Working through Damasio's writings, he uses 'image' for *any* form of sensory representation (i.e. visual, auditory, kinesthetic, proprioceptive, smell, taste). Damasio explains his reasons for his caution with the term 'representation' in *The Feeling of What Happens* (1999, p. 320).

58. Johnson (2007) takes issue with Damasio's (1999) claiming, fairly it is suggested, that it can be misleading. He provides a useful clarification: 'Consequently, when Damasio says that a feeling is an idea of the body, he is not saying that it is some abstract inner mental picture of the state of the body. Rather, the feeling is our felt awareness of something going on in the body.' (2007, p. 64)

59. This point is raised many times by Damasio. There is an implicit suggestion that he has been accused of being a dualist on more than one occasion.

These maps, or representations, of the body and external objects become the basis for developing the proposed levels of consciousness. Damasio (ibid., p. 202) suggests there is a need for an 'intermediate self process' between the protoself and the autobiographical self because of the implied gulf between the levels of complexity of the processing involved. So, he proposes a **core self** and it this processing that Damasio suggests allows the protoself to *'stand out'*, connect with events, and 'protagonise' (ibid.). Distilling the message, this can be interpreted as suggesting that the core self processes are those that allow the mind to be aware of the body being engaged with and affected by an object. As is emphasised later, the object is not necessarily always external; it can be an imagined scenario that changes the state of the body and, consequently, the feelings experienced. The principal sequence for core self processing is described:

> … an object engaged the body when that object was looked at, touched, or heard, from a specific perspective; the engagement caused the body to change; the presence of the object was felt; the object was made salient.
>
> (ibid., p. 203)

During these core self processes, a coherent pattern forms that links the object and organism for a time. In referring to this as a 'pulse' of core self, Damasio emphasises the transitory nature of the processing producing this particular representation and feelings of the body in relation to some object.[60] The main components of this composite core self state are '*feelings of knowing*', '*saliency* of the object', and '*perspective* and the sense of *ownership* and *agency*' (2010, p. 206). All this processing to re-represent the organism, the object, and the relation between them is considered to be a nonverbal process.

The third level of consciousness in this model is the **autobiographical self** which Damasio considers the 'traditional notion of self [which] is linked to the idea of identity and corresponds to a nontransient collection of unique facts and ways of being which characterise a person' (1999, p. 17). It involves conventional memory and working memory. Language (that is, words experienced through sight or hearing) plays a major role; although Damasio (ibid.) suggests that language is not necessary for this level of consciousness and describes being able to communicate with patients with global aphasia[61], even with complete removal of the left hemisphere, as part of his evidence.

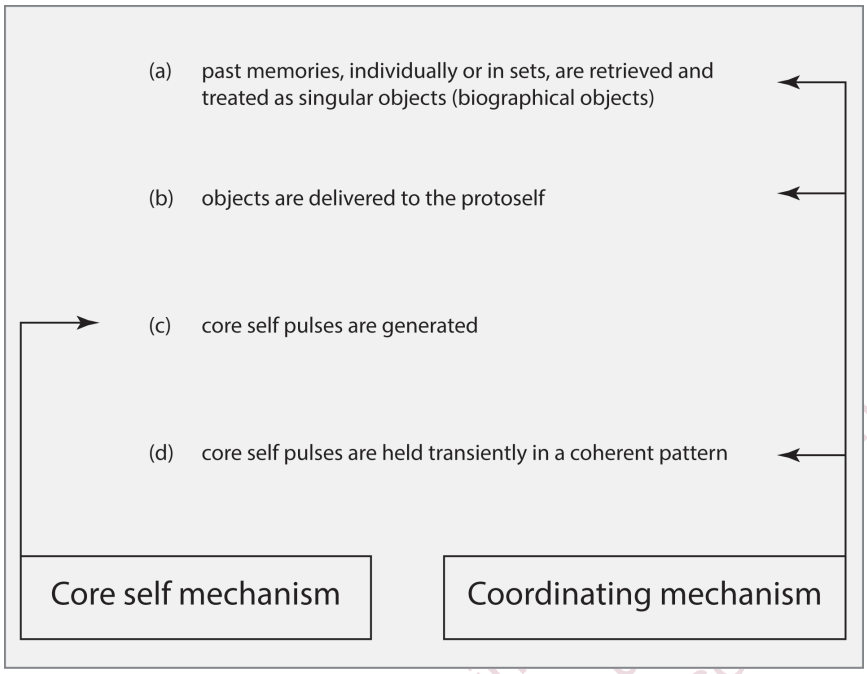
As might be expected, this autobiographical level is much more complex in its operations. Damasio (2010) suggests two mechanisms working in conjunction: one ensures that biographical memories are treated as objects and made conscious in core self processes ('in

60. Damasio's use of 'pulse' appears to be inspired by James' use of 'pulse of consciousness' (James 1890/1950a, Chapter X)

61. The inability to comprehend language either auditorily or visually.

a core self pulse'); a second coordinates memories and their representations and allows them to interact with the protoself. He depicts them thus:

Figure 14: Neural Mechanisms for the Autobiographical Self



(After ibid., p. 213)

Scherer (2004; Leventhal & Scherer 1987) too, as noted in Section 3.2.3, separates emotions from feelings. He highlights an important challenge in suggesting the need to 'resynthesise the elements or components of a phenomenon' and, in the case of emotion shows how those components 'form a phenomenologically indivisible whole – feeling' (2004, p. 154). Although we can acknowledge that much work remains to fully explicate how levels of consciousness can arise from neurophysiological phenomena, Damasio's model takes us a long way towards the synthesis suggested by Scherer. There is an increasingly detailed understanding of how the neuroanatomical structures are responsible for receiving signals, producing signals, and coordinating internal communication. The full test of Damasio's model will come with further developments in this area. However, taken in conjunction with models such as Panksepp's (on page 157) and Northoff's (2006), we have a much stronger model for discussing how feelings become evidence that contribute to forming and maintaing beliefs.

Starting with a model that is founded in the body seems almost too obvious when it is noted that *'the representation of the world external to the body can come into the brain only via the body itself'* (Damasio 2010, p. 91). The model is all the more plausible when we note that a person can imagine themselves in an emotion-inducing situation (for example,

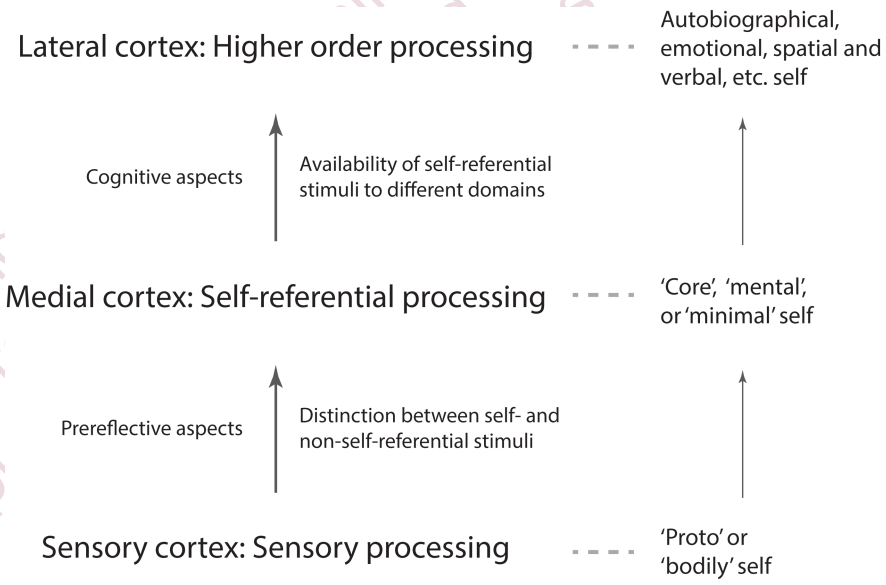
anxiety about giving a speech) and that this emotion is recognised through what is *felt* in the body along with any accompanying thoughts. Again, the feelings become the indicator of our relationship with the current situation. Prinz suggests that:

> The trick is to use the body as an indicator of how we are faring in the world.
>
> (Prinz 2006, p. 147)

Scherer also proposes that feeling is 'a monitoring system that consists of a central representation of the response organisation, including the underlying cognitive processes in an emotion episode' (2004, p. 137). Prinz (2006) suggests that the distinctive emotional patterns associated with a given set of circumstances indicate the relation between an organism and the environment that bear on its well-being; as such he labels the relations 'concerns'. Damasio's model has explanatory power in explaining how emotion is a reaction to the model of the situation as suggested by Clore and Gasper (2000) and how it may shape their response.

A useful overview of levels of processing with neural correlates is provided by Northoff (2006):

Figure 15: Cortical Localisation and Concepts of Self



(ibid., p. 450)

When it comes to emotions' influence on beliefs, Frijda and Mesquita suggest that 'concern relevance is the core notion' (2000, p. 73). Concerns can be as simple as recognising the danger associated with a sharp object, through to imagined catastrophes resulting from a forthcoming restructure in one's workplace or the likely social disintegration from the most recent election results.

4.1.1.3 Not-So-Simple Concerns

The apparently simple reflex action of suddenly withdrawing one's hand from a hot saucepan clearly causes a body state. So the question arises as to whether such reflex actions qualify as emotions. According to Damasio they do not (Damasio 1999, 2003b). Such startle reflexes are part of regulatory responses that comprise simple behaviours such as (rapid) removal of a limb from a source of damage. Even a paramecium will move away from something that impinges on its one-celled body ('Poke a paramecium, and it will swim away from the poke.' (Damasio 2010, p. 257)). However, while reflexes do not constitute emotions, they can contribute to the homeostatic signals that collectively constitute an emotion (Damasio 1999), as suggested in the earlier discussion of Damasio's proposal of 'nesting' (see Figure 3 and related discussion in Section 3.2.1).

From a similar perspective, neither pain in itself nor pleasure in itself counts as an emotion (ibid.; Damasio 2003a). Pain is generally associated with potential or actual tissue damage (punishment) and behaviour that causes one to withdraw or freeze. But, the stimulus that causes pain can *also* cause an emotion. Craig (2003, p. 306), in fact, concludes that pain is a 'homeostatic emotion reflecting an adverse condition in the body that requires a behavioural response'. Pleasure, on the other hand, generally arises from the correction of an imbalance: for example, eating when hungry, drinking when thirsty. It is associated with reward and with seeking or approaching behaviours (Damasio 1999).

> Pain or pleasure are prompted by many causes – glitches in some body function, optimal operation of metabolic regulation, or from external events that damage the organism or protect it. But the *experience* of pain or pleasure is *not the cause of the pain or pleasure* ***behaviours***, and is by no means necessary for those behaviours to occur.
>
> (Damasio 2003b, p. 33, bold emphasis added)

In the hot saucepan case, the reaction is instigated unconsciously (if the saucepan is hot enough!) before we are fully conscious of what is occurring. All that is really needed is to let the reflex occur and possibly run cold water on the burnt hand. But, there can be more: now that the person is aware of the feelings of being burnt, there is vast range of possible responses that are easily classed as emotional (for example: *How am I going to play tennis with a burnt hand? I'm so angry with myself.* Or *Stupid saucepan, why were you there in the first place?* etc.) The same stimulus produces both a reflex action and an emotion.

Safety is clearly a concern here. At a basic level the feeling of being burnt provides information about the degree of care needed: a slight burn may need only running water for a few minutes; a severe burn may need medical attention and the accompanying pain suggests the wisdom in restricting the use of the hand for a time. However, suppose the

person says something like *'I'm so stupid; how did I do that? Damn, I'm so clumsy!'* Now we begin to see the full relevance of Damasio's model (Figure 14). Previous instances of the person burning themself are retrieved; the representation of these instances alters the state of the protoself; the core self is now operating not just from the awareness of the burn but also from the awareness of the feelings instigated by the memories of previous failures of coordination and the consequent damage. The complexities of the emotion state are not reflected in either Figure 13 or Figure 14; they are captured better by Frijda:

> Emotions are always the outcome of a balance of multiple appraisals, multiple meanings, and relevance to multiple concerns.
>
> (Frijda 2004, p. 164)

So now we now hear utterances relating to the person's self, a general judgment on their intellectual capacity, and an expression of their belief regarding their clumsiness. There may well be considerations of treating the burn and decisions about seeking medical attention, but as Frijda puts it, there are multiple appraisals, meanings and concerns being processed.

As implied in Figures 13 and 14, these suggested 'levels' of consciousness derive from suggested 'second-order' mapping processes (Damasio 1999, 2003a). Damasio also suggests that a third-order of processing may occur when 'the second-order nonverbal narrative of consciousness' is converted to language (1999, p. 185). In suggesting second-order processing, Damasio is far from alone (Barrett *et al.* 2005b; Charland 2005; Critchley *et al.* 2001, 2004; Lambie & Marcel 2002; Northoff 2008; Panksepp 2005a).[62]

As might be anticipated there is some considerable debate about aspects of this approach. In the reviewed literature, particularly the more recent, there appears to be general acceptance of Damasio's contributions. Criticism largely appears directed to aspects that are considered omissions, explanations that need clarifying, and lack of rigour in some terminology (for example: Legrand 2007; De Preester 2007; Mosca 2000). Entry into the depth of philosophical debate is beyond the current scope: the intention here is to maintain the focus on how emotion and feelings as defined by Damasio contribute to understanding how beliefs are formed and maintained. The details of exactly how the brain creates self-awareness and self-consciousness will continue to be the subject of ongoing theoretical and empirical investigation. Gallagher (2000), in urging collaboration between philosophers and cognitive scientists, suggests that 'at present the neuroscientist, like the philosopher, can offer, at best, informed speculation on these processes' (ibid., p. 20). With respect to Damasio's contribution to this collaboration he considers that *The*

62. These are representative of relevant articles in the reviewed literature.

Feeling of What Happens (Damasio 1999) 'has insightfully captured the difficulty involved in expressing the interrelations between the minimal ('core') self and the narrative ('autobiographical') self (2000, p. 20). Here, the last word on this aspect of the debate is given to Gallagher and Zahavi:

> The recognition of the existence of a primitive form of pre-reflective self-consciousness is an important starting point for an understanding of more elaborate forms of self-consciousness that are concept- and language-dependent.
>
> (Gallagher & Zahavi 2010, Conclusion)

4.1.2 Following Feelings

Russell Ackoff[63] says this of problems:

> Problems are to reality what atoms are to tables. We experience tables, not atoms. Problems are abstracted from experience by analysis. We do not experience individual problems but complex systems of those that are strongly interacting. I call them *messes*.
>
> (Ackoff 1999, p. 117)

Ackoff's use of 'mess' is not perjorative; it is simply powerfully descriptive. In the areas of neuroanatomy and neurophysiology we do not yet have the metaphorical equivalent of laser diffraction crystallography that can see the 'atoms'. There is little doubt, however, that feelings$_D$ act as information, as feedback for the brain–body system. Clearly, there are still specific details of the feedback mechanisms and the representational mechanisms to be determined. Nevertheless, the literature reviewed encourages following the implications of the information value of feelings$_D$.

Much of the groundwork has already been laid in the discussion of the principal contributions to this thesis thus far from Clore and Gasper's (2000) and Frijda and Mesquita (2000) (see Section 3.4). Given the mess in which emotions influence beliefs, and in which beliefs may also be considered 'part of emotions' (ibid.), the more straightforward aspects of the influence of feelings are considered first.

Returning to the example of burning one's hand on a hot saucepan. Although it has been established that reflex actions are not considered emotions, it was noted that feelings$_D$ follow quickly. In the simplest instance, the person may decide to run cold water on the affected hand. Of course, such a response could be habitual.[64] But even in the unfortunate circumstance that it were a habit, Overskeid (2000) suggests there is no decision to be made in following a habit and, therefore, there is no expected influence from feelings. In

63. Renowned for his work in systems thinking

64. Although one hopes not, as that does not auger well for the hand's long-term well-being

this case, though, it is unlikely to be following a habit. It is far more likely that there were some quick appraisals based on the perceived circumstances: for example, deciding whether the saucepan needs to be removed from the hotplate; determining if the food is burning; assessing the urgency for treating the hand. And, this last consideration is most likely informed by the degree of pain felt, and possibly the sight of the burn site. This point is sufficiently obvious as to be almost trivial; but, it is a valid example of feeling informing action and initiating thoughts that will alter the body state (Damasio's protoself) and the awareness of it (Damasio's core self).

More importantly, the action of running cold water on the burn has been initiated by the belief that it is efficacious in alleviating the pain and limiting the tissue damage. Since it is highly improbable that the person was dwelling on a such a belief before the incident, it follows that this belief was brought to mind by the feeling of the burn. It would be a belief that is part of 'background knowledge' (Churchland & Churchland 2013) or a background belief as described by Morton (2003). Again, none of this is remarkable but it does give scope for investigating the application of Panksepp's hierarchy (see Figure 13): one or several incidents of being burnt helped the person learn that hot objects can be injurious; this, in turn, led to learning remedial action to take, and then applying that remedial action in the relevant circumstances.

If there were an 'hierarchy' of beliefs in this circumstance, the belief that running water on the hand is beneficial would likely be 'low' down the hierarchy. It is certainly one of the first beliefs to be engaged even though it might not have been consciously brought to mind: having accepted that belief in the past, the person could well act on it and might not articulate it unless, for example, someone asks the person why they are running water on their hand. However, as already suggested there are other beliefs that can be formed or reinforced rapidly. Take the example of the person saying *I can't count how many times I've done that, I'm so clumsy; I shouldn't be trusted in the kitchen.*

Now, Damasio's model for mechanisms for the autobiographical self (see Figure 14) becomes more immediately relevant. There are now at least two more autobiographical beliefs involved in this, now more strongly emotional, experience: *I'm so clumsy* and *I shouldn't be trusted in the kitchen*. It is one thing to suggest that memories of past similar occasions (and related beliefs) are retrieved and 'presented' to the protoself and generate a different (possibly more upset or angry, or both) core self process; but the question arises as to what triggers the retrieval. This is not an instinctive evolutionarily-derived response, not is it a case of classical conditioning; yet something triggered the retrieval of the relevant memories. It is here that Damasio's somatic marker hypothesis has explanatory force.

4.1.2.1 The Somatic Marker Hypothesis

Bechara and Damasio (2005, p. 351) describe the somatic marker hypothesis as being concerned with 'emotion that is integral to the decision-making task at hand'. Damasio et al. (1996) summarise the hypothesis:

> The key idea in the hypothesis is that 'marker' signals influence the processes of response to stimuli, at multiple levels of operation, some of which occur overtly (consciously, 'in mind') and some of which occur covertly (non-consciously, in a non-minded manner). The marker signals arise in bioregulatory processes, including those which express themselves in emotions and feelings, but are not necessarily confined to those alone. This is the reason why the markers are termed somatic: they relate to body-state structure and regulation even when they do not arise in the body proper but rather in the brain's representation of the body.
>
> (ibid., p. 1413)

The neural structures related to establishing somatic markers are covered in detail in the writings of Damasio and others (Damasio *et al.* 1996; Damasio & Carvalho 2013; Damasio 1994, to list a few). Of interest here is the function of somatic markers. One of Damasio's earlier definitions helps towards gaining a clearer understanding:

> … *somatic markers are a special instance of feelings generated from secondary emotions.* Those emotions and feelings *have been connected, by learning, to predicted future outcomes of certain scenarios.*
>
> (Damasio 1994, p. 174)

By 'secondary emotions', Damasio is referring to what he later labelled as *social emotions* (1999). This term seems preferable as he views somatic markers are largely associated with 'effective personal and social behaviour' (1994, p. 174). Somatic markers are acquired and adapted over time and their development occurs under the control of an 'internal preference system' as well as under the influence of external circumstances comprising other entities, events, social conventions and ethical rules.

Prinz helps by simply observing that there is evolutionary advantage in being able to record *patterns* of bodily changes as certain patterns are linked to particular circumstances. In such circumstances, or similar ones, the brain could 'register danger [for example] and … use that signal to tell action-selecting systems to search for response strategies that are useful …' (2006, p. 148). Damasio et al. (1996) list the following operations of somatic markers:[65]

- They act as a repository between factual knowledge and bioregulatory states.

65. They also list the associated neural structures.

- They lead to recall of pertinent facts associated with the current circumstances.
- When overt (that is, conscious), the somatic state can act as an alarm signalling potential harm or danger, or it can as an incentive signal. When covert (that is, non-conscious), the somatic marker acts as a bias (also suggested in: Damasio 1994; Damasio 2010)
- They may 'boost' attention and working memory processes.
- They *may* facilitate logical reasoning (though this is not always the case). Their hypothesis is that 'somatic markers normally help constrain the decision-making space by making that space manageable for logic-based, cost-benefit analyses' (1996, p. 1415).

Somatic markers, it seems, are particularly geared to sending signals when possible outcomes are uncertain; and this is especially the case in personal and social matters as they are 'frequently linked to punishment and reward … and the regulation of homeostatic states' (ibid., p. 1416).

Note that, as mentioned earlier, these somatic markers are not simple reflex actions, nor are they instances of classical conditioning. The individual has recorded memories formed from the recorded facts and the associated body states. Damasio et al. suggest that while the facts (presumably as perceived by the individual) may be stored, the body state memories do not need to be permanently stored (presumably because the states can be re-created whenever the relevant stimulus occurs). What the somatic marker, which is stored permanently, does is provide the *link* between 'certain classes of situation and certain body states' (ibid., p. 1417).

Prinz (2004b), who also defends a somatic theory, suggests that to explain why emotions are so meaningful, they are 'not merely perceptions of the body but also perceptions of our relations to the world' (ibid., p. 20). In an approach similar to Damasio's, he proposes that the brain needs 'a way of recording which stimuli are positive reinforcers and which are negative' (ibid., p. 173) and postulates an 'inner negative reinforcer' and an 'inner positive reinforcer'. These reinforcers are stored in memory along with related 'embodied appraisals' and exercise an influence on future behaviour.[66] When imagining an experience that one has valued in the past (for example, a particularly enjoyable type of food), the positive reinforcer acts to value the available option highly. Similarly, when deciding on something that has evoked negative affect before (for example, a dinner with company

66. When Prinz refers to *embodied appraisal*, he means that the appraisal occurs as a result of registering bodily changes (ibid., p. 78).

that one previously found unutterably boring), the negative reinforcer and associated embodied appraisal signal avoidance.

A useful comparison of two ways that a negative reinforcer could work is given by Prinz (ibid.) in examining two instances of fear. In one case, a person might be bushwalking and suddenly see a snake. As a result they might freeze, their heart rate will increase, their breathing pattern changes, and their eyes may widen. For a different case, consider a student about to sit an exam that they must pass to enter a course they greatly want. On seeing the exam desk and book, they exhibit similar physical behaviour to the person seeing the snake. However, despite superficial appearances, the cases are quite different. In the first (the snake), the fear system goes into action. As LeDoux (1998, p. 128) puts it, this system 'is not, strictly speaking, a system that results in the experience of fear. It is a system that detects danger and produces responses that maximise the probability of surviving a dangerous situation in the most beneficial way'. As noted earlier in the burnt hand example, it is unlikely that the person would *remain* in the simplest response state, and the emotional response will develop. But, the second case is different in that there is no immediately visible physical threat. Prinz (2004b) suggests that the exam case inherently requires cognition in order to recognise the exam as an object of fear. We can safely propose that the state of fear at the sight of the exam books, which are not physically threatening, has more to do with beliefs about the importance of the exam along with other possible related beliefs about the person's self efficacy in an exam situation.

We can also safely propose that something intervened between seeing the exam books and the resulting state of fear. In Prinz's (ibid.) terms, an inner negative reinforcer has signalled some situation to avoid. In Damasio's terms, a somatic marker has signalled a situation that, at the least, requires attention and beyond that may require decisions and appropriate action. In such situations which involve far more than reflex actions, classic conditioning, and habits, the following aspects now seem uncontroversial:

- Emotions are signalling concerns as discussed in Chapter 3, and the existence of a concern is announced via feelings. Prinz sums it up well with:

  > {Emotions] are like bodily radar detectors that alert us to concerns. When we listen to our emotions, we are not being swayed by meaningless feelings. Nor are we hearing the cold dictates of complex judgments. We are using our bodies to perceive our position in the world.
  >
  > (Prinz 2004a, p. 240)

- Beliefs may be intricately involved in emotional experiences when $\text{feelings}_{\text{D}}$ are *felt*.
- Signals—and the term 'somatic markers' is accepted here—direct attention, cause the retrieval of memories and related beliefs, and guide decision making.

Feelings$_D$, then, signal concerns that can be related back to basic survival and homeostatic balance. However, the concern about safety and balance goes beyond simple physical safety. That we can identify with what Damasio refers to as 'social' emotions, indicates that the perception of balance and safety goes beyond the fear response that is appropriate when confronted with a highly venomous snake or beyond the reflex action when burning one's hand. In unravelling how feelings$_D$ contribute to, and help maintain, beliefs, Damasio's 'As If Body Loop' assists.

4.1.2.2 'As If' Body Loops

The example of the fear experienced at the sight of the exam books suggests that the person can experience, for example, the feelings of whatever unpleasant emotion (perhaps shame, disappointment or sadness, anger, or a combination of all these) might accompany the imagined failing of the exam. Similarly, a pleasurable state can begin in *anticipation* of eating or drinking favourite beverages and food, for example. While it is arguably obvious that we can anticipate these emotional states and experiences, Prinz (2006) proposes that the systems involved activate before the body changes:

> It is extremely plausible that, when we encounter familiar emotion elicitors, our interoceptive systems become active before the body has had time to change. The brain anticipates what the body will do.
>
> (ibid., p. 145)

So, Prinz suggests that the brain both registers *and* anticipates bodily changes. In Damasio's terms, 'the brain allows us to *hallucinate* certain body states by a variety of means' (2003b, p. 118). More precisely:

> … the brain can *simulate*, within somatosensing regions, certain body states, *as if* they were occurring; and because our perception of any body state is rooted in the body maps of the somatosensing regions, we perceive the body state as actually occurring even if it is not.
>
> (2010, p. 102)

Damasio (ibid.) admits that initially there was only circumstantial evidence in favour of this 'as if body loop' hypothesis when he first proposed it. However, there is now considerable supporting evidence from studies in the mirror neuron system (Fogassi 2011; Gallese & Goldman 1998; Gallese 2001; Gazzoa *et al.* 2006; Gazzola *et al.* 2007; Knapp & Corina 2010; Ocampo & Kritkos 2011; Rizzolatti & Craighero 2004; Rizzolatti 2005) and from studies in *efference copy* (for example, Niziolek *et al.* 2013; Wolpert & Flanagan 2001; Wolpert *et al.* 2001) in which the brain creates predictions of motor movements (for example, lifting an object when one is uncertain of the weight) and uses sensory feedback to adjust the movement. Wolpert and Flanagan even suggest:

> Not only is prediction essential for motor control, it may also be fundamental for high level cognitive functions including action observation and understanding, mental practice, imitation and social cognition.
>
> (Wolpert & Flanagan 2001, p. R732)

This suggestion points to Damasio's social emotions (1999, 2003b) and to the general issues of concerns (Frijda & Mesquita 2000; Prinz 2004b, 2006). Bechara and Damasio (2005) refer to primary and secondary inducers of emotion:

> Primary inducers are innate or learned stimuli that cause pleasurable or aversive states. Once present in the immediate environment, they automatically and obligatorily elicit a somatic response.
>
> (ibid., p. 340)

Examples of the primary inducers include: fear objects such as a snake, or even a stimulus that predicts some fear object; concepts or knowledge such as learning that one has won the lottery. Following LeDoux (1998) and Morris (1999), they suggest that the amygdala is an important trigger structure for somatic states arising from primary inducers.

> Secondary inducers … are entities generated by the recall of a personal or hypothetical emotional event, i.e., "thoughts" and "memories" of the primary inducer, which when brought to working memory elicit a somatic state. Examples of secondary inducers include the emotional response elicited by the memory of encountering a snake, or the memory of losing a large sum of money. The imagination of being attacked by a bear, winning an award, or losing a large sum of money, are also examples of secondary inducers.
>
> (Bechara & Damasio 2005, p. 340)

Bechara and Damasio (ibid.) suggest that the ventromedial prefrontal cortex is the trigger structure for the states arising from secondary inducers. The neuroanatomical details need not be pursued here; it suffices to take up their suggestion that this structure operates as a 'convergence-divergence zone' that brings together:

- a certain category of event based on memory records in high order association cortices,
- effector structures that execute the somatic state, and
- the neural patterns related to the non-conscious or conscious feeling of the somatic state (ibid., p. 341).

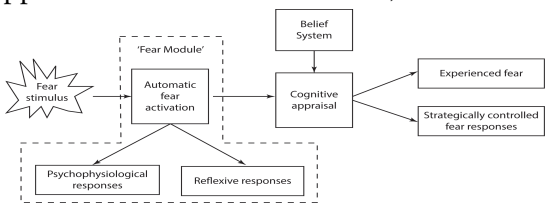
That is, the ventromedial prefrontal cortex links knowledge of events associated with the secondary inducers to the feeling of being in the situation. In other words, the brain can simulate a body state so that we perceive it as occurring even though it is not (Damasio &

Damasio 2006). This integrated representation of the body can serve as a 'stand in' for the self in relation to objects that are perceived (ibid.); and it has been suggested that objects include concepts such winning the lottery. Thus, there is evidence for mechanisms that trigger the kinds of somatic states that arise subsequent to events such as burning one's hand, approaching an important exam, hearing that one has won the lottery, and so on.

Importantly, Bechara and Damasio record that 'evidence suggests that, in a normal brain, primary and secondary inducer processing can be elicited by the same stimulus and at the same time' (Bechara & Damasio 2005, p. 340). This may explain how it is then that in an extremely short time after executing a reflex withdrawal of a hand from a hot saucepan, the person can very quickly be angry with or disappointed in themself for being so clumsy.

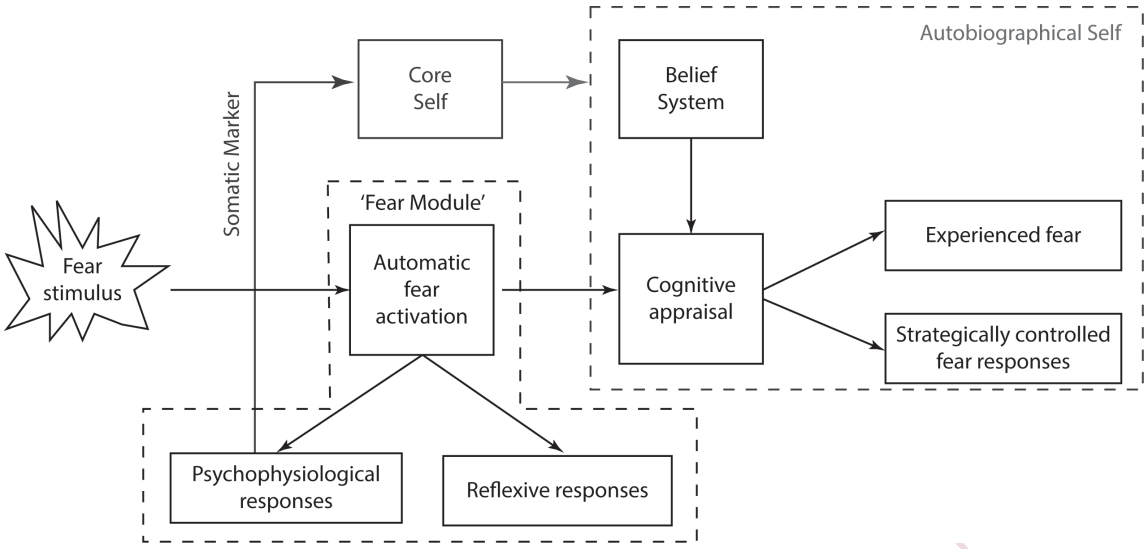
Because somatic markers are connecting the memory of circumstances and objects with the feelings of the related emotions, learning is required (Damasio 1994). It is reasonable then to propose in the burnt hand example that there has been more than one instance of such an event and conceivably the person has established a representation of themselves in relation with hot saucepans that involves pain. The secondary inducer evokes not only the memory of previous similar events but may also retrieve a belief they have formed relating to their clumsiness or worthiness as a cook. Damasio poses the concept that somatic markers draw our attention, either through positive or negative feelings, to our concerns by referring to 'life-managment needs' (2010, p. 175).

Reviewing the models presented in Chapter 3, the concept of somatic markers suggests possible answers to questions (some of which appeared not to have been asked). For example, in the Öhman and Wiens fear module shown in Figure 5,[67] the input from the belief system seems indisputable; but, the question is: what triggers its input? Given the neurophysiological support for somatic markers, it is plausible that the thoughts and memories induced by the initial stimulus proceed to induce the introduction of associated beliefs into working memory. An adaptation of this model could include an arrow going from the psychophysiological responses through the core self to the autobiographical self containing the belief system:



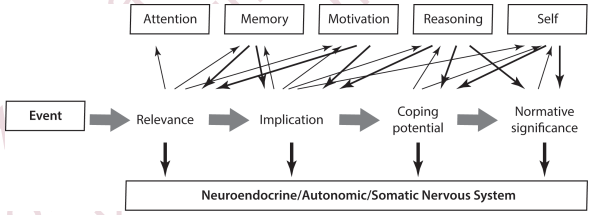
67. And shown in miniature here for ease of reference.

Figure 16: Variation to Fear Module suggesting possible influence of somatic marker signals

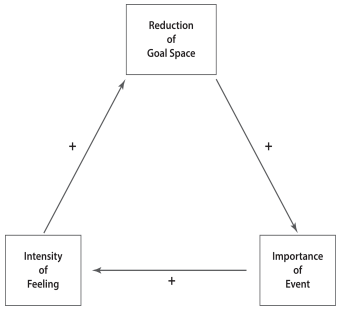


(Developed from Öhman & Wiens 2004, Fig 5.3(b), p. 75)

While the fine details of such an induction still elude us, it is safe to conclude that the felt experience—Damasio's feelings$_D$—are likely to be a significant signal in the process. Similarly, with Scherer's model shown in Figure 7, somatic markers appear to go some way towards answering the question about what exact stimulus is being checked in Scherer's (2001, 2004) proposed 'stimulus evaluation checks' (relevance, implications, coping potential, normative significance). For example, one component of Scherer's relevance detection is an 'intrinsic pleasantness check' that aims to determine whether the event is more likely to produce pleasure or pain. Scherer suggests that such a check is 'so basic to many emotional responses that affective feeling is often equated with the positive or negative reaction towards a stimulus' (2001). In other words, we have a positive or negative reaction to the feelings$_D$ evoked in the situation, and it is suggested here that those feelings are linked to memories and beliefs about the current scenario *as well as 'as if' projections* of how the individual might feel in possible future resultant states. The complex of memories (factual or perceived as factual), body states, and invoked beliefs of a somatic marker can be easily envisaged as contributing to all the elements of Scherer's model; and it is reasonable to propose that it is the feelings$_D$ that direct attention and greatly influence decisions.



Clore and Gasper's (2000) attentional funnel (see Figure 10) now gives more insight into the process of how attention is directed. Having established that the feelings of emotion direct and focus attention, we can infer, from Clore and Gasper's view, that whatever is part of the somatic marker complex that produces the most intense feelings will have the greatest influence on how the emotional experience unfolds and the decisions that are made because of it. Consequently, it is suggested that the proposed reduction of the goal space and the increased importance of the event (resulting from the reinforcing feedback loop depicted) will greatly influence attention. Then, as is depicted in Figure 9, the beliefs that are engaged in working memory are selected accordingly.



It is possible that both of these models are encapsulated in the somatic marker hypothesis. The detailed nature of the interactions—especially the feedback loops linking memories and beliefs with recorded body states—still awaits discovery. However, what can be safely accepted is the pivotal role of feelings and, more specifically, $feelings_D$.

Returning to the burnt hand example: there is clearly a life-management issue. If the person genuinely is so clumsy, they are potentially a danger to themselves and to others. So, it is no wonder then that beliefs about potential future limitations on their activities in the kitchen might come into play. This could be an issue for who they are as an entity—as a 'self'—in relation to the kitchen environment. But, it extends beyond that: if they can't be trusted in the kitchen, then how can they safely cook for family and friends, and what are the long-term implications for their social life, for example? And while such beliefs and the attendant scenarios are played out, the immediate pain from the burn, along with the emotional pain, captures their attention and acts as confirming feedback that what they believe about themself is true.

4.2 The Self

As noted above, in investigating the neural underpinnings of consciousness, Damasio's attention is strongly on the emergence of the 'self'. He views the 'self process' as indispensable to consciousness with the conscious mind arising 'when a self process is added onto a basic mind process' (2010, p. 8).[68] With respect to this, it is important to emphasise that Damasio sees the self as a *process:*

> There is indeed a self, but it is a process, not a thing, and the process is present at all times when we are presumed to be conscious.

68. There is scope here for arguing that a 'self process' cannot occur until the neural structures enabling self reflection are in place, but that would be a diversion from the focus of the thesis.

(ibid.)

Metzinger, who holds that 'nobody has ever *been* or *had* a self' (2009, p. 1), does allow that for 'ideological and psychological reasons' (ibid., p. 208) the self could be described as:

> … a self-organising and self-sustaining physical system that can represent itself on the level of global availability. The self is not a thing but a process.
>
> (ibid.)

Slaby, too, offers a process view that accords with Damasio and Metzinger:

> … we certainly do not want to claim that there is such a thing as "a self", apart from the organism or person whose self-consciousness is in question. What people usually have in mind when they use the expression "the self" is what constitutes the core of an individual's specific personality: that what makes this individual the specific person that he or she is. And this is not a thing, nor an entity, but rather a (more or less) stable pattern in the sequence of relations this person entertains with her surroundings, its social and physical environment.
>
> (Slaby 2008, p. 507)

Damasio acknowledges the influence of William James' writing, and it is interesting to note the attention that James devotes to the consciousness of the self (James 1890/1950b, Chapter X). It is clear from the last sentence in Chapter X, (Volume 1), that James, like Damasio, is pursuing a biologically-based understanding of consciousness. He proposes two classes of *constituents* of the self: a phenomenal self comprising the material self, the social self and the spiritual self;[69] and the pure 'Ego' that encompasses the sense of personal identity. Northoff (2011), on the basis of extensive literature, suggests that James' material self, social self (mental self) and spiritual self may correspond to distinct brain regions. Here we can see the seeds of Damasio's protoself, core self, and autobiographical self.

Extended philosophical debate on the nature of the self is beyond the scope of this thesis. Clearly, in daily language people perceive themselves as an entity to which they refer to as their 'self'. The stance taken here accords with those above: that there is a process of consciousness that enables one to perceive oneself as an entity in dynamic relationship with the environment and with objects that may be external (in the environment) or may be internal (for example, memories, imagined scenarios and beliefs). Therefore, for the most part, the distinctions of a core self (as defined by Damasio) and an autobiographical, or narrative, self will used be in the remaining discussion. It is argued below that it is this

69. The spiritual self is 'a man's inner or subjective being, his psychic faculties of dispositions, taken concretely …' (ibid., p. 296).

entity—the self —that is at the centre of the conscious and unconscious efforts to maintain beliefs that support the safety and integrity of the self, regardless of the objective truth of those beliefs.

4.2.1 Beliefs from Feelings

In the hierarchical models presented (Figures 13, 14, 15), we can see a relationship that allows for learning and the acquisition of beliefs. With respect to basic homeostatic mechanisms, as long as such hierarchical levels of consciousness are accepted, there seems little doubt that straightforward learning will lead to sustaining beliefs. For example, and using Damasio's labels, thirst will be registered in the protoself and then the core self takes action to drink. Repetition of this leads to learning that the feelings associated with thirst call attention to the need for hydration. As language is added, we can articulate what is understood by the autobiographical self: for example, 'I'm thirsty, I shouldn't go too much longer without drinking'. Implied in this example is the thought, for example, that it is bad for the body to allow oneself to dehydrate too much. On the basis of the selected definition for belief (page 27), this thought *is a belief*: it is stated in a way that asserts that it is true; it is driven by an underlying need/want (in this case, a homeostatic need); and it is guiding action. It is also indicative of the point made earlier that it is signalling the need for a decision: in this case, that is a decision about the timing of drinking. In each instance of feeling thirsty, it is the feelings of thirst that prompt action. It is reasonable to suppose also that, subsequent to drinking, the feelings of relieving thirst are injected into a feedback loop that confirms that the body feels better—and by implication *is* better—when it is not dehydrated. Such an effect would be even greater if the person was genuinely dehydrated: the feelings of relief not only signal that the body is now more homeostatically balanced but those same feelings could be an inducer for the somatic marker carrying memories of previous occasions. This marker, in turn, may bring to mind the belief that 'I really do have a bad habit of working too long without drinking'. Then, the associated memory of feeling dehydrated and the feeling of relief are experienced evidence for that belief. As Clore and Gasper put it:

> Knowing something directly from first-hand experience (experiential knowledge) is a different form of knowledge than knowing about the same thing indirectly (propositional knowledge) … Of these two forms, experiential knowledge often has priority.
>
> (Clore & Gasper 2000, p. 24)

They go on to conclude that'beliefs are adjusted to be compatible with internal evidence in the form of feelings, just as they are adjusted to be compatible with external evidence

from perceptual experience' (ibid., p. 25). Similarly they provide a succinct reminder of a point that has been made previously:

> Emotional feelings guide attention, not so much to themselves, but to what they signify.
>
> (ibid., p. 31)

In the hypothetical but realistic example above,[70] it is interesting to note that the person did not simply drink some water without comment, and then proceed to their next task; they attended to and articulated their belief about their bad habit of allowing themself to dehydrate. We cannot know what intricacies of appraisal occurred over repeated such instances that caused them to create that specific belief. However, it is significant that they did not put the wording some other way (for example, 'I'm so good at focusing on my work that I loose track of time'). What can be reasonably proposed in this case is that the feeling of being dehydrated followed by the relief-induced feelings confirmed the belief.

With both the drinking example and the burnt hand example, it is relatively straightforward to accept that a few repeated instances and the corresponding feelings, provide information to the brain that instigates beliefs that, in the first case, it is good to remain hydrated and, in the second case, it is wise to handle hot saucepans carefully. This may really be all that is *required*. Such beliefs could guide actions that would maintain bodily homeostasis and safety. But, in both cases, there are utterances of beliefs about the people themselves. In both cases, the response to the emotional experience that arises is to attend to their beliefs about life management. It is proposed that it is the somatic markers that carry the feelings and memories that the self-reflective processes use to create these beliefs.

Now, there is the issue of how feelings_D can maintain or even strengthen beliefs that appear to be of no service to the person by virtue of being:

- true up to the current time, but debilitating or limiting if maintained
- implausible, or
- false as shown by evidence.

4.2.2 Feelings of Self

In the literature reviewed, the role of feelings became an inescapable topic. Bearing in mind that the literature is predominantly focused on a neuroanatomical understanding of emotions and their effects on beliefs[71], that is not surprising. It is also even less surprising

70. The author has heard people say these things.

71. Rather than a strict philosophical approach.

given that if we are conscious of any emotion, feeling the emotion is also unavoidable (assuming no bodily or neural pathologies). Given that we are usually *aware* of ourselves, that awareness necessarily includes feelings of relationship of our body with other objects in our environment and internal objects (thoughts, memories).[72]

In investigating what he calls the 'stream of thought', James (1890/1950b) gives bodily feelings a most significant role:

> Our own bodily position, attitude, condition, is one of the things of which *some* awareness, however inattentive, invariably accompanies the knowledge of whatever else we we know. We think; and as we think we feel our bodily selves as the the seat of the thinking.
>
> (ibid., pp. 241–242)

Later, he adds:

> There is not a conjunction or a preposition, and hardly an adverbial phrase, syntactic form, or inflection of voice, in human speech, that does not express some shading or other of relation which we at some moment actually feel to exist between the larger objects of our thought …
>
> We ought to say a feeling of *and*, a feeling of *if*, a feeling of *but*, and a feeling of *by*, quite as readily as we say a feeling of *blue* or a feeling of *cold*.
>
> (ibid., pp. 245–246)

Ratcliffe (2005) strongly defends James' view of the role of feelings against the view taken in some philosophical approaches. As an example he opposes Ben Ze'ev's (2004) view that 'the feeling dimension has no significant cognitive content' and that feelings are not intentional (amongst other aspects of Ze'ev's views). Ratcliffe builds a case that feelings are intentional; that there is a difference between the location of a feeling and what the feeling is *'of'*; and that it is *'through'* feelings that we become conscious of other things. Ratcliffe's view is introduced here as it is largely in accord with James and Damasio (with one apparent misunderstanding of Damasio)[73] and his argument for 'existential feelings' leads to a greater understanding of the possible role of feelings and somatic markers in the beliefs that support the autobiographical self.

---

72. Excluding any neuropathological conditions that reduce or eliminate physical awareness, of course.

73. He suggests that Damasio 'interprets feeling as "the perception of a certain state of the body" … [and] thus omits the possibility that the body need not be the object of perception and that feelings can manifest themselves through the manner in which another object of perception, or the world in general, is perceived' (Ratcliffe 2005, p. 53). However, he appears to lack a full understanding of Damasio's first-principles definition. An external object will be perceived through the 'sensory portals' (as Damasio calls them), causing a change in the protoself. The feeling from the core self produced by the changes in the protoself *is* the means for perceiving another object whether that object is a recalled memory or an external object (e.g. the sight of a friend approaching).

To clarify what is meant by 'existential feeling' and to show how bodily feeling can contribute to experience as a whole, Ratcliffe invites consideration of visual perception, but in isolation from proprioception and other senses.

> If one were to model one's general conception of intentionality on the visual modality, considered in abstract isolation from proprioception and the other senses, a distinction between perceiving things in the world and having bodily feelings might well seem plausible. When one perceives an extremely desirable object X and has a bodily feeling Y, one has two distinct perceptions, the perception of X in the world and the perception of Y in the body. However, a different story is to be told if touch is taken as a starting point for phenomenological exploration and a more general account of bodily feelings is modelled on this modality. In the case of touch, bodily feeling and experience of the world are inextricable.
>
> (ibid., p. 47)

He goes on to suggest that 'to touch is to experience a relation between one's body and an object it comes into contact with' (ibid., p. 48). Ratcliffe is suggesting that *tactile* feeling contributes to a sense of self and other, thus contributing to a sense of the relationship of the self with the 'world as a whole' (ibid., p. 49). While this point is accepted, it does not seem to be appropriate to contrast tactile perception to visual perception in the way that he does to make the point. Damasio proposes that:

> Signals from the outside are thus *double*. Something you see or hear excites the special sense of sight or sound as a "nonbody" signal, but it also excites a "body" signal hailing from the place in the skin where the special signal entered …
>
> When you see, you do not just see: *you feel you are seeing something with your eyes*. Your brain processes signals about your organism's being engaged at a specific place on the body reference map (such as the eyes and their controlling muscles), and about the visual specifics of what it is that excites your retinas.
>
> (Damasio 1994, p. 232)

James, too, suggested that he could not think in 'visual terms, for example, without feeling a fluctuating play of pressures, convergences, divergences, and accommodations in my eyeballs' (James 1890/1950a, p. 300). With tactile input, the felt object might just be a glass; but one can indeed sense oneself in relation to the glass but separate from it, as Ratcliffe suggests. However, sensing oneself in relation to another object is also the case for seeing and hearing. For example, we can see an object and the fact that we can feel ourselves seeing allows us to make the distinction between ourselves and the object; the object is not some extremely lifelike hallucination from within. The brain has the role of *'representing the outside world in terms of the modifications it causes in the body proper …* whenever an interaction between organism and environment takes place' (Damasio 1994, p. 230). Part of that is ensuring that there is constant background monitoring of the body

state so that we know where our boundary (for the body) is (ibid.). Whether an external object is perceived through tactile interaction or through visual, auditory, olfactory or gustatory detection, the resulting changes in the protoself and core self allow the distinction between self and other. So, it seems that Damasio's description supports Ratcliffe's point more than Ratcliffe allows.

A difference between Ratcliffe and others' (notably, Prinz 2004a; Prinz 2006; Slaby 2008) view of feelings from Damasio's is whether bodily feelings are intentional. They claim feelings are intentional and that Damasio is not making that claim. Pursuing that debate is not a focus for this enterprise; what is significant is the general agreement with the view that emotions signal matters of *concern* and that the relationship of the self to the world/environment is important—both of which points are in agreement with Damasio. This point is succinctly emphasised by Prinz:

> Emotions are designed to provide information about our relation to the world. They represent concerns, and they do so in a way that captures our attention and influences our behaviour. By drawing attention to concerns, emotions disrupt our ongoing plans and force us to adopt coping strategies.
>
> (Prinz 2006, pp. 154–155)

Therefore, given that the human organism perceives that it is a 'self', the signalling provided by the emotions is likely to relate to more than just physical safety such as running from attacking animals and handling hot saucepans carefully. Thinking of oneself in relation to the environment carries feelings that are about the *self*. If we are being chased by an attacking animal, rather than saying 'The body is in danger' or 'My body is in danger', the more likely cry is '*I* am in danger'—the attention is on protecting the self.

The retort here could be that this is simply avoiding an extreme disruption of homeostasis (a chewed and mangled body is severely out of balance). However, it is important to recall the implications of Damasio's 'social' emotions (1999, 2003a, 2010) (introduced in Section 3.2.3.1) and Ratcliffe's existential emotions (2005). Amongst the social emotions Damasio includes compassion, embarrassment, shame, guilt, contempt, jealousy, envy, pride, and admiration (2010). Ratcliffe's examples[74] include the feeling of being 'complete, flawed and diminished, unworthy, humble, separate and in limitation, at home, a fraud, slightly lost, overwhelmed, abandoned, stared at, torn, disconnected from the world, invulnerable, unloved, watched, empty, in control, powerful, completely helpless, part of the real world again, trapped and weighed down, part of a larger machine, at one with life, at one with nature, there, familiar, and real' (Ratcliffe 2005, p. 45). Some of Ratcliffe's examples could well be subsumed under Damasio's social emotions or his background

74. Obtained by doing an internet search on the 'feeling of being'!

emotions and feelings (1994, 1999). At first sight, these might not be considered to be related to the kinds of emotions that alert us to the need for decisions and action. However, they clearly signal states of pleasure or pain, though of varying intensity. Many of the feelings, then, are signalling some somatic state arising from the marker associated with the object. In this case, we can safely assume that a most likely object is the self as it is in relation to society and/or the world.

James (1890/1950b) suggests that to understand the self, it is useful to consider three parts: its constituents (page 174); the *self-feelings* which are the emotions and feelings aroused by the constituents; and *self-seeking* and *self-preservation* which are actions prompted by the self-feelings. Now there is a case to say that feelings and the underlying emotions can be viewed as signals guiding decisions and actions that further the cause of the self. Once the self is being considered in the social context—with social emotions and existential emotions, if they be granted—the autobiographical self (or third-order processing, or narrative self) is necessarily the focus of attention.

4.2.3 Preserving the Autobiographical Self

Once again it is James who shows a way to helping unravel the interactions between self, feelings and beliefs. Of self-feelings, James considers there are primarily two: 'self-complacency' and 'self-dissatisfaction' (1890/1950a, p. 303). For James, the former has synonyms such as pride, vanity, and self-esteem while the latter has synonyms such as modesty, humility, confusion, shame, and contrition. In these, we can see Damasio's social emotions. Interestingly, James views these emotions as being as primitive as rage or pain (ibid., Chapter X) because each exhibits its own characteristic physical expression. Damasio's model (depicted in Figure 4) suggests that there is likely to be a contribution to these social emotions from higher order processing of the autobiographical self (in Damasio's model). In either case, James's view is interesting: that these emotions 'cover a large number of our fundamental instinctive impulses' (ibid., p. 307) and they include emotions of bodily self-seeking, social self-seeking, and spiritual self-seeking. Reading more deeply into James it can be argued that the self-seeking and self-preservation behaviours entail higher order processing that shapes the beliefs associated with them.

Bodily self-feelings are clearly related to physical preservation; but James also suggests that emotions such as anger and fear can be included with 'instincts' such as hunting and the construction of homes that seek to provide for the future. Consequently, he considers they contribute to the development of the self in the 'widest possible sense of the word' (ibid., p. 308.) That is James' segue to the social self-seeking which he describes as:

> … our desire to please and attract notice and admiration, our emulation and jealousy, our love of glory, influence, and power, and indirectly through whichever of

> the material self-seeking impulses prove serviceable as means to social ends. That the direct social self-seeking impulses are probably pure instincts is easily seen.[75]
>
> (ibid.)

Seeing the social self-seeking impulses as 'pure instincts' can be interpreted as being analogous to Damasio's view of the social emotions being built on homeostatic mechanisms and those emotions that direct our attention to survival-related matters. Now, this points to the position of the self in one's society (environment, family, etc.) as being a matter of homeostatic equilibrium or disequilibrium, of physical survival, and—most significantly—of survival of the narrative self (which entails James' social self). James suggests that people arrange their selves—bodily, social, spiritual—hierarchically according to worth. At a basic level, we can concede this point immediately: if one is ejected from one's society, one's physical existence can be directly threatened by lack of food, shelter and protection. So too, though, one's social self—or existential self (Ratcliffe 2005)—which James suggests ranks higher than bodily self in the hierarchy, can be threatened.

Yet again James pre-empted a number of the concepts in the current literature. He suggests that the thoughts in others' minds about oneself, despite being external to oneself, are objects of interest to one's autobiographical self ('*social* self-love' in James' terminology). And, although another's thoughts are external to oneself, James claims to perceive changes in their thoughts 'just as I perceive any other outward change' (1890/1950a, p. 321). Using the outward appearance of another to vicariously experience what is occurring within their thoughts is supported to some extent by the research into mirror neurons (Rizzolatti & Craighero 2004; Rizzolatti 2005; Rossi & Rossi 2006). Beyond reading the other's view of him, James suggests the complex of reactions that we would expect from Damasio's somatic marker.

> But the pride and shame which I feel are not concerned merely with *those* changes. I feel as if something else had changed too, when I perceived my image in your mind to have changed for the worse, something in me to which that image belongs, and which a moment ago I felt inside of me, big and strong and lusty, but now weak, contracted and collapsed …
>
> … it is as being an I who has always been treated with respect, who belongs to a certain family and 'set', who has certain powers, possessions, and public functions, sensibilities, duties, and purposes, and merits and deserts. All this is what your disdain negates and contradicts; this is 'the thing inside of me' whose changed

75. Further on, in a quote too entertaining to omit, James gives insight in the 19th century equivalent of the modern social media frenzy. 'We are crazy to get a visiting-list which shall be large, to be able to say when any one is mentioned, "Oh!, I know him well," and to be bowed to in the street by half the people we meet. Of course distinguished friends and admiring recognition are the most desirable—' (ibid., p. 308)

> treatment I feel the *shame* [emphasis added] about; this is what was lusty, and now, in consequence of your conduct, is collapsed; and this certainly is an empirical objective thing. Indeed, the thing that is felt modified and changed for the worse during my feeling of shame is often more concrete even that this—*it is simply my bodily person* [emphasis added], in which your conduct immediately and without any reflection at all on my part works those muscular, glandular, and vascular changes which together make up the 'expression' of shame.
>
> (James 1890/1950a, pp. 321–322)

In this there are all the elements of the somatic marker: an emotionally competent stimulus (the disapproving look of another, in this case); the bodily (somatic) response; an imagined image of what the other's image is; and implied judgments of himself as evidenced in the resulting emotion of shame. It can be assumed too that since the somatic marker brings to working memory the memories of past similar experiences and related beliefs, there will be considerable interactions between lower-order processes and higher-order processes.

Any number of beliefs could be created or changed in such a reaction. For example, they may now have further proof that they are shameful in some way; the proof being partly the disapproving countenance of the other person, but mainly the intensity of the bodily feelings. Alternatively, they may become angry, indignant at being unfairly judged and decide—that is, come to believe—that the other person is unpleasant, unkind and not worthy of acquaintance.

The second belief of these two highly probable examples could be considered as deriving from taking a defensive stance. If self-preservation is a strong drive, as suggested in James (ibid.) and Damasio, it can easily follow that assuming a belief that blames the other person is a form of defence. Such a belief can prepare someone to be more assertive, or even aggressive; to be more prepared when dealing with similar other people; to be less subject to feelings that 'weaken' one; and, most importantly, maintain a coherent perception of one's own integrity and thereby protect one's social self. The underlying proposition here is that such a belief is formed to protect the structural integrity, metaphorically speaking, of one's own perceived—or believed—autobiographical self. In the example given, if the feelings-as-evidence hypothesis holds, then the feelings of anger and indignation are experienced (felt) proof that the view held of the other person is true.

However, a problem remains with the first example response. How is it that the person—indeed, any person—takes on a belief that diminishes themself in their own eyes and is almost certainly likely to limit their perception of the options available to them in society? Such a response hardly seems to bear the marks of self-preservation or self-protection. It is no wonder, then, that Forgas says:

> Beliefs about the self represent a particularly complex, elaborate and problematic domain …
>
> (Forgas 2013, p. 7)

In considering the effects of society, Cromby agrees with Day and Coleman (2010) who perceive a problem when belief is treated outside its cultural context. Day, in fact, concludes that 'belief is not separate from identity or social context but a way of creating who "I" am relative to "you" here and now' (2010, p. 26). This view appears to fit well with the James and Damasio models relating to the autobiographical self and social emotions. Cromby, picking up from James and the concept of the 'feeling of knowing',[76] offers the suggestion that beliefs are 'a structure of socialised feeling, contingently allied to discursive practices and positions' (2012, p. 945). The discourse, which includes thinking (as in an internal discussion), is how Cromby sees belief being 'worked up rhetorically [to] provide both its core content and nuances' (ibid., p. 946). But the discourse commingles with feelings of which Cromby distinguishes three types. Emotional feelings he suggests are the 'embodied components of the enculturated complexes of discourse, intention and action' (ibid., p. 948); extra-emotional feelings are those associated with urges such as hunger and with responses such as pain; and, feelings of knowing are those such as the feeling of *but*, for example. With respect to the last of these, Cromby refers to the James quote given above (page 177) and Johnson's (2007) development of the idea that feelings 'help organise and direct our thinking' (2012, p. 948). While Cromby's distinctions of feelings vary from Damasio's distinctions for types of feelings (1994, 1999), they provide the opportunity to suggest a significant role for feelings of knowing in apparently dispassionate discourse:

> … even the most abstract forms of reasoning, indexed by seemingly content-free words such as 'but', 'and' or 'not', rely upon feelings of (for example) obstruction, connection, disjunction or direction. These feelings help organise and direct our thinking so that when, for instance, we want to say 'but' in response to a claim, the urge to do so is not solely a discursive move. It is simultaneously the verbal enactment of a felt sense of obstruction to the smooth movement of thought, a subtle embodied feeling that something is extraneous, missing or wrong.
>
> (Cromby 2012, p. 948)

Cromby suggests all his types of feeling are socialised, pointing most significantly to how feelings shape expression of emotions that are considered to be *basic* emotions defined by biology (Ekman 1992, 2003): anger can be tight-lipped or be furiously expressive, for

76. These 'feelings of knowing' are different from those that Damasio (2010, p. 9) uses and which refer to the distinguishing of self from non-self.

example. This is offered by Cromby as support for the socialisation view, as are the different responses to pain in different contexts and cultures.

While the value of Cromby's definition of belief is debated (Forgas 2013), he draws a conclusion that is aligned to the development of the argument that the social context will influence beliefs by the emotions and feelings that occur in it; namely, that beliefs are 'not singular cognitive entities: they are complex, variegated habits of felt thinking: far more contingent upon the flows of social practice, and far more rooted in our bodies, than psychology usually allows' (2012, p. 952). Forgas does concede that Cromby's view of the role emotions in shaping beliefs is 'interesting' and 'largely consistent with recent evidence suggesting that affect is involved in all cognitive activity, including mental representations of the social world' (2013, p. 5). This, then, suggests that the examination of the effects of emotions and feelings$_D$ on the autobiographical self in context is worth further pursuit.

4.2.3.1 Feelings in the Social Context

Johnson (2007) is greatly supportive of Damasio's view of the role of emotion[77] and reinforces the support cited from Prinz (2004a, 2005, 2006) in Section 4.1.2.1 and the general case for the role of emotion and feeling presented in Chapter 3.

> Emotions are key components of complex processes of bodily perception, assessment, internal monitoring, self-transformation, motivation, and action.
>
> (Johnson 2007, p. 65)

The variance that Johnson appears to have from Damasio is in considering that 'emotions arise from and are directed out into the world' (ibid.) while he interprets Damasio as suggesting that emotions are only about a person's body state. This interpretation seems to miss the significance of Damasio's definition: the emotion is about the body state *but* it can arise not just from homeostatic imbalance but also from the events in the external world such as physical threat (for example, a charging lion) or from perceived threat (for example, the boss saying *'Can you come to a meeting to discuss the business's forthcoming restructure?'*). Moreover, the ability to recall past similar experiences and the associated beliefs allows a person to engage an 'as if loop' to imagine how they might act *and feel* in a possible future situation. So, Damasio's insistence that the emotion is the body state in no way precludes an outward direction of attention. In fact, as amply shown in Chapter 3, the feelings of the emotion will direct attention depending on the object of the emotion.

Usefully, however, Johnson argues a case arising from Dewey's work that it is *'situations* that are characterised by emotionality, rather than merely a person's mind or psychic state'

77. Apart from some issue about potentially misleading use of some terminology (e.g. 'image')

(ibid., p. 66). Johnson clarifies this by suggesting that 'body states connected with feelings are states of both response to and remaking of experience' (ibid.). In a given situation, he suggests that one could say *'I am frightened'*, for example, but also say *'The situation is frightening'*. It is not clear if Johnson intends to support Damasio's somatic marker hypothesis, but his discussion does. There are external conditions that present a clear danger and these invoke a somatically marked body state; the feelings of this state draw attention to the conditions while memories and beliefs relevant to the situation are brought to working memory in order to guide interpretation and decisions. Such a situation can apply even when there is no genuine danger: for example, seeing a highly venomous snake safely behind glass in a zoo still does not stop the feelings of fear for many people.

Beyond immediate physical danger, perceived danger in the social context can evoke similar outward attention. For example, the concern that a person's job might be threatened by an imminent organisational restructure[78] frequently evokes strong emotions. The word 'restructure' is conceivably sufficient to act as an emotionally competent stimulus that triggers a somatic marker with a complex of emotions containing, in all probability, fear, anger and some indignation at yet more unnecessary rearrangement. The 'as if loop' will then feed off imagined scenarios involving a feared long period without regular work or pay and the attendant *feelings* will help confirm the beliefs, for example, that organisations treat their staff unfairly, that the government does nothing to maintain social justice, and that the individual is bound to have trouble finding new work in the current economic climate. The emotions and feelings are at the core of this complex of responses; and attention is directed both inward and outward. What is proposed here is that once the feedback loop is set up between the imagined and undesirable scenarios and the feelings of the emotions evoked, the feelings can be the key force in maintaining the emotional response and in confirming existing beliefs and, possibly, directing attention to new information that contributes to forming new beliefs (for example, *'I never thought that that particular manager would treat us like that; they must be the same as the others'*). Therefore, the interest is now on the autobiographical self (encompassing the 'social self') in its social context. Johnson's argument leads the way in shifting the focus:

> … once we see that emotions exist precisely because of the ways they are connected to our shared world and permit us to function within it, then it becomes possible to recognise their crucial roles in our communal well-being.
>
> (ibid.)

78. A curiously regular and frequent phenomenon.

Johnson's conclusion that 'emotions exist precisely because of the ways they are connected to our shared world …' might not withstand close examination; but the overall significance of them in shaping the well-being of the autobiographical self in its social context must be granted.

Section 3.4 examined the general agreement in the literature that emotion affects cognition and that the two are tightly interrelated. These points do not need further discussion other than to note that the support in the literature for them continues to amass (for example: Bless & Fiedler 2006; Clore & Huntsinger 2007; Clore & Ortony 2008; Clore & Palmer 2009; Clore & Storbeck 2006; Gasper 2004; Gasper & Bramesfeld 2006; Panksepp 2005a, b, 2012; Phelps 2005, 2006; Phelps *et al.* 2006; Prinz 2004a; Scherer 2004).

From a model such as that offered by LeDoux in Figure 11, it can be hypothesised that in a case involving fear the 'low road' gives primacy to feelings. The initial reaction, first consciously experienced as the feelings of fear invoke memories and beliefs associated with the situation. Even in a situation where there may be a quick realisation that there is no threat (for example, it is not a snake, it is a small branch), the lingering feelings of fear can still affect the next actions (for example, *'Ok, that's not a snake, but I'll bet there are some around; so let's get out of here'*.)

Prinz (2006), in arguing that perception is 'quasi-modular',[79] suggests that:

> … we cannot, under ordinary circumstances, have a perceptual representation generated top-down when the current stimulus is disposed to induce an incongruent representation.
>
> (ibid., p. 140)

This is associated with his pertinent response to Fodor's view that 'knowledge cannot penetrate the perception system' (ibid.):

> An alternative explanation is that bottom-up inputs *trump* trump top-down inputs when the two come into conflict.
>
> (ibid.)

Here we have echoes of Clore and Gasper's view that 'one can argue with logic, but not with feeling' (op. cit.). In the same vein:

> The intensity of a belief cannot be counted on to reflect its supporting evidence any more than its causes can.
>
> (Quine & Ullian 1978, p. 16)

79. A discussion beyond scope of this thesis.

4.2.3.2 Autobiographical Beliefs

In Section 1.2, the issue of discriminating memory from knowledge and belief was discussed. The epistemological debate over these terms is extensive. As well, there is a growing literature from the neurosciences regarding differing neural circuits that may account for an individual's distinction between knowledge and belief. To extend either of those enquiries at this stage is not required for the focus of this thesis.[80] However, reviewing some salient points assists.

A proposition is accepted as knowledge if:

1. the belief or the proposition is held to be true
2. available data are examined or interpreted to justify the claim as knowledge rather than belief. The data may be perceived external events or may be internal processes.

It was decided to take the criteria for knowledge to be those outlined on page 33 (and shown in the side box for easy reference). However, the issue was noted of someone saying *'I **know** it's the case'* when the only evidence that they have is what is in their imagination and the accompanying feelings. Eichenbaum and Bodkin (op. cit.) acknowledged the difficulty in separating knowledge from belief (see page 35). Based on their work and Davies and Colheart's (2000) work, it was determined to accept the use of 'know' (for example, *I know it*) to be equivalent to 'believe (strongly)'.[81] A further justification for this position is that treating this *experience* of knowledge fits with the definition of belief selected to work with (see page 27): it is taken to be true; it may be related to desires and wants; and it can inform decisions and influence preparation for action.

An important aspect in employing the definitions for belief and for knowledge, is the distinction of viewing evidence for beliefs or knowledge *from the perspective of the speaker* (introduced on page 38). This allows for criterion two of the knowledge definition to be satisfied by internal processes. Of course, those internal processes may be accurate memories, to the extent that such things exist. More importantly, in the light of the feelings-as-evidence hypothesis, the examined data for justifying the belief or knowledge may simply be feelings.

These approaches were guided by the exhortation from Westbury and Dennett (2000) to account for context. For example, under some circumstances a person may recall that they attended a particular dinner party and that recollection may fit only some of the criteria for belief: they might have a diary entry so they 'know' they were there because of that entry along with some recalled representations (images of who else was there, what some of the food was, recalled snippets of some conversations). However, in this case there is no particular link to the desires or wants, nor is there any preparation for action; therefore, in

80. However, as noted in the concluding chapter, these areas are ripe for further research.

81. See page 38

using the selected definition for belief this memory barely qualifies. But, now suppose there is a recollection of an argument from the another dinner party and the person says something like *'Fred is a seriously obnoxious person, the way he spoke to me at the dinner party and made me feel was unforgivable – I still get upset when I think about it; one thing's for sure, I know I'll never go to any party again that he's going to be at*'. In this case, we have what may be classed as a memory but under challenge to the accuracy of the memory, the person may respond with *'I **know** what happened; I'll never forget what he said, the way he said it, and the way it made me feel'.* So, we have a declaration of knowledge with evidence being the recalled representations of the event and, in particular, the recalled and reawakened feelings—notably all internal processes at this stage. This memory fits the criteria for a belief: they consider they have evidence; it is related to the desire to be treated a certain way; and there is a preparation for action in avoiding the offender's presence again. From the perspective of the speaker, it is also considered to be knowledge.

A contrasting type of belief was discussed previously (see page 40) in which there is a lack of contextual detail to support clear memory. Someone might say, for example, that they believe they were at an event but that they are unable to recall any details. Sufficient friends may attest that the person was there to support that person's belief that they were present despite their have minimal or no recollection. This is a very different experience from the 'certain knowledge' (based on recalled images and feelings) that Fred (in the first example) is an unpleasant person.

Given these examples, the approach of Eichenbaum and Bodkin (discussed in Section 1.2) of treating both knowledge and belief as memories is more pragmatic than seeking a strict delineation between autobiographical belief/knowledge and autobiographical memory. It is considered that treating these terms according to the context permits development of the thesis and does not affect the strength of the case presented.

In the first example (Fred, the offender), we have all the elements covered so far. The mention of the offender's name, for example, is an emotionally competent stimulus which triggers the somatic state with all its associated memories and beliefs. Implied in this account is the metacognitive belief that the person can trust their account of events based largely on remembering their feelings. Also implied is a belief, or set of beliefs, about how people should treat each other. It is interesting to postulate, too, that the declaration *I'll never forget what he said* also reflects a belief in the ability to maintain a particular attitude to the event and the other person. So, here we also have the emotional experience drawing attention to and directing decisions about the speaker and their relations in that particular context. Beyond this, it is quite probable that there is an autobiographical belief about their 'I': for example, *'At least, I'm not like him'*—implying kinder, or more understanding, or better-mannered, or more civilised.

This draws attention back to the issue of the 'self' and autobiographical beliefs. The catalyst that initiated this thesis is the interest in autobiographical beliefs with the characteristics described in the Introduction: mainly that they are limiting to the person's options for development and action by virtue of being false, implausible, or falsely formed. Much of the discussion so far has been around beliefs in general and the influences that act to facilitate their creation, maintenance or dissolution. However, beliefs 'in general' belong to the sets of belief that any individual can hold that determine how they perceive their context (and what they believe about it) and what they perceive and believe about themself in relation to their context.

4.2.4 The Self in Context

A question was posed earlier (page 182) as to why a person might develop and even strengthen a belief about themself that can be limiting or, at the extreme, debilitating or endangering. While we might more easily understand how someone might develop self-aggrandising or narcissistic beliefs, beliefs that lessen the person in their own or others' eyes, especially if they are patently false, can defy credulity. No wonder, as already noted, that Forgas says 'beliefs about the self represent a particularly complex, elaborate and problematic domain' (2013, p. 7).

In the literature reviewed there is agreement that there is no such thing as a 'self'. As already established, what is perceived as oneself is the result of the processes from which consciousness emerges. Damasio (2010) suggests two perspectives on the self: the 'self-as-knower' and the 'self-as-object', this latter being:

> a dynamic collection of integrated neural processes, centred on the representation of the living body, that finds expression in a dynamic collection of integrated mental processes.
>
> (ibid., p. 9)

The self-as-knower—the 'I'—Damasio admits, is a more elusive concept and despite being 'a very real presence' it is 'more difficult to capture than the plain me' (ibid.). These distinctions appear to be inspired by James' view that there is a 'material me' and an Ego which is the sense of personal identity (James 1890/1950b). James, too, was more challenged in addressing the Ego—presumably the equivalent of the 'I'. He admits that the Ego was a topic he had 'always shied away from and treated as a difficulty to be postponed' (ibid., p. 330).

Metzinger introduces the term *phenomenal self-model* (PSM) which he describes as 'the conscious model of the organism as a whole that is activated by the brain' (2009, p. 4)—a description clearly aligned with Damasio's views of core consciousness and the

autobiographical self. He captures the ephemeral nature of the Ego/self-as-knower in suggesting that this conscious self-model can represent the process of representation itself so that 'we can catch ourselves—as Damasio would call it—in the act of knowing' (ibid., p. 5). When Metzinger uses the term 'Ego' he says it is:

> … simply the content of your PSM at this moment (your bodily sensations, your emotional state, your perceptions, memories, acts of will, thoughts). But it can become the Ego only because you are constitutionally unable to realise that all this is just the content of a simulation in your brain.
>
> (ibid., p. 8)

Previously, it was also noted that Metzinger described the self as a 'self-organising and self-sustaining physical system that can represent itself on the level of global availability' (ibid., p. 208). It is suggested here that the words ***self-sustaining*** are highly significant: if the autobiographical self—the narrative self—is indeed built upon the processes that allow awareness of, say, the protoself and core self, then what we experience as social emotions can be viewed as the mechanisms that guide the passage of the autobiographical self safely through its environment. In other words, these emotions could conceivably be viewed as, for want of a better term, higher-order *homeostatic alerts attending to the coherency and security of the 'self'*.

If I am embarrassed or made fun of in my social setting, in most cases it is unlikely that I will be in physical danger; but the somatic markers are triggered regardless. The physical reaction may even be that I want to run away; and it may even include the experience of pain—despite the lack of physical danger. What is in danger is *who I am*: will my friends still regard me well; will I still get work with these people; will my overall security be threatened because someone regards me less? While such considerations are occurring, the brain can be busy running simulations of what is supposed to be occurring in the others' minds (based on what I am seeing and hearing) and simulations of possible future undesirable and uncomfortable situations. Of course, the signal that the current situation and imagined future scenarios are undesirable is the *feelings*$_D$.

While Chapter 3 clearly established the influence of emotions and mood on cognition, it is notable that development of our understanding of this interaction continues, especially with respect to the differing effects of positive and negative affect. Forgas (2013, p. 6) observes that positive affect appears to encourage 'more superficial processing strategies', whereas negative affect appears to utilise a 'more effortful, systematic, analytic and vigilant processing style'. Drawing on the work of Fielder and Bless (2000) and Bless and Fiedler (2006) he suggests that:

> … positive affect, signalling a safe and familiar environment, generally promotes a more *assimilative*, schema-based, top-down thinking style, where pre-existing beliefs,

> attitudes and representations dominate thinking. In contrast, negative affect functions like a mild alarm signal and produces a more accommodative, bottom-up and externally focused information processing strategy where attention to situational information drives thinking.
>
> (Forgas 2013, p. 6)

This finding agrees with the types of effects reported in Chapter 3 and it suggests how it might be that someone makes or strengthens beliefs that seem to be counter to their best interests and well-being. Bless and Fielder (2006) hold the view that people need access to regulators that are neither too complex nor strenuous in order to be able to adjust quickly and smoothly to circumstances. Hence:

> Feeling good suggests assimilating the environment to the organism's internal state. Feeling bad suggests accommodating the internal state to the requirements of the problematic external state.
>
> (ibid., p. 79)

Reviewing work by DeSteno and Salovey (1997), Forgas notes that when people are in *neutral* affective states, descriptive features such as 'achievement' and 'affiliation' tend to be organising influences for beliefs about the self. But, *positive or negative affect had the effect of structuring self-beliefs in terms of feeling good or bad*. As observed previously, he suggests that 'affect may function as a key organising principle of self-related beliefs' (2013, p. 7). Similarly, work by Trope et al. (2001) supported a mood-as-a-resource hypothesis and showed, for example, that feeling good can help overcome defensiveness and help a person be more willing to accept information that may lessen their positive mood: 'It seems, then, that instead of trying to maintain a positive mood, people are willing to exchange their positive mood for useful information about themselves' (ibid., p. 272).

In Forgas's (2002) view, the complexity of many social situations makes affective influences more likely in shaping interpersonal behaviours.

> The principle appears to be that the more complex and ambiguous a social situation, the more likely it is that people will need to engage in open, elaborate, and constructive thinking, drawing on their own memory-based ideas to produce an appropriate response.
>
> (ibid., p. 2)

Forgas (2000c, 2002; and also Forgas & Ciarrochi 2002) goes on to present the concept of *homeostatic mood management* based on work by Sedikides (1994) and work of his own[82] which showed an alternation between affect infusion and affect control. Affect infusion

82. Cited as being in press at the time.

(substantive processing) maintains and amplifies the current state while affect control (motivated processing) can reduce the current state; and Forgas suggests that people alternate between these two processes to maintain homeostatic mood management (see Figure 18).

As established in Chapter 3, Forgas argues a strong case for the way in which the Affect Infusion Model accounts for the effects of emotions and feelings in given situations. However, an interesting enquiry is to ask how the 'memory-based ideas' (quoted above) come into consciousness for people to use in complex and ambiguous social situations. It would seem that Damasio's somatic marker provides at least part of the answer since the feelings involved are associated with context-relevant memories and beliefs. Given the proposed operations of a somatic marker (see page 166), it is suggested here that the information available to the system (that is, the body and brain of the individual) that is associated with one or more somatic markers, provides the information used (accurately or not, wisely or not) to address each of the decision points in Forgas's flowchart for deciding which type of processing to use in making a judgment. In fact, if the somatic marker operations are accepted, it is difficult to conceive how they would not be involved at each of the decision points (familiar prior judgment, relevance, specific motivation, target characteristics, affect state, motivation for accuracy), with the possible exception of 'cognitive capacity'.[83]

However, the question posed on page 182 and recalled at the beginning of this section still has not been answered: why create or strengthen a belief that lessens oneself or one's opportunities, especially if it is implausible or false? Relatively recent literature on self-enhancement and self-protection sheds additional light on this issue.

4.2.4.1 Self-enhancement and Self-protection

Alicke and Sedikides define self-enhancement and self-protection as:

> … interests that individuals have in advancing one or more self-domains or defending against negative self-views
>
> (Alicke & Sedikides 2009, p. 1)[84]

Or, to put the terms more simply:

83. Development of this line could be quite extensive. However, given the reported research demonstrating the role of emotions and feelings, and the broad (albeit not complete) support of Damasio's hypothesis, this is the stance taken herein.

84. There appears to be a relatively recent burgeoning of literature in this area . (e.g. Alicke & Sedikides 2011a; Sedikides & Spencer 2007). The overview provided by Alicke and Sedikides is the most pertinent for this thesis: going into the details of motivations is beyond the focus and required scope.

> Although self-enhancement and self-protection are complex motives, they have at their roots the assumption that people want to feel good, or avoid feeling bad, *about themselves* [emphasis added].
>
> (Alicke & Sedikides 2011b, p. 3)

The types of psychological interests Alicke and Sedikides have in mind include 'security/ love, social status, and popularity, as well as the possession of high skill levels and capacities (e.g., intelligence, athleticism, musicality)' (2009, p. 6). In advancing these interests, these researchers take up the distinctions of 'primary control' and 'secondary control' where *primary* control refers to 'changing an objective state of affairs by taking effective or instrumental action, whereas secondary control substitutes psychological mechanisms that control events by altering how one perceives or interprets them' (ibid.). They consider that self-enhancement and self-protection are straightforward and non-controversial with respect to the primary control distinction. A person simply acts to achieve the desired changes. *Secondary* control, on the other hand, engages when primary control does not achieve the desired result. It is these secondary control processes that are relevant here.

> When people cannot promote themselves objectively, they have recourse to construal mechanisms such as reinterpreting the meaning of social or task feedback, misremembering or reconstructing events in a self-serving way, and making excuses for poor behaviour or performance.
>
> (ibid.)

To extend one of Alicke and Sedikides' examples; if someone has an interest in being popular they may work harder in their workplace in an effort to please and impress, or maybe throw more dinner parties to impress and please friends. However, if that does not work (or is perceived not to work), the suggestion is that they may misperceive or misremember events, or construe comparisons with other people in order to achieve the same goal (ibid.).

The interests may be: specific (for example, do well in a test) or general (for example, be a good employee); private (for example, meeting one's own standards) or public; involve close groups or extended groups; and, positive (things to attain) or negative (things to avoid). In a another distinction, Alicke and Sedikides point to a possible contrast in the same interest:

> The same interest can be cast in either a positive frame (e.g., a desire to be viewed as assertive) or a negative frame (e.g., a desire to avoid being seen as unassertive).
>
> (ibid., p. 7)

This final distinction is highly suggestive that the views a person has of their relations between themself and their social context(s) are intricately entwined with their specific beliefs related to those contexts (for example, *'It's wrong to be too assertive with your in-laws'*). Alicke and Sedikides hypothesise five levels for interests (see Figure 19). Of significance here is that they put 'global self-esteem'[85] at the top of the hierarchy and consider that all 'self-enhancement and self-protection processes are geared towards maintaining the most favourable conception of self and positive affect that reality constraints will comfortably allow' (2009, p. 7). Then self-enhancement, with respect to any specific interest, is aimed at preserving the current perceived level of functioning or achieving the level aspired to. By contrast, Alicke and Sedikides see self-protection processes as the analogue of the sympathetic nervous system (self-enhancement being the parasympathetic analogue). Self-protection is a form of 'damage control' (ibid., p. 14) engaged when the fulfilment of an interest is below a minimum acceptable level or is not being advanced sufficiently.

It is important to note that self-enhancement and self-protection 'never provide complete accounts or behaviour or judgment' (ibid., p. 17), and Alicke and Sedikides highlight that other effects such as expectancies and biased memories can lead to self-enhancement and self-protection. However, when self-enhancement and self-protection processes are invoked, and more particularly self-protection, there is a strong likelihood of self-deception.

> When extreme self-deceptions occur they are more likely to be enlisted for self-protection than self-enhancement. Whereas self-enhancement is a luxury, self-protection is a necessity. Except for psychotic delusions and extreme personality disorders, self-enhancement includes mundane tendencies such as thinking that one is slightly better than others, choosing to compare with worse-off others, and construing events in a way that frames one's actions and attributes in a positive light. These tendencies can sometimes be subverted simply by pointing out the facts to the self-enhancer.
>
> *The self-deception involved in self-protection, on the other hand, can be far more potent and hence much more difficult to overcome* [emphasis added].
>
> (ibid., pp. 21–22)

As these researchers say, 'examples abound': parents' belief that their children can do no wrong; denial of serious health conditions; excuses for improper behaviour that leave the listener incredulous; belief in multiple lives via reincarnation; resort to moral judgments when rationality is questioned.

85. Others seem to depict this as global 'self-integrity' (Sherman & Cohen 2006; Sherman & Hartson 2011).

While affect is considered by Alicke and Sedikides to play a 'relatively minor role' in self-enhancement, *negative affect necessarily accompanies self-protection*. In reporting work that is directly relevant to the question about maintaining a belief that lessens the self, Alicke and Sedikides (ibid.) report on studies examining the perseverance of negative self-belief following the discrediting of negative feedback subsequent to a performance task. They note that despite a significant reduction in perseverance, 'some degree of perseverance on the negative characteristic remains, showing that perseverance is indeed a viable phenomenon, and that self-protection interests are only partially able to surmount this tendency' (ibid., pp. 34–35).

The reduction in perseverance of negative affect could be partly explained by Sherman and Hartson's (Sherman & Hartson 2011) observation that people are willing to be self-critical if they can feel *globally* self-affirmed. They make a point that is also emphasised by Sherman and Cohen (2006) that the self-system, as they view it (see Figure 20), is flexible. As they put it:

> Because the goal of the self-system focuses on maintaining the *overall* [emphasis added] worth and integrity of the self, people can respond to threats in one domain by affirming the self in another domain.
>
> (ibid., p. 188)

Now, in this light, we can expand an example from Section 4.2.3. Take the person who took offence at the treatment received from a manager in the discussion about their fate in the forthcoming organisational restructure. They might find that the interest of furthering their work goals has been thwarted or at least has slowed down. However, they could affirm a different interest—for example, adhering to their personal standards—with a view such as *'Well, if I take longer to become a manager at least I now have more information on how to be a humane and ethical manager'*. Such a shift is commensurate with the mood management hypothesis aimed at homeostatic balance. Furthermore, when we recall that these emotions/moods[86] are accompanied by feelings, and more particularly $feelings_D$, it can be seen how these feelings can be a central driver in this shift. But it is not just the feelings at play: somatic markers ensure that memories and beliefs are playing an active role.

This type of example fits with much of what is proposed in self-affirmation theory (Aronson *et al.* 1999; Aronson 1999; Nail *et al.* 2004; Sherman & Cohen 2006; Sherman & Hartson 2011) and in the cognitive dissonance literature (Cooper 2007; Harmon-Jones & Mills 1999; Harmon-Jones 2000, 2001; Stone & Cooper 2001, 2003; Stone 1999;

86. See comments on distinguishing emotions from moods on page 114.

Thibodeau & Aronson 1992) with respect to the involvement of the concept of the self.[87] What is generally agreed in this literature is the importance of the concept of the self, and the connection with this thesis is succinctly put by Cooper:

> People will distort their cognitions about themselves in the service of protecting their self-system.
>
> (Cooper 2007, Chapter 5)

Cooper (ibid.) and Stone and Cooper (2001) propose a Self-Standards Model (SSM) that aims to account for some of the discrepancies found when contrasting self-consistency theory with self-affirmation theory.[88] While Nail et al. (2004) suggest further research is required to determine applicability of Stone and Cooper's work depending on the source of the cognitive dissonance (self or other), they do acknowledge that regardless of the eventual verdict on this work 'the present findings and much contemporary research point to the critical role of the self in dissonance processes' (ibid., p. 1903).

A particularly pertinent point from Cooper (2007) and Stone and Cooper (2001) which is captured in their SSM is the consideration that a person may act according to personal standards *or* cultural standards (external to themselves). This returns us to the importance of the awareness of the self in relation to the environment (society, culture), and it points to the intricacies of the interrelationships of the feelings evoked by a situation (via a somatic marker) and the associated beliefs and memories.

4.2.4.2 False beliefs in belief systems – for the sake of the self

Regardless of whatever neurological and physiological mechanisms combine to adjust beliefs, the suggestion from Clore and Gasper (see page 175) that we adjust beliefs to be compatible with both internal evidence (in the form of beliefs) and external evidence is compatible with the views above that we act and adjust beliefs to help protect our concept of self. From earlier discussion we can emphasise that it is the self in relation to the context that is being protected.

Threat can arise externally. But it is not always the obvious and mundane that need attention. Attacking animals, people attacking one physically, heights, dangerous driving practices, for example, are all threats that we would expect to initiate at least a fear response and, hopefully, a subsequent set of decisions and actions to protect the body

87. Extensive discussion of this literature is not warranted for this thesis as there is considerable debate relating to self-protection vs self-affirmation; self-consistency vs self-affirmation; experimental methodology; and more. While interesting, that debate would be a diversion here. The elements of the literature that are relevant to this thesis appear to have largely subsumed by the self-enhancement and self-protection literature.

88. See Figure 21 and Figure 22.

physically. When considering the effect on beliefs, the real concerns are those instances of perceived threat to the individual's self-concept, or self-identity.

Clearly, there is broad agreement about the propensity to act in a way that protects the self. An 'act' does not need to be a physical act: it can be a change of attitude or the creation of a belief. If, for example, I can create a belief that the person who criticises my golf technique is a total idiot whose opinion is worthless, then I can maintain a value on all the years I have been practising (no matter how badly) and can prevent or minimise feelings of damage to my self-esteem or, indeed, my self. Such a type of belief is commonplace. There is no enquiry into the criteria for categorising an 'idiot'; nor is there consideration from a detached perspective in order to assess my actual skill level; nor is there any overt effort to take into account factors that may affect communication (for example, fatigue, alcohol). So at the very least, I have created a belief that has no justification in explicit evidence. What was the evidence? Predominantly, the feelings.

This type of belief is relatively uncomplicated and other than emphasising the role of feelings could be considered unremarkable. Consideration then needs to be given to the system in which any belief may reside. Frijda and Mesquita (2000) highlight one of the principal factors contributing to the resistance to changing a belief within a belief system even if it is clearly contradicted by evidence.

> It probably is the anticipation of emotional situations and corresponding emotions that can give beliefs their enormous power when challenged … **Changing one belief would require changing a whole system of beliefs** [emphasis added]… Major beliefs are linked to action. They guide everyday judgments and decisions and modes of interpersonal conduct … Generally speaking, challenging major beliefs challenges an individual's world view, which challenges sense of security.
>
> (ibid., p. 63)

Quine and Ullian (1978), too, observe the challenge when finding an inconsistent set of beliefs. They suggest that if one is decisive then the whole set must be either kept or switched to disbelief. In the event of being indecisive, judgment on some beliefs may be suspended, leaving them as non-belief.[89]

A striking example of the challenge to a belief system that can occur is the case of the a person who believes they have been sexually abused by a parent. Following a court case it becomes evident that the person had a false memory and the evidence exonerating the

89. There is a temptation here to indulge in considerable discussion of belief systems. However, an analytical philosophical approach is beyond the requirements needed for this thesis; and an extensive discussion from a more general approach risks being an indulgence. The reviewed literature clearly shows acceptance of the view that any one belief is not held in isolation; that point is sufficient for the purposes here.

parent is unequivocal. Despite this, the person is unable to admit that they were in error nor are they able to forgive the alleged offending parent. To change the belief that their memory was accurate would most likely have extensive implications: if they were wrong about something like that, how can they trust themself in the future; what sort of person are they that they could treat an innocent parent in that way; how can they explain this to their therapist and support group(s); what does this say about their therapist who encouraged them to take legal action?

Frijda and Mesquita (Frijda & Mesquita 2000) place concern relevance at the centre of emotional influences on beliefs. At this stage, it can be hypothesised that *the* central concern in a self-system is, naturally enough, the self. And it may be that even when the self is not directly in focus, it cannot not be peripheral. On a relatively mundane level, if I am running late to a meeting, I might well be concerned about the outcome of the meeting. However, it is also very likely that I will be concerned about how *I* will be judged if late. The somatic marker associated with the anxiety about being late could well entrain memories of judgments made by others in relation to latecomers (for example, *'How inconsiderate of them', 'If they're late for an important meeting like that, they'll be too unreliable for this job'*). The anticipation of emotional consequences that Frijda and Mesquita mention (op. cit.) is then a natural part of the sequence of the 'as if loop' simulations instigated by the anxiety somatic marker and perpetuated by the feedback loop between feelings and imagined unwanted consequences.

Perhaps a simpler way of expressing the consequence of Damasio's postulated 'as if loop' is the plain speaking from Prinz: 'Positive thoughts lead to positive affect, and negative thoughts lead to negative affect' (Prinz 2004a, p. 30), a conclusion that Prinz suggests hardly needs an experiment to show. Nevertheless, it is supportive of the concept of an 'as if loop' which allows a person to simulate how they might feel in a given imagined situation. Positive or negative, the ensuing appraisal informs how the thought may bear on their concerns.

Concerns go beyond the demands of traffic and meetings: for example, political, personal relationships, religious convictions (Frijda & Mesquita 2000). On the subject of religious conversion Frijda and Mesquita suggest that a new belief can be:

> … rendered indubitable by the network of detailed beliefs, social habits and modes of interpersonal interaction in which it is embedded. Acceptance of such a belief is supported by living according to the entire network.
>
> (ibid., p. 51)

Metzinger adds to the impact of considering the system of beliefs and influences that interrelate with it when he says this of religion:

> Religious belief is an attempt to endow your life with deeper meaning and embed it in a positive metacontext—it is the deeply human attempt to finally feel *at home*. It is a strategy to outsmart the hedonic treadmill. On an individual level, it seems to be one of the most successful ways to achieve a stable state—as good as or better than any drug so far discovered.
>
> (Metzinger 2009, p. 211)

Similarly, for the areligious or anti-religious person, it can be argued that intense political affiliation or even strong adherence to environmental ethics perform a similar role. Such commitment to an accepted context, whether real or perceived, gives a place for the self. There is a context in which the person can find themselves, their *I*, in relations with the elements (people, events) of that context that stabilise the self, that give it security.

As much as we might aspire to a world in which beliefs were based on logic, it is clear from the literature that logic has a mighty challenge to overcome the survival of homeostatic processes that pervade the whole self-system, from the physical body to the perceived entity known as the self. In looking at this system in its context, we can postulate that the formation and maintenance of beliefs ultimately aim at securing and stabilising this entity. Overt autobiographical beliefs (for example, *'I'm an excellent manager'*) have a prominent and obvious role in this system. Even beliefs about mundane elements such as the traffic flow can be traced back to concerns about the self (for example, *'If the traffic were better, I'd spend less time commuting, get more sleep, be more alert, and therefore better at my job.'*)

Since emotions and feelings are such strong drivers in (1) engaging existing belief systems, (2) creating new beliefs, (3) maintaining or strengthening existing beliefs, and (4) changing beliefs, the 'truth' of the belief is of little consequence if it is coherent with the existing and immediately relevant belief systems and helps protect and/or enhance the self. What is the principal signal that a belief satisfies these criteria? From the research reviewed, the conclusion is that *it feels 'right'*. No wonder then that it is easy and common for people to create false beliefs at every level.

Unless a person has a core concern to live in a context that is governed by a commitment to logic and evidence and has a personal interest in living up to the attendant personal and contextual standards, it seems that truth and reality are subject to physical and self-conceptual homeostasis processes. From this perspective, further pursuit of the mechanisms for creating false beliefs versus accurate beliefs is arguably irrelevant. If the aim is to understand how false beliefs are formed and maintained, it is almost as if any belief is good enough as long as feels as though it fits.

4.3 Summary

This chapter began by focusing on the feelings-as-information hypothesis. It was observed that as well as widespread lay use of the words 'feeling' and 'feelings' with respect to making judgments and confirming beliefs, there are research results showing that feelings are indeed a principal criterion in confirming judgments and beliefs.

The proposition was put forward that Damasio's view of emotions and feelings gives weight to this argument as feelings are part of a somatic marker that entails context-relevant memories and beliefs. Models that were introduced in Chapter 3 were re-examined in the light of the suggestion that the information value of feelings means that feelings often take priority over propositions and that they draw attention to relevant goals and concerns. The importance of self-representation emerged, and this pointed to investigating how Damasio's view of feelings may further the understanding of how they act as evidence for autobiographical beliefs that contribute to the construct that we call the 'self'. In doing this, the suggestion that Damasio's definitions for emotion and feeling should be considered as being from 'first principles'. This was emphasised because some commentators appear to attempt to understand Damasio through the filters of existing definitions. The stance taken here is that Damasio's definitions offer a 'clean slate' and because of their explanatory power they were taken as the operating definitions of *emotion* and *feelings* for the remainder of the thesis. An important point from Damasio's hypotheses is that feelings$_D$ privilege interoception over exteroception—a point which has not been fully appreciated (or perhaps understood) by some commentators but which is pivotal to understanding the degree to which feelings can dominate appraisals and confirm beliefs.

LeDoux's work on fear was briefly reviewed to further indicate how there can be a physical reaction before there is conscious awareness of feelings of fear (or some other emotion). This provides at least one mechanism that can explain how a somatic marker can give rise to the feelings, the awareness of the feelings, and post-perceptual cognitive processes. While a relatively straightforward emotion such as fear is easily understood in these terms, the recognition that cognitive processes can follow bodily response opened the way to examine the more abstract social emotions such as embarrassment, shame, and satisfaction. This pointed to the value of Damasio's hypothesis for the development of consciousness in explaining these more abstract emotions.

The hypothesis of a protoself, core self, and autobiographical self provides a structure with a reference point for action and mechanisms for being able represent changes occurring in the body–mind system. Not only does this hypothesised system allow registration (to use Prinz's (2004a) term) of changes occurring in the body but it also enables the awareness of

the associated feelings, the bringing of associated memories and beliefs into working memory, and the ability to 'observe' the changes occurring not only in the physical system but also the cognitions as an emotional state progresses. Damasio's hypothesis does far more than provide a model for thinking about hierarchical levels of consciousness; it is at the forefront of offering a neuroanatomical explanation of how levels of consciousness arise. It also finds correlation in other hierarchical models of consciousness or representational processes. While a full explication of the details of the neurophysiological mechanisms that create second- and third-order representational processes is still in the future, the current models support a view of an autobiographical, or narrative, self.

Concerns related to the self were then examined. An example used was that of a person burning their hand and progressing from the reflex action of withdrawing their hand from the source of the heat to emotional experiences that lead to an autobiographical belief related to their self-efficacy in the kitchen. It was then shown that Damasio's Somatic Marker Hypothesis (SMH) offers a plausible account of how the feelings initially associated with a burn can entrain memories of similar past incidences and beliefs about themself in that context.

A significant aspect of Damasio's SMH is the concept of an 'as if loop' which ties the experience of bodily feelings, and associated emotions, to the simulation of experience. Thus a person (a self) has a mechanism for anticipating how they might feel and react to an imagined situation and make judgments about the situation, themself, or others. It was noted that the 'self', as each of us perceives it, is a process; but in perception it does live as an entity. With respect to preserving and securing that entity, it was shown how a repeated instance of a simple homeostatic need (for example, thirst) could lead to an autobiographical belief (for example, 'I don't take enough care of myself'). A strong view in the literature is that these types of beliefs result from the emotions and feelings because these guide attention to concerns.

It was shown that the concerns of the 'self' go well beyond physical homeostatic balance to concerns about one's place in the world, the interactions with others in various contexts, social concerns, and even those concerns that may be labelled existential. Therefore, the autobiographical self becomes the focus of attention in many instances, and the signals from emotions and feelings draw attention in a way that is aimed at protecting that same autobiographical self. It was then shown how the elements of a somatic marker could be instrumental in this process with the objects of an emotion being such things as the imagined thoughts of others about oneself or even the self itself (for example, 'I look so incompetent in meetings'). This was followed by extensive examination of the influence of emotions on the autobiographical self in context.

The primacy of feelings in guiding attention and the style of cognitive processing was established; but there remained the question of how it is that a person might form and maintain a belief that appears to be to their detriment. Homeostatic mood management is a concept presented by Forgas (op. cit.) and is supportive of the view expressed here that emotions and feelings are involved in higher order processes aimed at stabilising the self. Valuable insights were gained from the literature on self-enhancement and self-protection arguing that when someone cannot protect or promote themselves objectively through action, they resort to reinterpretations, excuses and reconstructions of events. The principal finding of relevance here was that in the interests of promoting global self-esteem, a person could form a belief in one domain that appears to work against their interests (for example, 'I'd be a hopeless manager') but which allows another domain of interest/concern to be promoted ('… but at least I don't have to become obsessed with budgets'). This provided a plausible explanation for the question about the rationale for the forming of apparently debilitating beliefs.

Finally, it was observed that any belief, especially about the self, resides in an interlocked system of beliefs. Changing one belief in a system necessarily involves changes in other beliefs in that system. This supports the proposition from the self-enhancement and self-protection literature that belief adjustments are made to achieve an overall feeling of good about oneself. Moreover, it has long been acknowledged as one of the principal sources of resistance to changing beliefs.

The literature showed that the influence of feelings on judgments and belief is not only powerful but unavoidable. Although a major focus of the thesis is false or implausible beliefs, there would appear to be little distinction made by feelings per se regarding 'truth' or 'reality'. From the self-enhancement and self-protection literature it would seem that for a person to have feelings that were more disposed to direct attention in a way that formed beliefs supported by evidence, they would have to have something like 'logic' or 'truth' or 'evidence' as a highly valued domain of interest. Without that, it is reasonable to conclude that when emotions take over, any belief is right if it feels right.

Chapter 5: Conclusion

> The falsehood has to start somewhere; the seed may be sown in hallucination, illusion, a normal variety of simple misperception, memory deterioration, or deliberate fraud, for instance, but the false beliefs that are reaped grow in a culture medium of true beliefs.
>
> (Dennett 1987, p. 18)

5.1 Summary

The thesis was introduced with reference to a lament in popular science literature (Mooney 2011) about the perceived growth of 'cognitive relativism' and 'truthiness'. This opened the discussion about the potential damage that can be consequent from false and irrational beliefs. The tension between views such as those of Clifford (1999) and James (1956) was noted and the significant observation was made that James draws attention to the fact that beliefs are influenced by processes other than the purely logical:

> … our non-intellectual nature does influence our convictions.
>
> (ibid., p. 11)

The profound role of beliefs in the lives of individuals, societies and countries was noted by Damasio (2000b) who was introduced because of his interest in studying belief from a neurological point of view. As noted in the Introduction, Damasio's suggestion is that what we recognise as beliefs are those things which are particularly important to us. This, along with the proposed role of emotions, indicated a path of inquiry that may help explain how it is that normal people can cling to beliefs that are either demonstrably false, or for which there is no available evidence and which are falsifiable through logically sound argument.

While ever we might might agree with Clifford's (op. cit.) view of rigour in basing beliefs on evidence, James' view that emotions and feelings come to bear on belief formation and maintenance accords with that of contemporary researchers and with our everyday observations. Of particular interest are those beliefs that cause concern, difficulty and disappointment for a person and which have the characteristic of being evidentially false or highly implausible. More broadly, however, the factors that contribute to forming and maintaining any belief were necessarily drawn into focus.

Chapter 1 began by acknowledging the prominence of beliefs in our thought processes (James 1890/1950a; Quine & Ullian 1978; Russell 1921/2005; Schwitzgebel 2010). Given the central role of beliefs in providing frameworks for decision making and for

action, review of some of the debate relating to defining beliefs was undertaken. The philosophical literature reviewed focused on the degree to which the truth of a proposition needed to be established in order to support a given belief. During this investigation, the value of incorporating the findings from psychology emerged (Damasio 2000b; Garfield 1988; Goldman 1986; Hirstein 2005; Russell 1921/2005; Westbury & Dennett 2000). Consequently, for the interest stated in the Introduction, a belief was taken to:

- involve a view that a proposition is true regardless of factual evidence
- often be related to desires and wants
- be a preparation for action.

There still remained the matter of distinguishing *belief* from *knowledge*, and it was observed that people often use 'I know' to express that they hold a belief. The complexity of the philosophical pursuit was noted and for the context relevant to this thesis *knowledge* was taken to satisfy two principal criteria:

1 The belief or the proposition is held to be true.

2 Available data are examined or interpreted to justify the claim as knowledge rather than belief. The data may be perceived external events or may come from internal processes such as emotional responses.

What was acknowledged is that there is not necessarily an inherently rational process used to justify the claim of knowledge. Importantly, a strong argument emerged for taking the following point of view into consideration: what is deemed to be evidence for the proposition is what the *speaker* takes as evidence.

The remaining issue of distinguishing memory from belief and knowledge gave more weight to the argument for enquiring into both the psychological and neurological mechanisms that give rise to such distinctions. From the reviewed literature, it was clear that beliefs are formed in a complex system involving emotions, perceptual processes, interaction with other beliefs and propositional attitudes, and metacognitive beliefs. Given that complexity, Westbury and Dennett's (2000) exhortation to account for context in defining and understanding belief was considered to be particularly apt. Aligned with this was the determination that investigating beliefs, using the chosen definition, from a predominantly *psychological* perspective would best serve the focus of this enquiry. From the psychological perspective, a natural starting point for the investigation of influences on belief formation was determined to be the role of perception. However, it is important to note that the philosophical literature reviewed also argued strongly for examining the role of perception and sense-data in forming beliefs (for example, Audi 2011; Berkeley 2009; Goldman 1986; Locke 1690; Morton 1997, 2003).

The simplest case of perception affecting beliefs is arguably that of the sensory dysfunction. Chapter 2 provided examples of how impaired sensory apparatus (for example, auditory or visual impairment) could give rise to false beliefs . It was also clear in the research that individual differences in sense perception—in the absence of any damage or pathology to the sense organs—can lead different individuals to markedly different experiences of and, consequently, beliefs about events. The types of differences that are common include, for example, taste perception and colour perception that arise from genetic differences.

Chapter 2 also reviewed the roles of sensory illusion and sensory limitations such as 'change blindness blindness'. It was established that such perceptual processes lead to inferences which can be the basis for false beliefs and erroneous claims of knowledge. Related to these perceptual processes were the heuristics that affect judgments as explicated by researchers such as Tversky and Kahneman. An extensive review their work, and related work, allowed the conclusion that the heuristics and biases to which humans are subject can easily lead to false beliefs and false claims to knowledge. The judgments that people make in many situations were shown to fit the selected criteria for a belief, thereby giving cause to suggest that heuristics and biases directly *cause* many beliefs.

Following that, a brief investigation of the other factors related to perception that could contribute to belief formation was undertaken. These included emotions, pre-existing beliefs, and language via its influence on perception. As the role of emotions and feelings emerged as being particularly significant, that became the key focus for Chapters 3 and 4.

Chapter 3 opened by establishing that there is general agreement that emotions affect beliefs (and vice versa), and emotions also affect memories (and vice versa). Given the significant role of emotions and feelings in affecting beliefs, definitions of these terms came into focus. Again, agreement was found in the literature; but this time it was that there is extensive *lack of agreement* with respect to definitions of the terms 'emotions' and 'feelings'. As well, the conflation of the two terms was observed.

There was broad agreement regarding the general features of emotion and also that emotions are separate from feelings. Examination of the hypothesis from Antonio Damasio suggested that Damasio's somewhat counter-intuitive definitions, based on neurophysiological findings, offered the most useful definitions for the purposes of examining the interaction between beliefs, emotions and feelings.

In a development of the observations made in Chapter 2, it was shown that emotions affect attention and perception. Given that it had already been established that perception can influence beliefs or directly lead to the creation of beliefs, there is a clear link between emotions and beliefs via emotions' effects on perception. Chapter 3 then continued to

establish that emotions affect memory and in doing so have a direct affect on at least some beliefs, especially where these beliefs are autobiographical. A key suggestion in the literature was that positive and negative emotions engage different forms of information processing and thereby can produce differences in memories.

This was followed with an examination of key literature relating to the interaction of emotion with beliefs specifically. What was established is that there is agreement that beliefs are involved in the processes occurring when a person experiences an emotion. The literature also provided a strong case for including beliefs as an integral part of experiencing and appraising emotion. A significant observation made is that the term 'appraisal' refers to both the content of the emotional experience and the emotion that is the *result* of an appraisal process. This led to noting the significant role of beliefs in affecting appraisal processes.

The final part of Chapter 3 focused on the role of feelings as feedback about the significance of events and related emotions. This returned the enquiry to the Damasian view of feelings and opened the way for investigating the effect of feelings on autobiographical memories and beliefs, and on a person's concept of self.

Chapters 4 opened with evidence from research that suggests that feelings are a principal criterion in confirming judgments and beliefs. This allowed for further enquiry into the feelings-as-information hypothesis through the lens provided by Damasio's work. Re-examining models that were introduced in Chapter 3 suggested that the information value of feelings means that feelings often take priority over propositions and that they draw attention to relevant goals and concerns. The importance of self-representation emerged and this increased the relevance of Damasio's work given his hypotheses around the development of the autobiographical self.

Having taken Damasio's definitions of *emotion* and *feeling*, his hypothesis for the role of feelings opened the enquiry into the development of consciousness and the autobiographical self. Damasio's view of a protoself, a core self, and an autobiographical (or narrative) self found parallels in existing models suggesting 'levels' of consciousness. In the relevant literature there was general agreement that there had to be neural mechanisms that allow awareness of body states and ultimately awareness of a 'self'. Feelings were shown to signal not only straightforward physical concerns but also matters that are of concern to the autobiographical self. This last view was supported through Damasio's Somatic Marker Hypothesis in which the feelings associated with emotion not only draw attention to matters of physical concern (for example, homeostasis needs) but also entail memories and beliefs associated with that body–mind state. The entrainment of memories and beliefs about a particular type of event in a given context—one example used was a person burning their hand while cooking—was shown to establish ways in which a person

could form beliefs about their capabilities in that context or, indeed, about their own worthiness, for example.

It was also shown how Damasio's hypothesised 'as if loop' allows a person to imagine how they *would* feel in a given situation. In relation to the 'self' this had particular relevance to the social emotions and the stability and security of one's self in society. That these concerns about interactions with others in social contexts goes well beyond physical homeostatic balance pointed to the role of signals from emotions and feelings drawing attention in a way that is aimed at protecting that same autobiographical self. Throughout this part of the investigation the primacy of feelings in guiding attention and the style of cognitive processing was established.

The chapter then turned to the literature on self-enhancement and self-protection to gain insight into how it can be that a person will form an autobiographical belief that appears to be against their best interests. The principal finding was that in the interests of promoting *global* self-esteem, a person could form a belief in one domain that appears to work against their interests but which allows another domain of interest or concern to be promoted. This provided a plausible explanation for the question about the rationale for the forming of apparently debilitating beliefs.

Finally, it was observed that changing one belief in a belief *system* necessarily involves changes in other beliefs in that system. This supported the proposition from the self-enhancement and self-protection literature that belief adjustments are made to achieve an *overall* feeling of good about oneself. It also highlighted one of the principal sources of resistance to changing beliefs.

Overall, the literature showed that the influence of feelings on judgments and belief is not only powerful but unavoidable. Feelings per se were shown to have no necessary reference back to 'truth' or 'reality'. On the contrary, for the most part, it was shown that feelings themselves are the indicators of 'truth'. So, it was concluded that unless using logic or seeking verifiable evidence is a highly valued domain of interest for a person, the feelings associated with emotions signal what the human system perceives to most serve the self and therefore they are a principal agent in the formation of beliefs and any effect on beliefs. It follows that they are also a principal agent in resistance to changing beliefs.

5.2 Conclusions

Having taken a predominantly psychological approach, this work has shown that there are numerous agents that contribute to forming, maintaining, strengthening, or changing beliefs. In a person who has no neurophysiological pathology, these agents include:

- perceptual phenomena including:
    - normal limitations to perceptual processes (for example, inattentional blindness)
    - genetic differences (for example, colour blindness, differences in taste perception)
- heuristics and biases as outlined in Chapter 2
- emotions and feelings through effects such as focusing attention
- pre-existing beliefs and memories.

With so many functional aspects contributing to belief formation, it is abundantly clear why arguing against a false belief through logic can be such a futile exercise.

In the pursuit of understanding how people create and maintain false beliefs, the advice of those recorded in Chapter 1 who encouraged the enquiry from a predominantly psychological perspective—arguing that philosophy ought to be informed by current scientific knowledge—has proved fruitful in a way that is difficult to imagine from a purely philosophical approach. Of course, the point has to be accepted that what many people call a 'belief' might not fit a definition supplied by a strict philosophical approach. But that does not stop people using the term and acting on beliefs that fit the definition chosen for this enquiry.

Tracing belief formation in the case of relying on perception is relatively straightforward. If, for example, a person recalls a scene and they are confident that they saw three people in a given situation, their conviction of that can be considered to be directly related to their perception. Of course, one key problem arises: if the person has no way of recognising that they have misperceived (for example, there were only two people), then from the objective observer's perspective they have a false belief, but from their own perspective they hold the proposition (that there were three people) to be true because they saw them. Here, there has to be a commitment to relying on their own perceptual processes. Whether they overtly express the belief that they trust their sense-data or whether they do so unconsciously, we can propose the existence of such a metabelief. A similar situation applies when acting on judgments that are subject to the heuristics and biases that are consequent on evolutionarily derived neural mechanisms. For the most part, as some researchers noted, this is not only reasonable but it is necessary for our general functioning in our world. What is important though is perhaps the *lack* of

willingness in many people to mistrust their senses, or to at least question them, when what they believe they experienced is false and could have seriously unfortunate consequences (for example, in a murder trial). The emotional reaction that many people exhibit when the accuracy of a memory, for example, is tested points to the major findings in Chapters 3 and 4: the powerful significance of emotions via feelings. If you ask someone who is adamant that their account of an event is accurate (that is, 'true') to question that account, it is suggested that they will frequently resort to the type of phrase discussed in Chapter 1 (for example, *'I just know'*), thus reinforcing our inference that there is at least an unconscious commitment to their perceptual record—that is, that they believe they can trust their senses.

If we persist with that line of questioning and the person remains adamant that they are right, there is ample evidence that ultimately they are relying on feelings as evidence for their belief. Interestingly, if someone becomes willing to reconsider a belief that is being challenged, general observation suggests that this is usually accompanied with noticeable changes in bodily response (posture, breathing, gestures, physiognomy), suggesting significant changes in experienced feelings.

The combination of Damasio's Somatic Marker Hypothesis and the literature on self-enhancement and self-protection give a credible account of how a belief, or indeed a belief system, will be maintained in order to protect a person's neural construct of their 'self'. And, as was suggested earlier in this thesis: as much as we might aspire to a world in which beliefs were based on logic, it is clear from the literature that logic has a mighty challenge to overcome the survival of homeostatic processes that pervade the whole self-system, from the physical body to the perceived entity known as the self.

We could conceive that in a folk psychological context this finding could be received with a response such as *'Well, of course, isn't that obvious?'*. After the fact, it may seem obvious; but it is not so if the enquirer steps aside from their own feelings of certainty and seeks evidence and considered philosophical and psychological enquiry. This is especially the case since there is no such *thing* as a self.

It is offered here that the proposition that the role of many beliefs—supported by emotions and feelings—in protecting the self is particularly significant. It accounts for so much relating to the seemingly undue influence of emotions and feelings, especially when emotions and feelings are defined as they are by Damasio.

We could say that, in a sense, it begins and ends with homeostasis. Once we move beyond simple reflex actions, our systems are designed to establish homeostasis in the body: that much is uncontroversial. However, it is noted that others (for example, Forgas 2000c; Forgas 2002; Forgas & Ciarrochi 2002) have also made the suggestion that homeostasis is

also sought in social emotions and the interaction of one's self in one's world. It is concluded that the ultimate purpose of many beliefs—particularly autobiographical beliefs—is arguably the homeostasis of the self construct. Feelings appear to be the principal signal as to whether a belief is fulfilling that purpose.

It must be noted here that although this thesis opened with an express interest in false or implausible beliefs, much of the discussion related to beliefs in general. In examining the influences that form and affect beliefs, it emerged that the influences (perceptions, heuristics, biases, emotions, feelings, memories, pre-existing beliefs) need not be especially different in their action in forming a true *or* a false belief. There are some observed differences relating to positive and negative affect; but, in general, we have observed that true or false beliefs can be formed or strengthened from either. The principal hypothesis is that beliefs are serving the (perceived) security and stability of the self: if this stands, then it seems not to matter whether the belief is true or false as long as it protects the self construct.

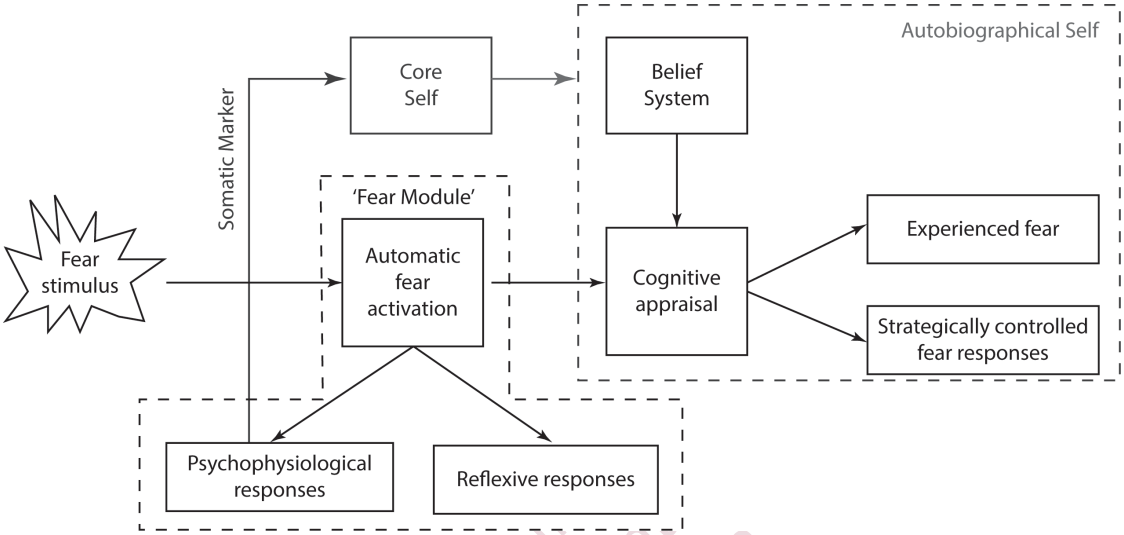
The original instigation for this interest was the potential application of the findings in the areas of coaching and counselling predominantly. Having a greater understanding of the processes by which false or implausible, debilitating beliefs are formed opens the way for additional approaches and, hopefully, for more effective approaches.

Clearly there is much more research to be done to understand the neural pathways and biochemical changes involved in creating, maintaining and changing beliefs. That notwithstanding, there is already early discussion of being able to change memories with drugs; but that is early discussion and inevitably the subject of much debate on the ethics of doing so.

While the fine structural and biochemical details continue to be discovered, the psychological approach benefits greatly from the continuing refinement of the models that guide psychotherapeutic approaches, ongoing research, and even general applications such as approaching political debate. Great orators and rhetoricians have known for centuries that for most people an appeal to logic alone is largely futile; the appeal must be made to the emotions. Acknowledging the role of feelings adds new a new depth of insight into many of the applications of the findings herein. If the aim is to help a person review and reevaluate a belief or belief system, knowing that it requires asking them to question their reliance on *felt feelings*$_{D}$ invites different approaches that more readily address the true state of neurophysiological affairs.

5.3 Further Research

An area for continuing research that stands out from this work is the continuing refinement of models to guide psychological applications of the findings, research pathways, and general debate. A starting point could be the adjusted model shown in Figure 16 (and repeated below).



As more is discovered about the neural mechanisms supporting proposed second- and third-order forms of consciousness, we may expect to test the validity of the overall structure of this model and, in particular, the additions made. Allied with this is continued investigation of Damasio's Somatic Marker Hypothesis with particular focus on the feedback loops between memories, beliefs and feelings. It is relatively straightforward to conceive that once a feeling is registered the neural circuitry associated with second- and third-order reflexive processes introduce related memories and beliefs (sometimes these being the same) into working memory. But, there is still much that could be done to devise suitable models to explain how those same autobiographical phenomena feed back to the proposed protoself in a way that, firstly, changes the state of the protoself and, secondly, results in further changes in consciousness (Damasio's core self, or second-order processes) which continue to affect the autobiographical phenomena. It is conceivable, for example, that such understanding could help interrupt cognitive processes in which belief systems feed back information that strengthens feelings that sustain irrational or false beliefs. The ideal outcome from such understanding would be to help people reevaluate strongly held beliefs that are not serving their best interests without being subject to the feelings that sustain resistance to changing those same beliefs.

That ideal outcome might be considered a philosopher's dream. Why that dream is so difficult for many to realise is much easier to understand given the findings of the

reviewed literature. What has not been mentioned so far, though, is that some people clearly do have the capacity to reevaluate beliefs despite initial feelings of resistance. An intriguing research project could be with subjects with contrasting responses to requests to reevaluate beliefs—those who can readily do so and those who strongly resist. Such a study would be ripe for both neuroscientific studies and as well as cognitive and social psychological studies.

In areas such as neuroscience it is expected that findings will contribute to our understanding of belief formation, maintenance and change. Damasio's hypotheses have evolved out of studies of structural damage to the brain. We can anticipate that he and others (e.g. Hirstein, Bermúdez, Bortolotti) will continue to develop this field. Allied with these areas are expected developments in areas such as those pursued by Eichenbaum into the brain circuits that may be responsible for distinguishing belief, memory and knowledge. As well, work such as Gazzaniga's that considers the left hemisphere as an interpreter can be expected to enrich our understanding, particularly if taken in conjunction with work on the role of language. The role of positive and negative affect, taking into account the self-enhancement and self-protection literature with possible extension into related neural circuits, is viewed as being another area that can contribute significantly to the understanding of the formation and role of autobiographical beliefs.

A complete marriage between the neuroscience findings and the models from cognitive and social psychology is still some distance away. However, it is offered that the literature reviewed shows that the engagement (if the metaphor may be continued) is confirmed and well-established. It is suggested here that one of the great contributions of Cognitive Science is its capacity to take a pan-disciplinary approach[90] that enables discovery of connections that might otherwise remain unobserved. Much of the literature reviewed already takes an interdisciplinary approach and it is hoped that such approaches will continue to develop the models in focus while the individual disciplines continue to provide the fine detail to enrich those models.

In researching this thesis a number of areas of literature have received some attention and some considerable restraint was exercised in not being seduced into pursuing those avenues too deeply. One area that is considered ripe for development is the role of language in affecting—or even directing—perception and in shaping beliefs. This is especially the case in light of James' view of felt understanding and the case for feelings proposed in Chapter 4, along with the notion that language serves to trigger somatic markers. Given the influence of language observed by researchers (for example, Barrett *et al.* 2007; Boroditsky 2001; Drivonikou *et al.* 2007; Gilbert *et al.* 2006; Levinson *et al.*

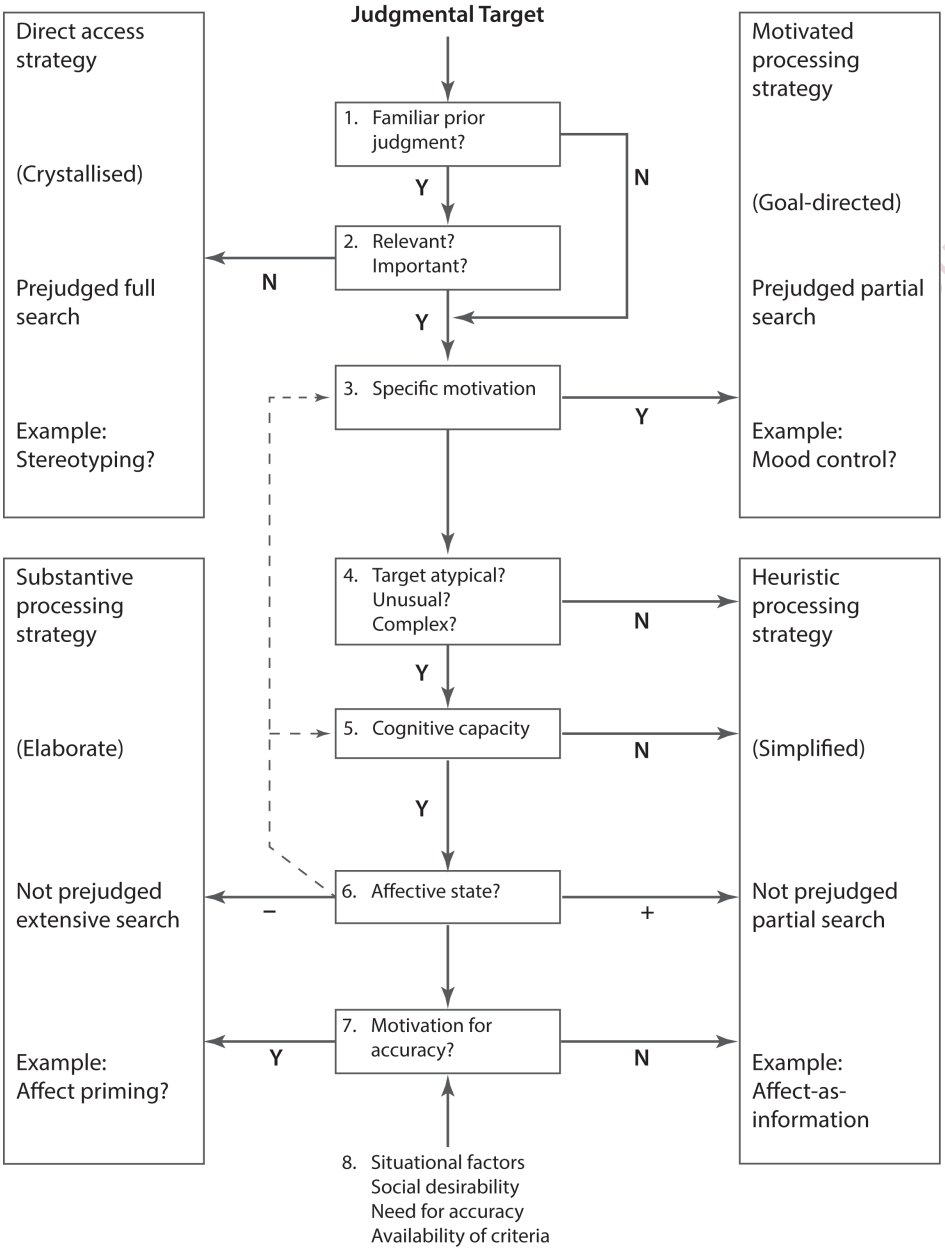
90. In the case of this thesis: philosophy, neuroscience, cognitive psychology, social psychology

2002; Mo *et al.* 2011; Tan *et al.* 2008), it is conceivably a key part of the mechanism of creating and, especially, maintaining beliefs because of the embodied representations linked to the language as well as the somatic markers that may be evoked. This is considered to be a strong candidate for research following on from this thesis.

Appendices

Appendix 1

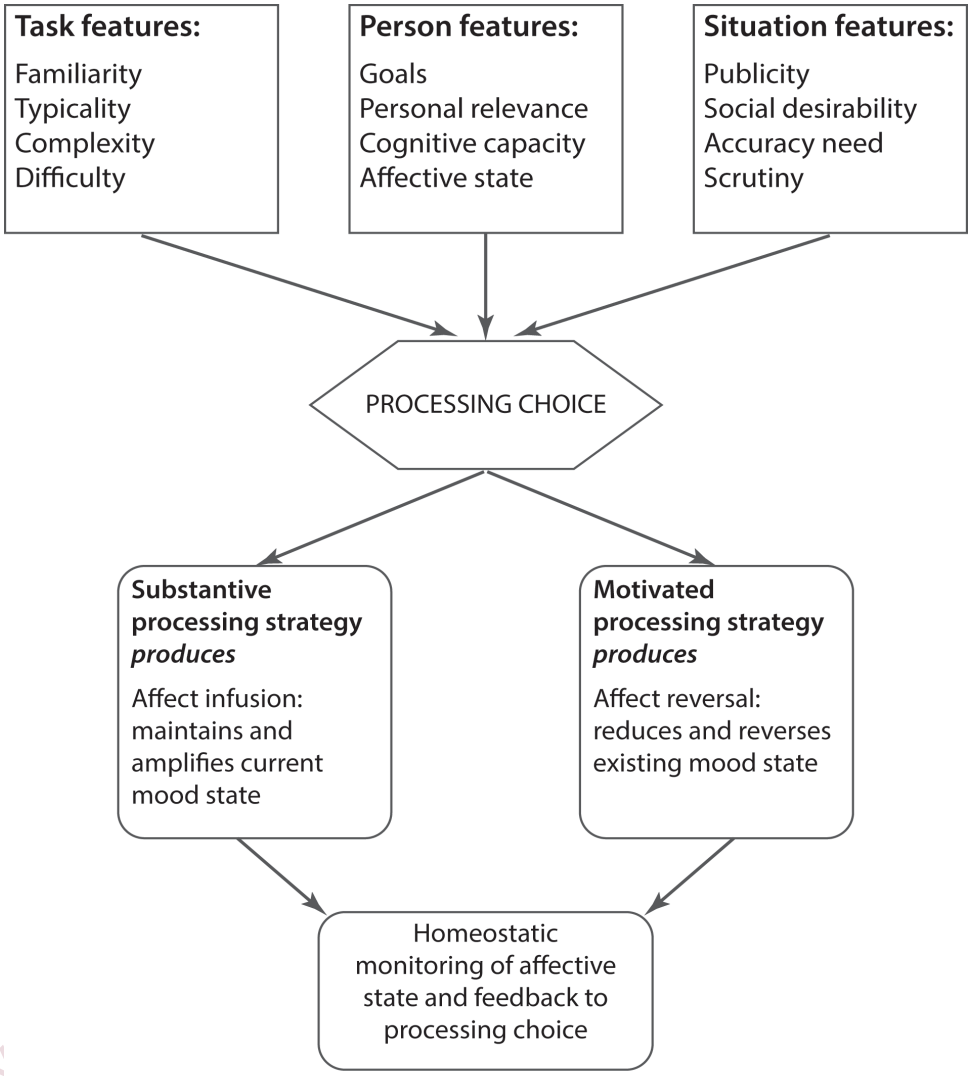
Figure 17: Hierarchical relations of antecedents determining processing choices



Flowchart illustrating the hierarchical relationship among various antecedent factors determining processing choices and the multiple informational and processing consequences of affect on judgments. Y = yes; N = no

(After Forgas 1995, p. 49)

Figure 18: Mood management hypothesis – outline

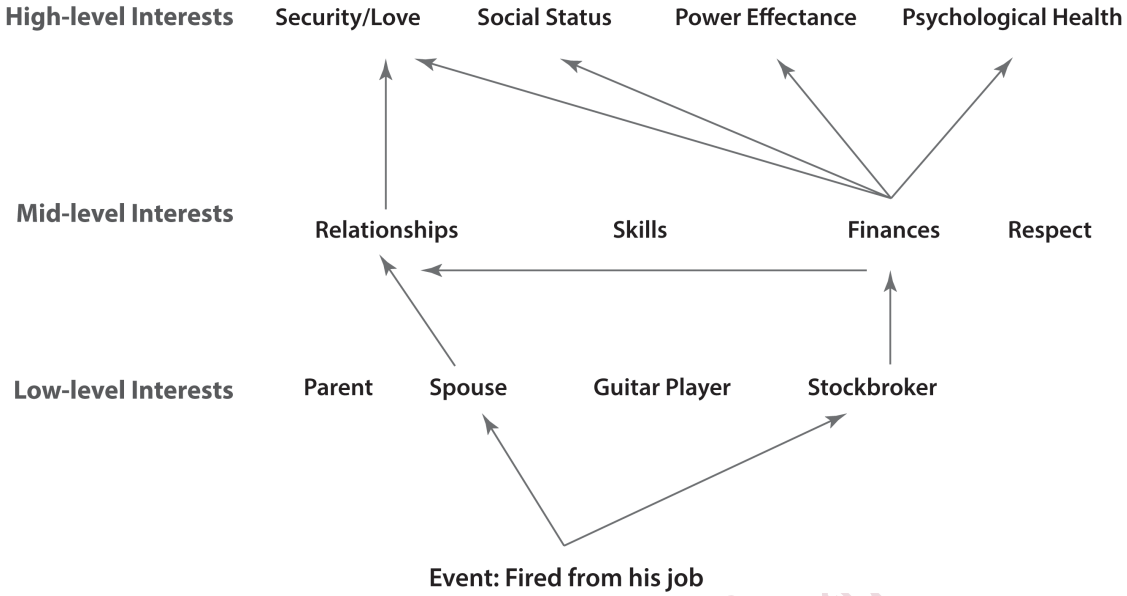


Outline of the mood-management hypothesis. Task, person, and situation features jointly influence the kind of information-processing strategy adopted to deal with a social task. Substantive processing produces affect infusion and mood congruity, whereas motivated processing produces affect control. Social actors automatically switch between these two processing strategies to maintain a balanced affective state. Thus, substantive processing and motivated processing operate in an interactive fashion, jointly constituting a homeostatic mood-management system.

(After Forgas 2002, p. 22)

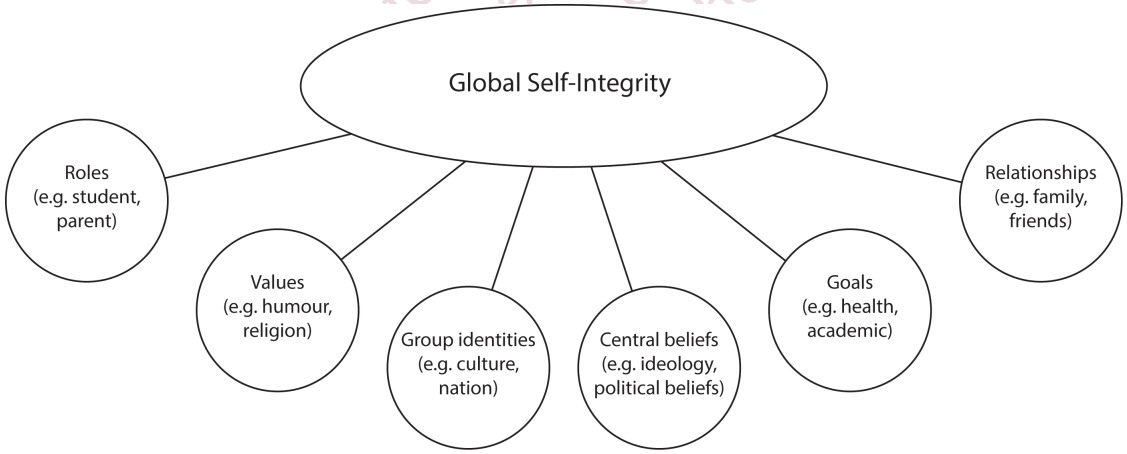
Appendix 2

Figure 19: Hypothetical Network of Interests



(Alicke & Sedikides 2009, p. 8)

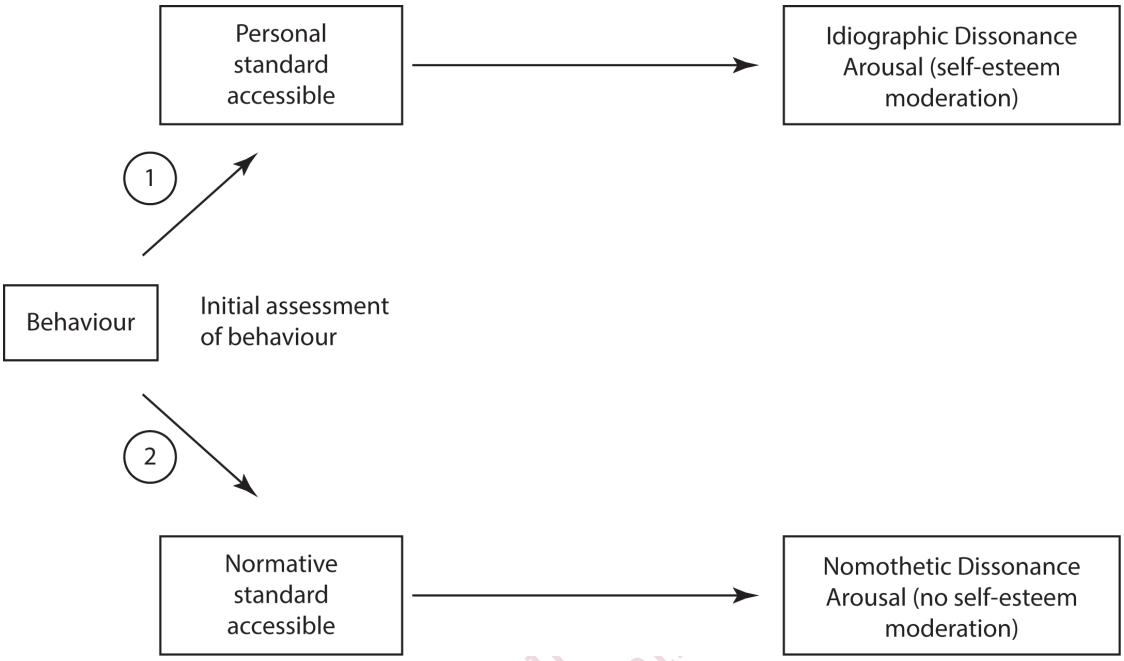
Figure 20: Schematic representation of a self-system



(After, Sherman & Cohen 2006, p. 188)

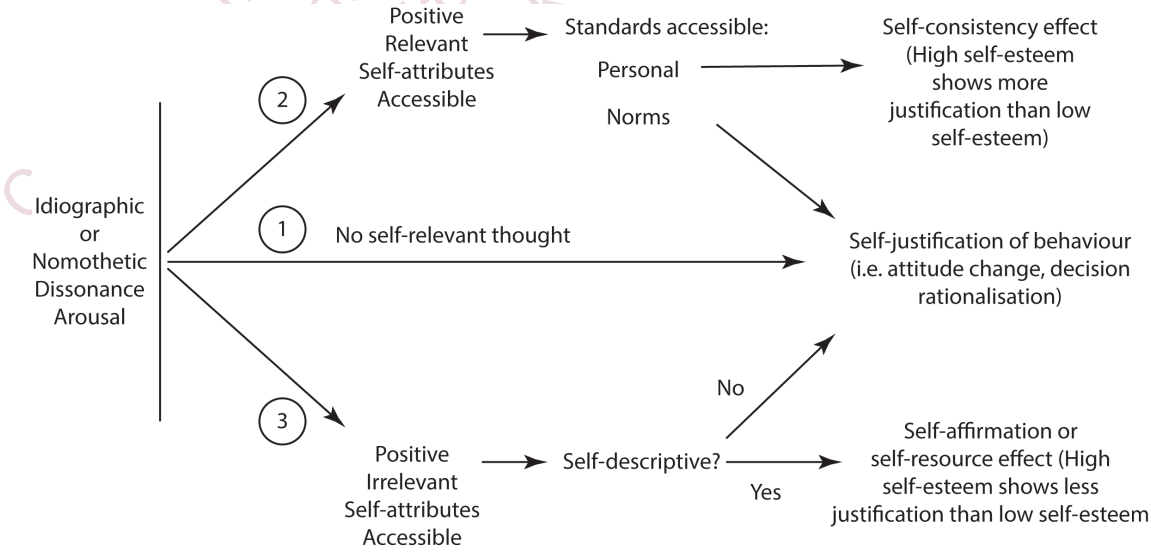
Appendix 3

Figure 21: Self-Standards Model 1



(Cooper 2007, Chapter 5, Figure 5.5)

Figure 22: Self-Standards Model 2



(ibid., Chapter 5, Figure 5.6)

References

Ackoff, RL 1999, *Ackoff's Best*, John Wiley & sons, Inc., New York

Adams, P 2007, Interview with Salmon Rushdie (repeat of program recorded in 1995), *Late Night Live*, from http://www.abc.net.au/rn/latenightlive/index/audio.htm, 5 October 2007

Alicke, MD & Sedikides, C 2009, Self-enhancement and self-protection: What they are and what they do, *European Review of Social Psychology*, Vol. 20**,** No. 1, pp. 1-48, accessed 2 May 2014 from Taylor & Francis Online, doi:10.1080/10463280802613866

Alicke, MD & Sedikides, C (Eds.) 2011a, *Handbook of self-enhancement and self-protection*, Adobe Digital Editions version, The Guildford Press, New York, accessed 5 May 2014

——— 2011b, Self-enhancement and self-protection: Historical overview and conceptual framework, in *Handbook of self-enhancement and self-protection*, (eds, Alicke, MD & Sedikides, C), Adobe Digital Editions version, The Guildford Press, New York, pp. 1-19 accessed 5 May 2014

Anderson, AK 2005, Affective Influences on the Attentional Dynamics Supporting Awareness, *Journal of Experimental Psychology: General*, Vol. 134**,** No. 2, pp. 258-281

Anderson, AK & Phelps, EA 2001, Lesions of the Human Amygdala Impair Enhanced Perception of Emotionally Salient Events, *Nature*, Vol. 411**,** pp. 305-309

Aronson, E 1999, Dissonance, Hypocrisy, and the Self-Concept, in *Cogntive Dissonance: Progress on a pivotal theory in social psychology*, (eds, Harmon-Jones, E & Mills, J), American Psychological Association, Washington, DC, pp. 103-126

Aronson, J, Cohen, G & Nail, PR 1999, Self-Affirmation Theory: An Update and Appraisal, in *Cogntive Dissonance: Progress on a pivotal theory in social psychology*, (eds, Harmon-Jones, E & Mills, J), American Psychological Association, Washington, DC, pp. 127-147

Atkinson, AP & Adolphs, R 2005, Visual Emotion Perception: Mechanisms and Processes, in *Emotion and Consciousness*, (eds, Barrett, LF, Niedenthal, PM & Winkielman, P), The Guilford Press, New York, pp. 150-182

Audi, R 2011, *Epistemology: A Contemporary Introduction to the Theory of Knowledge*, 3rd edn, Kindle version, Routledge, New York

Baddeley, A 2007, *Working Memory, Thought, and Action*, Oxford University Press, Oxford

Bandler, R & Grinder, J 1982, *Reframing*, Real People Press, Moab, UT

Barrett, LF, Lindquist, KA & Gendron, M 2007, Language as context for the perception of emotion, *TRENDS in Cognitive Sciences*, Vol. 11**,** No. 8, pp. 327-332

Barrett, LF, Niedenthal, PM & Winkielman, P (Eds.) 2005a, *Emotion and Consciousness*, The Guilford Press, New York

——— 2005b, Introduction, in *Emotion and Consciousness*, (eds, Barrett, LF, Niedenthal, PM &

Winkielman, P), The Guilford Press, New York, pp. 1-18

Bayne, T & Fernández, J (Eds.) 2009, *Delusion and self-deception: Affective and motivational influences on belief formation*, Psychology Press, New York

Bechara, A & Damasio, AR 2005, The somatic marker hypothesis: a nerual theory of economic decision, *Games and Economic Behavior*, Vol. 52**,** pp. 336-372

Beedie, CJ, Terry, PC & Lane, AM 2005, Distinctions between Emotion and Mood, *Cognition and Emotion*, Vol. 19**,** No. 6, pp. 847-878

Ben-Ze'ev, A 2004, Emotion as a subtle mental mode, in *Thinking about feeling: Contemporary philosophers on emotions*, (ed, Solomon, RC), Oxford Unversity Press, Oxford, chapter 16, accessed 16 March 2014 from www.amazon.com

Bennett, MR & Hacker, PMS 2003, *Philosophical Foundations of Neuroscience*, Blackwell Publishing, Malden, MA

Berkeley, G 2009, *Principles of Human Knowledge and Three Dialogues,* Oxford University Press, Oxford

Blackburn, S 2005, *Truth: A Guide for the Perplexed*, Adobe Digital Editions version, Allen Lane, London, accessed 26 March 2009 from www.ebooks.com

Bless, H & Fiedler, K 2006, Mood and the Regulation of Information Processing; and Behavior, in *Affect in Social Thinking and Behavior*, (ed, Forgas, JP), Kindle 2012, Pyschology Press, New York, pp. 64-84 from www.amazon.com

Boroditsky, L 2001, Does Language Shape Thought?: Mandarin and English Speakers' Conception of Time, *Cognitive Psychology*, Vol. 43**,** pp. 1-22

Bortolotti, LI 2010, *Delusions and Other Irrational Beliefs*, Oxford University Press, Oxford

Bower, GH 1981, Mood and Memory, *American Psychologist*, Vol. 36**,** No. 2, pp. 129-148, doi:0003-066X/81/3602-0129S00.75

Bower, GH 1991, Mood Congruity of Social Judgments, in *Emotion and Social Judgments*, (ed, Forgas, JP), Pergamon Press, Oxford, pp. 31-54

Bruner, JS & Goodman, CC 1947, Value and Need as Organizing Factors in Perception, *The Journal of Abnormal and Social Psychology*, Vol. 42**,** No. 1, pp. 33-44

Chabris, C & Simons, DJ 2010, *The Invisible Gorilla: And Other Ways Our Intuition Deceives Us*, Kindle version, HarperCollins, London, accessed 19 December 2010 from www.amazon.com

Chapman, GB & Johnson, EJ 2002, Incorporating the Irrelevant: Anchors in Judgments of Belief and Value, in *Heuristics and Biases: The Psychology of Intuitive Judgment*, (eds, Gilovich, T, Griffin, D & Kahneman, D), Cambridge University Press, Cambridge, pp. 120-138

Charland, LC 2005, Emotion Experience and the Indeterminacy of Valence, in *Emotion and Consciousness*, (eds, Barrett, L, Niedenthal, PM & Winkielman, P), The Guildford Press, New York, pp. 231-254

Churchland, PS & Churchland, PM 2013, What are Beliefs?, in *The Neural Basis of Human Belief Systems*, (eds, Krueger, F & Grafman, J), Kindle version, Psychology Press, Hove. East Sussex, chapter 1, accessed 21 April 2014 from www.amazon.com

Cialdini, RB 1984, *Influence: The Psychology of Persuasion*, The Business Library, Melbourne

Clark, MS & Brissette, I 2000, Relationship beliefs and emotion: Reciprocal effects, in *Emotions and Beliefs*, (eds, Frijda, NH, Manstead, ASR & Bem, S), Cambridge University Press, Cambridge, pp. 212-240

Clifford, WK 1999, *The Ethics of Belief and Other Essays*, Prometheus Books, New York

Clore, GL & Gasper, K 2000, Feeling is believing: Some affective influences on belief, in *Emotions and Beliefs*, (eds, Frijda, NH, Manstead, ASR & Bem, S), Cambridge University Press, Cambridge, pp. 10-44

Clore, GL, Gasper, K & Garvin, E 2001, Affect as Information, in *Handbook of Affect and Social Cognition*, (ed, Forgas, JP), Taylor & Francis e-Library 2009, Lawrence Erlbaum Associates, Inc., Mahwah, NJ, pp. 122-145 from www.amazon.com

Clore, GL & Huntsinger, JR 2007, How emotions inform judgment and regulate thought, *TRENDS in Cognitive Science*, Vol. 11**,** No. 9, pp. 393-399, doi:10.1016/j.tics.2007.08.005

Clore, GL & Ortony, A 2008, Appraisal Theories: How Cognition Shapes Affect into Emotion, in *Handbook of Emotions*, (eds, Lewis, M, Haviland-Jones, JM & Barrett, LF), 3rd, The Guilford Press, New York, pp. 628-642

Clore, GL & Palmer, J 2009, Affective guidance of intelligent agents: How emotion controls cognition, *Cognitive Systems Research*, Vol. 10**,** pp. 21-30, accessed 7 April 2014 from Elsevier ScienceDirect, doi:10.1016/j.cogsys.2008.03.002

Clore, GL & Storbeck, J 2006, Affect as Information about Liking, Efficacy, and Importance, in *Affect in Social Thinking and Behavior*, (ed, Forgas, JP), Kindle 2012, Pyschology Press, New York, pp. 122-142 from www.amazon.com

Coltheart, M & Davies, M (Eds.) 2000, *Pathologies of Belief*, Blackwell Publishers, Oxford

Conway, MA & Pleydell-Pearce, CW 2000, The Construction of Autobiographical Memories in the Self-Memory System, *Psychological Review*, Vol. 107**,** No. 2, pp. 261-288, doi:10.1037/0033-295X.107.2.261

Cooper, J 2007, *Cognitive Dissonance: Fifty Years of a Classic theory*, Kindle version, Sage Publications, London, accessed 2 July 2013 from www.amazon.com

Craig, AD 2003, A new view of pain as a homeostatic emotion, *TRENDS in Cognitive Science*, Vol. 26**,** No. 6, pp. 303-307, accessed 28 March 2014 from Elsevier ScienceDirect, doi:10.1016/S0166-2236(03)00123-1

——— 2008, Interoception and Emotion, in *Handbook of Emotions*, (eds, Lewis, M, Haviland-Jones, JM & Barrett, LF), 3rd, The Guilford Press, New York, pp. 272-288

Critchley, HD, Mathias, CJ & Dolan, RJ 2001, Neuroanatomical basis for first- and second-order

representations of bodily states, *Nature Neuroscience*, Vol. 4**,** No. 2, pp. 207-212

Critchley, HD, Wiens, S, Rotshtein, P, Öhman, A & Dolan, RJ 2004, Neural systems supporting interoceptive awareness, *Nature Neuroscience*, Vol. 7**,** No. 2, pp. 189-195

Cromby, J 2007, *Toward a Psychology of Feeling*, *International Journal of Critical Psychology*, **21**, 94-118., http://homepages.lboro.ac.uk/~hujc4/Toward%20a%20Psychology%20of%20Feeling.pdf, Accessed 23 March 2014

——— 2012, Beyond Belief, *Journal of Health Psychology*, Vol. 17**,** No. 7, pp. 943-957, accessed 4 September 2012 from Sage Journals, doi:10.1177/1359105312448866

Damasio, A 1994, *Descartes' Error: Emotion, Reason, and the Human Brain*, Quill, New York

——— 1999, *The Feeling of What Happens*, William Heinemann, London

——— 2000a, A Second Chance for Emotion, in *Cognitive Neuroscience of Emotion*, (eds, Lane, RD & Nadel, L), Oxford University Press, New York, pp. 12-23

——— 2000b, Thinking about Belief: Concluding Remarks, in *Memory, Brain and Belief*, (eds, Schacter, DL & Scarry, E), Harvard University Press, Cambridge, Massachusetts, pp. 325-333

——— 2003a, Feelings of Emotion and the Self, *Annals of the New York Academy of Sciences*, Vol. 1001**,** pp. 253-261

——— 2003b, *Looking for Spinoza*, Vintage, London

——— 2004, Emotions and Feelings: A Neuolobiological Perspective, in *Feelings and Emotions: The Amsterdam Symposium*, (eds, Manstead, ASR, Frijda, N & Fischer, A), Cambridge University Press, Cambridge, pp. 49-57

——— 1998, Commentary on "Mind, Body, and Mental Illness, *Philosophy, Psychiatry, & Psychology*, Vol. 5**,** No. 4, p. 343

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

Damasio, A & Carvalho, GB 2013, The nature of feelings: evolutionary and neurobiological origins, *Nature Reviews Neuroscience*, Vol. 14**,** pp. 143-152, accessed 21 April 2014 from Nature, doi:10.1038/nrn3403

Damasio, A & Damasio, H 2006, Minding the Body, *Daedalus*, Vol. 135**,** No. 3, pp. 15-22

Damasio, A, Everitt, BJ & Bishop, D 1996, The somatic marker hypothesis and the possible functions of the prefrontal cortex, *Philosophical Trasnactions: Biological Sciences*, Vol. 351**,** No. 1346, pp. 1413-1420

Damasio, A *et al.* 2000, Subcortical and cortical brain activity during the feeling of self-generating emotions, *Nature Neuroscience*, Vol. 3**,** p. 1049

Davies, M & Coltheart, M 2000, Introduction: Pathologies of Belief, in *Pathologies of Belief*, (eds, Coltheart, M & Davies, M), Blackwell Publishers, Oxford, pp. 1-46

Day, A 2010, Propositions and performativity: Relocating belief to the social, *Culture and Religion*,

Vol. 11, No. 1, pp. 9-30, accessed 22 March 2014 from Taylor & Francis Online, doi:10.1080/14755610903528812

Day, A & Coleman, S 2010, Introduction: Broadening boundaries: creating inter-disciplinary dialogue on belief, *Culture and Religion*, Vol. 11, No. 1, pp. 1-8, accessed 24 March 2014 from Taylor & Francis Online, doi:10.1080/14755610903528952

De Preester, H 2007, The deep bodily origins of the subjective perspective: Models and their problems, *Consciousness and Cognition*, Vol. 16, pp. 604-618

Dennett, DC 1987, *The Intentional Stance*, 1989 Paperback edn, The MIT Press, Cambridge, MA

——— 1990, Three Kinds of Intentional Psychology, in *Foundations of Cognitive Science: The Essential Readings*, (ed, Garfield, JL), Paragon House, New York, pp. 88-110

——— 1995, Review of Damasio, *Descartes' Error*, *Times Literary Supplement*, Vol. August 25, pp. 3-4 from http://ase.tufts.edu/cogstud/papers/damasio.htm

DeSteno, DA & Salovey, P 1997, The effects of mood on the structure of the self-concept, *Cognition & Emotion*, Vol. 11, No. 4, pp. 351-372, accessed 30 April 2014 from Taylor & Francis Online, doi:10.1080/026999397379836

Dreisbach, G & Goschke, T 2004, How Positive Affect Modulates Cognitive Control: Reduced Perseveration at the Cost of Increased Distractibility, *Journal of Experimental Psychology: Learning, Memory and Cognition*, Vol. 30, No. 2, pp. 343-353

Drivonikou, GV *et al.* 2007, Further evidence that Whorfian effects are stronger in the right visual field than the left, *PNAS*, Vol. 104, No. 3, pp. 1097-1102

Droit-Volet, S & Gil, S 2009, The Time-Emotion Paradox, *Philosophical Transactions of the Royal Society, B*, Vol. 364, pp. 1943-1953, doi:10.1098/rstb.2009.0013

Droit-Volet, S & Meck, WH 2007, How Emotions Colour Our Perception of Time, *Trends in Cognitive Sciences*, Vol. 11, No. 12, pp. 504-513

Eakin, PJ 2000, Autobiography, Identity and the Fictions of Memory, in *Memory, Brain and Belief*, (eds, Schacter, DL & Scarry, E), Harvard University Press, Cambridge, Massachusetts, pp. 290-306

Eich, E & Macauley, D 2000, Fundamental Factors in Mood-Dependent memory, in *Feeling and Thinking: The Role of Affect in Social Cognition*, (ed, Forgas, JP), Cambridge University Press, Cambridge, pp. 109-130

Eichenbaum, H 2002, *The Cognitive Neuroscience of Memory: an Introduction*, Oxford University Press, Oxford

Eichenbaum, H & Bodkin, JA 2000, Belief and Knowledge as Distinct Forms of Memory, in *Memory, Brain and Belief*, (eds, Schacter, DL & Scarry, E), Harvard University Press, Cambridge, Massachusetts, pp. 176-207

Ekman, P 1992, An Argument for Basic Emotions, *Cognition and Emotion*, Vol. 6, No. 3/4, pp. 169-200, accessed 22 May 201 from Informa, doi:10.1080/02699939208411068

——— 2001, *Telling Lies*, W W Norton & Company, New York

——— 2003, *Emotions Revealed*, Henry Holt and Company, New York

Elster, J 2004, Emotions and Rationality, in *Feelings and Emotions*, (eds, Manstead, ASR, Frijda, N & Fischer, A), Cambridge University Press, Cambridge, pp. 30-48

Eysenck, MW 2000, Anxiety, cognitive biases, and beliefs, in *Emotions and Beliefs*, (eds, Frijda, NH, Manstead, ASR & Bem, S), Cambridge University Press, Cambridge, pp. 171-184

——— 2006, *Fundamentals of Cognition*, Psychology Press, Hove, East Sussex

Fiedler, K & Bless, H 2000, The formation of beliefs at the interface of affective and cognitive processes, in *Emotions and Beliefs*, (eds, Frijda, NH, Manstead, ASR & Bem, S), Cambridge University Press, Cambridge, pp. 144-170

Finke, RA 1980, Levels of Equivalence in Imagery and Perception, *Psychological Review*, Vol. 87**,** No. 2, pp. 113-132

Finke, RA, Pinker, S & Farah, MJ 1989, Reinterpreting Visual Patterns in Mental Imagery, *Cognitive Science*, Vol. 13**,** pp. 51-78

Fogassi, L 2011, The mirror neuron system: How cognitive functions emerge from motor organization, *Journal of Economic Behavior & Organization*, Vol. 77**,** pp. 66-75, accessed 25 April 2014 from Elsevier ScienceDirect, doi:10.1016/j.jebo.2010.04.009

Forgas, JP 1995, Mood and Judgment: The Affect Infusion Model (AIM), *Psychological Bulletin*, Vol. 117**,** No. 1, pp. 39-66

Forgas, JP (Ed.) 2000a, *Feeling and Thinking: The Role of Affect in Social Cognition*, Cambridge University Press, Cambridge

——— 2000b, Feeling is believing? The role of processing strategies in mediating affective influences on beliefs, in *Emotions and Beliefs*, (eds, Frijda, NH, Manstead, ASR & Bem, S), Cambridge University Press, Cambridge, pp. 108-143

——— 2000c, Managing moods: Toward a dual-process theory of spontaneous mood regulation, *Psychological Inquiry*, Vol. 11**,** No. 3, pp. 172-177, accessed 11 May 2014 from Taylor & Francis Online

Forgas, JP (Ed.) 2001, *Handbook of affect and social cognition*, Kindle version, Taylor & Francis e-Library 2009, Lawrence Erlbaum Associates, Inc., Mahwah, NJ, accessed 24 April 2014 from www.amazon.com

——— 2002, Feeling and Doing: Affective Influences on Interpersonal Behavior, *Psychological Inquiry*, Vol. 13**,** No. 1, pp. 1-28

Forgas, JP (Ed.) 2006, *Affect in social thinking and behavior*, Kindle version, Pyschology Press, New York, accessed 13 May 2013 from www.amazon.com

——— 2013, Belief and affect: On the mental pre-cursors of health-related cognition and behaviour, *Journal of Health Psychology*, Vol. 18**,** No. 1, pp. 3-9,

doi:10.1177/1359105312448869

Forgas, JP & Ciarrochi, JV 2002, On Managing Moods: Evidence for the Role of Homeostatic Cognitive Strategies in Affect Regulation, *Personality and Social Psychology Bulletin*, Vol. 28**,** pp. 336-345, accessed 9 May 2014 from Sage Journals, doi:10.1177/0146167202286005

Frijda, NH 1993, The Place of Appraisal in Emotion, *Cognition and Emotion*, Vol. 7**,** No. 3/4, pp. 357-387, doi:10.1080/02699939308409193

——— 2004, Emotions and Action, in *Feelings and Emotions*, (eds, Manstead, ASR, Frijda, NH & Fischer, A), Cambridge University Press, Cambridge, pp. 158-173

——— 2005, Emotion Experience, *Cognition and Emotion*, Vol. 19**,** No. 4, pp. 473-497

Frijda, NH, Manstead, ASR & Bem, S (Eds.) 2000a, *Emotions and Beliefs*, Cambridge University Press, Cambridge

——— 2000b, The influence of emotions on beliefs, in *Emotions and Beliefs*, (eds, Frijda, NH, Manstead, ASR & Bem, S), Cambridge University Press, Cambridge, pp. 1-9

Frijda, NH, Manstead, ASR & Fischer, A 2004, Epilogue: Feelings and Emotions: Where Do We Stand?, in *Feelings and Emotions*, (eds, Manstead, ASR, Frijda, NH & Fischer, A), Cambridge University Press, Cambridge, pp. 455-467

Frijda, NH & Mesquita, B 2000, Beliefs through emotions, in *Emotions and Beliefs*, (eds, Frijda, NH, Manstead, ASR & Bem, S), Cambridge University Press, Cambridge, pp. 45-77

Frith, C & Dolan, RJ 2000, The Role of Memory in the Delustions Associated with Schizophrenia, in *Memory, Brain and Belief*, (eds, Schacter, DL & Scarry, E), Harvard University Press, Cambridge, Massachusetts, pp. 115-135

Gallagher, S 2000, Philosophical conceptions of the self: implications for cognitive science, *TRENDS in Cognitive Science*, Vol. 4**,** No. 1, pp. 14-21, accessed 21 April 2014 from Elsevier ScienceDirect

Gallagher, S & Zahavi, D 2010, Phenomenological Approaches to Self-Consciousness, *The Stanford Encyclopedia of Philosophy*, (ed, Zalta, EN), Winter 2010 Edition, http://plato.stanford.edu/archives/win2010/entries/self-consciousness-phenomenological/, Accessed 21 April 2014

Gallese, V 2001, The 'Shared Manifold' Hypothesis, *Journal of Consciousness Studies*, Vol. 8**,** No. 5–7, pp. 33-50, accessed 25 April 2014 from Imprint Academic

Gallese, V & Goldman, AI 1998, Mirror Neurons and the Simulation Theory of Mind Reading, *Trends in Cognitive Science*, Vol. 2**,** No. 12, pp. 493-501 from doi:10.1016/S1364-6613(98)01262-5

Garfield, JL 1988, *Belief in Psychology*, The MIT Press, Cambridge, Massachusetts

Gasper, K 2004, Permission to Seek Freely? The Effect of Happy and Sad Moods on Generating Old and New Ideas, *Creativity Research Journal*, Vol. 16**,** No. 2–3, pp. 215-229, accessed 24 March 2014 from Informa, doi:10.1080/10400419.2004.9651454

Gasper, K & Bramesfeld, KD 2006, Should I follow my feelings? How individual differences in following feelings influence affective well-being, experience, and responsiveness, *Journal of Research in Personality*, Vol. 40**,** pp. 986-1014, accessed 24 March 2014 from Elsevier ScienceDirect, doi:10.1016/j.jrp.2005.10.001

Gasper, K & Clore, GL 2002, Mood and Global Versus Locual Processing of Visual Information, *Psychological Science*, Vol. 13**,** No. 1, pp. 34-40

Gazzaniga, MS 1998, *The Mind's Past*, electronic edn, University of California Press, Berkeley, CA, accessed 12 October 2009 from www.ebookmall.com

Gazzoa, V, Aziz-Zadeth, L & Keysers, C 2006, Empathy and the somatotopic auditory mirror system in humans, *Current Biology*, Vol. 19**,** pp. 1824-1829, accessed 25 April 2014 from Elsevier ScienceDirect, doi:10.1016/j.cub.2006.07.072

Gazzola, V, Rizzolatti, G, Wicker, B & Keysers, C 2007, The anthropomorphic brain: The mirror neuron system responds to human and robotic actions, *NeuroImage*, Vol. 35**,** pp. 1674-1684, accessed 25 April 2014 from Elsevier ScienceDirect, doi:10.1016/j.neuroimage.2007.02.003

Gigerenzer, G 2007, *Gut Feelings: The Intelligence of the Unconscious*, Viking Penguin, London

Gigerenzer, G, Todd, PM & ABC Research Group 1999, *Simple Heuristics That Make Us Smart*, Oxford University Press, Oxford

Gilbert, AL, Regier, T, Kay, P & Ivry, RB 2006, Whorf hypothesis is supported in the right visual field but not the left, *PNAS*, Vol. 103**,** No. 2, pp. 489-494

Gilbert, DT 1991, How Mental Systems Believe, *American Psychologist*, Vol. 46**,** No. 2, pp. 107-119

Gilbert, DT, Krull, DS & Malone, PS 1990, Unbelieving the Unbelievable: Some Problems in the Rejection of False Information, *Journal of Personality and Social Psychology*, Vol. 59**,** No. 4, pp. 601-613

Gilbert, DT, Tafarodi, RW & Malone, PS 1993, You Can't Not Believe Everything You Read, *Journal of Personality and Social Psychology*, Vol. 65**,** No. 2, pp. 221-233

Gilovich, T 1991, *How We Know What Isn't So: The Fallibility of Human Reason in Evergyday Life*, The Free Press, New York

Gilovich, T & Griffin, D 2002, Introduction – Heuristics and Biases: Then and Now, in *Heuristics and Biases: The Psychology of Intuitive Judgment*, (eds, Gilovich, T, Griffin, D & Kahneman, D), Cambridge University Press, Cambridge, pp. 1-18

Goldman, AI 1986, *Epistemology and Cognition*, Harvard University Press, Cambridge, Massachusetts

Gray, JR 2004, Integration of Emotion and Cognitive Contro, *Current Directions in Psychological Science*, Vol. 13**,** pp. 46-48, doi:DOI: 10.1111/j.0963-7214.2004.00272.x

Griffin, D & Tversky, A 2002, The Weighing of Evidence and the Determinants of Confidence, in

Introduction – Heuristics and Biases: Then and Now, (eds, Gilovich, T, Griffin, D & Kahneman, D), Cambridge University Press, Cambridge, pp. 230-249

Hammond, C 2012, *Time Warped: Unlocking the Mysteries of Time Perception*, Kindle edn, Canongate Books Ltd, Edinburgh from www.amazon.com

Harmon-Jones, E 2000, A cognitive dissonance theory perspective on the role of emotion in the maintenance and change of beliefs and attitudes, in *Emotions and Beliefs*, (eds, Frijda, NH, Manstead, ASR & Bem, S), Cambridge University Press, Cambridge, pp. 185-211

——— 2001, The role of affect in cognitive-dissonance processes, in *Handbook of affect and social cognition*, (ed, Forgas, JP), Kindle version, Taylor & Francis e-Library 2009, Lawrence Erlbaum Associates, Inc., Mahwah, NJ, pp. 239-257, chapter 11, accessed 24 April 2014 from www.amazon.com

Harmon-Jones, E & Mills, J (Eds.) 1999, *Cogntive Dissonance: Progress on a pivotal theory in social psychology*, American Psychological Association, Washington, DC

Hirstein, W 2000, *Self-Deception and Confabulation*, Vol 67, Supplement. Proceedings of the 1998 Biennial Meetings of the Philosophy of Science Association. Part II: Symposia Papers (Sep., 2000) version, The University of Chicago Press,

——— 2005, *Brain Fiction*, The MIT Press, Cambridge, Massachusetts

Hirstein, W (Ed.) 2009, *Confabulation*, Oxford University Press, Oxford

Hollingham, R 2004, In the Realm of Your Senses, *New Scientist*, (2432), Issue 2432, http://www.newscientist.com/article/mg18124326.100-in-the-realm-of-your-senses.html, Accessed 4 October 2011

Hume, D 1739,1740/1985, *A Treatise of Human Nature*, Penguin Books, London

James, IA & Barton, S 2004, Changing Core Beliefs with the Continuum Technique, *Behavioural and Cognitive Psychotherapy*, Vol. 32, pp. 431-442

James, W 1890/1950a, *The Principles of Psychology*, Vol 2, Authorised edn, Dover Publications, Inc., New York

——— 1890/1950b, *The Principles of Psychology*, Vol 1, Authorised edn, Dover Publications, Inc., New York

——— 1956, *The Will to Believe and Other Essays in Popular Philosophy*, Dover Publications, Inc., New York

Janis, IL & Mann, L 1977, *Decision Making: A Psychological Analysis of Conflict, Choice, and Commitment*, The Free Press, New York

Jha, A 2005, *Where belief is born*, Guardian, http://www.guardian.co.uk/science/2005/jun/30/psychology.neuroscience, Accessed 9 Feb 2007

Johnson, MK & Raye, CL 2000, Cognitive and Brain Mechanisms of False Memories and Beliefs, in *Memory, Brain and Belief*, (eds, Schacter, DL & Scarry, E), Harvard University Press, Cambridge, Massachusetts, pp. 35-86

Johnson, M 2007, *The meaning of the body: Aesthetics of human understanding*, Kindle version, The University of Chicago Press, Chicago, accessed 27 April 2014 from www.amazon.com

Kahneman, D 2003a, A Perspective on Judgment and Choice, *American Psychologist*, Vol. 58**,** No. 9, pp. 697-720 from DOI: 10.1037/0003-066X.58.9.697

——— 2003b, A Perspective on Judgment and Choice: Mapping Bounded Rationality, *American Psychologist*, Vol. 58**,** No. 9, pp. 697-720

——— 2011, *Thinking, fast and slow*, Kindle version, Allen Lane, London, accessed 21 December 2011 from www.amazon.com

Kahneman, D & Frederick, S 2002, Representativeness Revisited: Attribute Substitution in Intuitive Judgment, in *Heuristics and Biases: The Psychology of Intuitive Judgment*, (eds, Gilovich, T, Griffin, D & Kahneman, D), Cambridge University Press, Cambridge, pp. 49-81

Kahneman, D, Slovic, P & Tversky, A (Eds.) 1982, *Judgment under uncertainty: Heuristics and biases*, Cambridge University Press, Cambridge, NY

Kahneman, D & Tversky, A 1982, Judgment under uncertainty: Heuristics and biases, in *Judgment under uncertainty: Heuristics and biases*, (eds, Kahneman, D, Slovic, P & Tversky, A), Cambridge University Press, Cambridge, pp. 3-20

——— (Eds.) 2000, *Choices, Values and Frames*, Cambridge University Press, Cambridge

Klein, G 2005, *The Power of Intuition*, Currency (Random House), New York

Knapp, HP & Corina, DP 2010, A human mirror neuron system for language: Perspectives from signed languages of the deaf, *Brain & Language*, Vol. 112**,** pp. 36-43, accessed 25 April 2014 from Elsevier ScienceDirect, doi:10.1016/j.bandl.2009.04.002

Kosslyn, SM 1980, *Image and Mind*, Harvard University Press, Cambridge, MA

——— 1983, *Ghosts in the Mind's Machine: Creating and Using Images in the Brain*, W W Norton & Company, New York

Lambie, JA & Marcel, AJ 2002, Consciousness and the Varieties of Emotion Experience: A Theoretical Framework, *Psychological Review*, Vol. 109**,** No. 2, pp. 219-259, accessed 24 March 2014 from OVID PsycARTICLES

Lazarus, RS 1982, Thoughts on the Relation Between Emotion and Cognition, *American Psychologist*, Vol. 37**,** No. 9, pp. 1019-1024

——— 1984, On the Primacy of Cognition, *American Psychologist*, Vol. 39**,** No. 2, pp. 124-129

LeDoux, J 1998, *The Emotional Brain*, Phoenix, London

——— 2000, Emotion Circuits in the Brain, *Annual Review of Neuroscience*, Vol. 23**,** pp. 155-184

——— 2012, Rethinking the Emotional Brain, *Neuron*, Vol. 73**,** pp. 653-676, doi:10.1016/j.neuron.2012.02.004

——— 1993, Cognition versus Emotion, Again- This Time in the Brain: A Response to Parrott

and Schulkin, *Cognition and Emotion*, Vol. 7**,** No. 1, pp. 61-64

——— 2002, *Synaptic Self*, Penguin Books, New York

Legrand, D 2007, Pre-reflective self-as-subject from experiential and empirical perspective, *Consciousness and Cognition*, Vol. 16**,** pp. 583-599

Lepper, MR, Ross, L & Lau, RR 1986, Persistence of Inaccurate Beliefs About the Self: Perseverance Effects in the Classroom, *Journal of Personality and Social Psychology*, Vol. 50**,** No. 3, pp. 482-491

Lerner, JS & Keltner, D 2000, Beyond valence: Toward a model of emotion-specific influences on judgement and choice, *Cognition and Emotion*, Vol. 14**,** No. 4, pp. 473-493

Leventhal, H & Scherer, KR 1987, The Relationship of Emotion to Cognition: A Functional Approach to a Semantic Controversy, *Cognition and Emotion*, Vol. 1**,** No. 1, pp. 3-28

Levin, DT 2002, Change Blindness Blindness As Visual Metacognition, *Journal of Consciousness Studies*, Vol. 9**,** No. 5–6, pp. 111-130

Levin, DT, Drivdahl, SB, Momen, N & Beck, MR 2002, False predictions about the detectability of visual changes: The role of beliefs about attention, memory, and the continuity of attended objects in causing change blindness blindness, *Consciousness and Cognition*, Vol. 11**,** pp. 507-527

Levin, DT, Momen, N, Drivdahl, SB & Simons, DJ 2000, Change Blindness Blindness: The Metacognitive Error of Overestimating Change-detection Ability, *Visual Cognition*, Vol. 7**,** No. 1–3, pp. 397-412

Levine, LJ & Pizarro, DA 2004, Emotion and Memory Research: A Grumpy Overview, *Social Cognition*, Vol. 22**,** No. 5, pp. 530-554

——— 2006, Emotional Valence, Discrete Emotions, and Memory, in *Memory and Emotion*, (eds, Uttl, B, Ohta, N & Siegenthaler, AL), Blackwell Publishing, Malden, MA, pp. 37-58

Levinson, SC, Kita, S, Haun, DBM & Rasch, BH 2002, Returning the Tables: Language Affects Spatial Reasoning, *Cognition*, Vol. 84**,** pp. 155-188

Lindström, BR & Bohlin, G 2011, Emotion processing facilitates working memory performance, *Cognition & Emotion*, Vol. 25**,** No. 7, pp. 1196-1204

Linville, PW 1985, Self-Complexity and Affective Extremity: Don't put all of your eggs in one cognitive basket, *Social Cognition*, Vol. 3**,** No. 1, pp. 94-120

Locke, J 1690, *An Essay Concerning Humane Understanding*, Vol 2, 2nd edn, Kindle version, Produced by Steve Harris and David Widger, accessed 29 February 2012 from www.amazon.com

Macknik, SL & Martinez-Conde, S 2010, *Sleights of Mind*, Henry Holt and Company, LLC, New York

Manstead, ASR, Frijda, N & Fischer, A (Eds.) 2004, *Feelings and Emotions*, Cambridge University Press, Cambridge

Markowitsch, HJ 2000, Neuroanatomy of Memory, in *The Oxford Handbook of Memory*, (eds, Tulving, E & Craik, FIM), Oxford University Press, Oxford, pp. 465-484

Markus, GE & Mackuen, MB 1993, Anxiety, Enthusiasm, and the Vote: The Emotional Underpinnings of Learning and Involvement During Presidential Campaigns, *The American Political Science Review*, Vol. 87**,** No. 3, pp. 672-685

Mazzoni, G & Kirsch, I 2002, Autobiographical memories and beliefs: a preliminary metacognitive model, in *Applied Metacognition*, (eds, Perfect, TJ & Scwartz, BL), Cambridge University Press, Cambridge, pp. 121-145

McDonald, RJ & White, NM 1993, A Triple Dissociation of Memory Systems: Hippocampus, Amygdala, and Dorsal Striatum, *Behavioral Neuroscience*, Vol. 107**,** No. 1, pp. 3-22, doi:10.1037/0735-7044.107.1.3

Menashe, I, Man, O, Lancet, D & Gilad, Y 2003, Different Noses for Different People, *Nature Genetics*, Vol. 34**,** No. 2, pp. 143-144

Metzinger, T 2009, *The ego tunnel: The science of the mind and the myth of the self*, Adobe Digital Editions version, Basic Books, New York, accessed 23 April 2011 from www.ebookmall.com

Mitchell, RLC & Phillips, LH 2007, The psychological, neurochemical and functional neuroanatomical mediators of the effects of positive and negative mood on executive functions, *Neuropsychologia*, Vol. 45**,** pp. 617-629

Mo, L, Xu, G, Kay, P & Tan, LH 2011, Electrophysiological evidence for the left-lateralized effect of language on preattentive categorical perception of color, *PNAS*, Vol. 108**,** No. 34, pp. 12026-14030

Mooney, C 2011, Irrationality vs Vaccines: Fighting for Reality, *New Scientist*, Vol. 2795**,** pp. 46-47, accessed 5 March 2011 from http://www.newscientist.com/article/mg20927955.400-irrationality-vs-vaccines-fighting-for-reality.html?full=true&print=true

Morris, JS, Öhman, A & Dolan, RJ 1999, A subcortical pathway to the right amygdala mediating "unseen" fear, *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 96**,** No. 4, pp. 1680-1685, accessed 25 April 2014 from Highwire Press National Academy of Sciences

Morton, A 1997, *A Guide Through the Theory of Knowledge*, 2nd edn, Blackwell Publishers Ltd, Oxford

——— 2003, *A Guide Through the Theory of Knowledge*, 3rd (Kindle) edn, Blackwell Publishers Ltd, Malden, MA

Mosca, A 2000, *A Review Essay on Antonio Damasio's The Feeling of What Happens: Body and Emotion in the Making of Consciousness*, http://psyche.cs.monash.edu.au/v6/psyche-6-10-mosca.html, Accessed 29 March 2007

Nail, PR, Misak, JE & Davis, RM 2004, Self-affirmation versus self-consistency: a comparison of two competing self-theories of dissonance phenomena, *Personality and Individual*

Differences, Vol. 36**,** pp. 1893-1905, accessed 12 May 2014 from Elsevier ScienceDirect, doi:10.1016/j.paid.2003.08.019

Nelson, K 2000, Memory and Belief in Development, in *Memory, Brain and Belief*, (eds, Schacter, DL & Scarry, E), Harvard University Press, Cambridge, Massachusetts, pp. 259-289

Niziolek, CA, Nagarajan, SS & Houde, JF 2013, What does motor efference copy represent? Evidence from speech production, *The Journal of Neuroscience*, Vol. 33**,** No. 41, pp. 16110-16116, accessed 25 April 2014 from <http://www.jneurosci.org/content/33/41/16110.full.pdf+html>

Northoff, G 2008, Is Appraisal 'Embodied' and 'Embedded'?, *Journal of Consciousness Studies*, Vol. 15**,** No. 5, pp. 68-99

——— 2011, Self and brain: what is self-relating processing?, *TRENDS in Cognitive Science*, Vol. 15**,** No. 5, pp. 186-187

Northoff, G *et al.* 2006, Self-referential processing in our brain—A meta-analysis of imaging studies on the self, *NeuroImage*, Vol. 31**,** pp. 440-457

Oatley, K 2000, The sentiments and beliefs of distributed cognition, in *Emotions and Beliefs*, (eds, Frijda, NH, Manstead, ASR & Bem, S), Cambridge University Press, Cambridge, pp. 78-107

Ocampo, B & Kritkos, A 2011, Interpreting actions: The goal behind mirron neuron function, *Brain Research Reviews*, Vol. 67**,** pp. 260-267, accessed 25 April 2014 from Elsevier ScienceDirect, doi:10.1016/j.brainresrev.2011.03.001

Öhman, A, Flykt, A & Esteves, F 2001, Emotion Drives Attention: Detecting the Snake in the Grass, *Journal of Experimental Psychology: General*, Vol. 130**,** No. 3, pp. 466-448

Öhman, A & Mineka, S 2001, Fears, Phobias, and Preparedness: Toward an Evolved Module of Fear and Fear Learning, *Psychological Review*, Vol. 108**,** No. 3, pp. 483-522, doi:10.1037//0033-295X.108.3.483

Öhman, A & Wiens, S 2004, The Concept of an Evolved Fear Module and Cognitive Theories of Anxiety, in *Feelings and Emotions*, (eds, Manstead, ASR, Frijda, N & Fischer, A), Cambridge University Press, Cambridge, pp. 58-80

Ortony, A, Clore, GL & Collins, A 1988, *The Cognitive Structure of Emotions*, Cambridge University Press, Cambridge

Overskeid, G 2000, The Slave of the Passions: Experiencing Problems and Selecing Solutions, *Review of General Psychology*, Vol. 4**,** No. 3, pp. 284-309

Oxford University Press 2007, *Shorter Oxford English Dictionary*, Electronic version 1.4 (1.4) edn, Oxford University Press, Oxford from www.wordwebsoftware.com

Palmer, SE 1975, The Effects of Contextual Scenes on the Identification of Objects, *Memory & Cognition*, Vol. 3**,** No. 5, pp. 519-526

Panksepp, J 2004, Basic Affects and the Instintcual Emotional Systems of the Brain, in *Feelings*

and Emotions, (eds, Manstead, ASR, Frijda, N & Fischer, A), Cambridge University Press, Cambridge, pp. 174-193

——— 2005a, Affective consciousness: Core emotional feelings in animals and humans, *Consciousness and Cognition*, Vol. 14**,** pp. 30-80

——— 2005b, On the Embodied Neural Nature of Core Emotional Affects, *Journal of Consciousness Studies*, Vol. 12**,** No. 8–10, pp. 158-184

——— 2012, What is an emotional feeling? Lessons about affective origins from cross-species neuroscience, *Motivation and Emotion*, Vol. 36**,** No. 1, pp. 4-15

Panksepp, J & Northoff, G 2009, The trans-species core self: The emergence of active cultural and neuro-ecological agents through self-related processing within subcortical-cortical midline networks, *Consciousness and Cognition*, Vol. 18**,** pp. 193-215

Paramel, GV, Bimler, DL & Mislavskaia, NO 2004, Colour perception in twins: individual variation beyond common genetic inheritance, *Clinical and Experimental Optometry*, Vol. 87**,** No. 4–5, pp. 305-312

Parrott, WG & Schulkin, J 1993, What Sort of System Could An Affective System Be? A Reply to LeDoux, *Cognition and Emotion*, Vol. 7**,** No. 1, pp. 65-69

Parvisi, J & Damasio, A 2001, Consciousness and the Brainstem, *Cognition*, Vol. 79**,** pp. 135-159

Phelps, EA 2005, The Interaction of Emotion and Cognition: Insights from Studies of the Human Amygdala, in *Emotion and Consciousness*, (eds, Barrett, LF, Niedenthal, PM & Winkielman, P), The Guilford Press, New York, pp. 51-66

——— 2006, Emotion and Cognition: Insights from Studies of the Human Amygdala, *Annual Review of Psychology*, Vol. 57**,** pp. 27-53

Phelps, EA, Ling, S & Carrasco, M 2006, Emotion Facilitates Perception and Potentiates the Perceptual Benefits of Attention, *Psychological Science*, Vol. 17**,** pp. 292-299

Prinz, JJ 2005, Are Emotions Feelings?, *Journal of Consciousness Studies*, Vol. 12**,** No. 8–10, pp. 9-25

——— 2006, Is Emotion a Form of Perception?, *Canadian Journal of Philosophy*, Vol. 36 Supplement 32**,** pp. 137-160

——— 2004a, *Gut Reactions*, Adobe Digital Editions version, Oxford University Press, Oxford, accessed 8 May 2010 from www.ebooks.com

——— 2004b, *Gut Reactions*, Kindle version, Oxford University Press, Oxford, accessed 23 April 2014 from www.amazon.com

Pylyshyn, Z 1999, What's in Your Mind?, in *What is Cognitive Science?*, (eds, Lepore, E & Pylyshyn, Z), Blackwell Publishers, Malden, Massachusetts, pp. 2-25

Quine, WS & Ullian, JS 1978, *The Web of Belief*, 2nd edn, McGraw-Hill, Inc., New York

Ramachandran, VS 2000, Memory and the Brain: New lessons from Old Syndromes, in *Memory, Brain and Belief*, (eds, Schacter, DL & Scarry, E), Harvard University Press, Cambridge,

Massachusetts, pp. 87-114

Ramachandran, VS & Blakeslee, S 1999, *Phantoms in the brain*, Fourth Estate, London

Ratcliffe, M 2005, The Feeling of Being, *Journal of Consciousness Studies*, Vol. 12**,** No. 8–10, pp. 43-60

Reisberg, D 1997, *Cogition: Exploring the Science of the Mind*, W W Norton & Company, New York

——— 2006, Memory for Emotional Episodes: The Strengths and Limits of Arousal-Based Accounts, in *Memory and Emotion*, (eds, Uttl, B, Ohta, N & Siegenthaler, AL), Blackwell Publishing, Malden, MA, pp. 15-36

Reisberg, D & Chambers, D 1991, Neither Pictures Nor Propositions: What Can We Learn from a Mental Image?, *Canadian Journal of Psychology*, Vol. 45**,** No. 3, pp. 336-352

Rizzolatti, G 2005, The mirror neuron system and its function in humans, *Anatomy and Embryology*, Vol. 210**,** No. 5, pp. 419-421, accessed 25 April 2014 from Springer Standard Collection

Rizzolatti, G & Craighero, L 2004, The mirron-neuron system, *Annual Review of Neuroscience*, Vol. 27**,** pp. 169-192, accessed 25 April 2014 from Annual Reviews, doi:10.1146/annurev.neuro.27.070203.144230

Rolls, ET 2000, Memory Systems in the Brain, *Annual Review of Psychology*, Vol. 51**,** pp. 599-630

Rossi, EL & Rossi, KL 2006, The neuroscience of observing consciousness & mirror neurons in therapeutic hypnosis, *American Journal of Clinical Hypnosis*, Vol. 48**,** No. 4, pp. 263-278, accessed 5 August 2012 from Taylor & Francis Online, doi:10.1080/00029157.2006.10401533

Rowe, G, Hirsch, JB & Anderson, AK 2007, Positive Affect Increase the Breadth of Attentional Selection, *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 104**,** No. 1, pp. 383-388

Russell, B 1921/2005, *The Analysis of Mind*, Dover Publications, Inc., Mineola, NY

Schacter, DL 1996, *Searching for Memory*, Basic Books, New York

Schacter, DL & Scarry, E 2000a, Introduction, in *Memory, Brain, and Belief*, (eds, Schacter, DL & Scarry, E), Harvard University Press, Cambridge, Massachusetts, pp. 1-10

——— (Eds.) 2000b, *Memory, Brain, and Belief*, Harvard University Press, Cambridge, Massachusetts

Scherer, KR 2001, Appraisal considered as a process of multilevel sequential checking, in *Appraisal Processes in Emotion: Theory, Methods, Research (Series in Affective Science)*, (eds, Scherer, KR, Schorr, A & Johnstone, T), Kindle version, Oxford University Press, Oxford, chapter 5, accessed 8 July 2012 from www.amazon.com , doi: 2224781-4923254

——— 2004, Feelings Integrate the Central Representation of Appraisal-driven Response Organization in Emotions, in *Feelings and Emotions*, (eds, Manstead, ASR, Frijda, N &

Fischer, A), Cambridge University Press, Cambridge, pp. 136-157

Schwartz, N & Vaughan, LA 2002, The Availability Heuristic Revisited: Ease of Recall and Content of Recall as Distinct Sources of Information, in *Heuristics and Biases: The Psychology of Intuitive Judgment*, (eds, Gilovich, T, Griffin, D & Kahneman, D), Cambridge University Press, Cambridge,

Schwitzgebel, E 2010, *Belief*, Stanford Encyclopedia of Philosophy, Winter 2010 edn, http://plato.stanford.edu/archives/win2010/entries/belief, Accessed 29 May 2011

Sedikides, C 1994, Incongruent effects of sad mood on self- conception valence: it's a matter of time, *European Journal of Social Psychology*, Vol. 24**,** No. 1, pp. 161-172, accessed 8 May 2014 from Wiley Online

——— 1995, Central and Peripheral Self-Conceptions Are Differentially Influenced by Mood: Tests of the Differential Sensitivity Hypothesis, *Journal of Personality and Social Psychology*, Vol. 69**,** No. 4, pp. 759-777

Sedikides, C & Spencer, SJ (Eds.) 2007, *The Self*, Kindle version, Taylor & Francis e-Library 2011, Psychology Press, New York, accessed 8 May 2014 from www.amazon.com

Sherman, DK & Cohen, GL 2006, The psychology of self-defense: Self-affirmation theory, *Advances in Experimental Social Psychology*, Vol. 38**,** pp. 183-242, accessed 10 May 2014 from Elsevier ScienceDirect, doi:10.1016/S0065-2601(06)38004-5

Sherman, DK & Hartson, KA 2011, Reconciling self-protection with self-improvement: Self-affirmation theory, in *Handbook of self-enhancement and self-protection*, (eds, Alicke, MD & Sedikides, C), Adobe Digital Editions version, The Guildford Press, New York, pp. 128-151 accessed 5 May 2014

Sherman, SJ, Cialdini, RB, Schwartzman, DF & Reynolds, KD 2002, Imagining Can Heighten or Lower the Perceived Likelihood of Contracting a Disease: The Mediating Effect of Ease of Imagery, in *Heuristics and Biases: The Psychology of Intuitive Judgment*, (eds, Gilovich, T, Griffin, D & Kahneman, D), Cambridge University Press, Cambridge, pp. 98-102

Simon, HA 1967, Motivational and Emotional Controls of Cognition, *Psychological Review*, Vol. 74**,** No. 1, pp. 29-39

Simons, DJ & Chabris, CF 1999, Gorillas in our midst: sustained inattentional blindness for dynamic events, *Perception*, Vol. 28**,** pp. 1059-1074

Simons, DJ & Levin, DT 1998, Failure to Detect Changes to People During a Real-world Interaction, *Psychonomic Bulletin & Review*, 5 (4)**,** Issue 4, 644-649, http://armannd.com/wp-content/uploads/2007/06/sl_1998.pdf, Accessed 9 October 2011

Slaby, J 2008, Affective intentionality and the feeling body, *Phenomenology and the Cognitive Sciences*, Vol. 7**,** pp. 429-444, accessed 26 April 2014 from Springer Standard Collection, doi:10.1007/s11097-007-9083-x

Sloman, SA 2002, Two Systems of Reasoning, in *Heuristics and Biases: The Psychology of Intuitive Judgment*, (eds, Gilovich, T, Griffin, D & Kahneman, D), Cambridge University Press,

Cambridge, NY, pp. 379-396

Slovic, P, Finucane, M, Peters, E & MacGregor, D 2002, The Affest Heuristic, in *Heuristics and Biases: The Psychology of Intuitive Judgment*, (eds, Gilovich, T, Griffin, D & Kahneman, D), Cambridge University Press, Cambridge, NY, pp. 397-420

Smart, JJC 1959, Sensations and Brain Processes, *The Philosophical Revew*, Vol. 68**,** No. 2, pp. 141-156

Solomon, A & Haaga, DAF 2003, Reconsideration of Self-Complexity as a Buffer Against Depression, *Cognitive Therapy and Research*, Vol. 27**,** No. 5, pp. 579-591

Solomon, RC 2004, On the Passivity of the Passions, in *Feelings and Emotions*, (eds, Manstead, ASR, Frijda, N & Fischer, A), Cambridge University Press, Cambridge, pp. 11-29

——— 2008, The Philosophy of Emotions, in *Handbook of Emotions*, (eds, Lewis, M, Haviland-Jones, JM & Barrett, LF), 3rd, The Guilford Press, New York, pp. 3-16

Stephens, GL & Graham, G 2004, Reconceiving Delusion, *International Review of Psychiatry*, Vol. 16**,** No. 3, pp. 236-241

Stone, J 1999, What Exactly Have I Done? The Role of Self-Attribute Accessibility in Dissonance, in *Cogntive Dissonance: Progress on a pivotal theory in social psychology*, (eds, Harmon-Jones, E & Mills, J), American Psychological Association, Washington, DC, pp. 175-200

Stone, J & Cooper, J 2001, A Self-Standards Model of cognitive dissonance, *Journal of Experimental Social Psychology*, Vol. 37**,** pp. 228-243, accessed 11 May 2014 from Elsevier ScienceDirect, doi:10.1006/jesp.2000.1446

——— 2003, The effect of self-attribute relevance on how self-esteem moderates attitude change in dissonance processes, *Journal of Experimental Social Psychology*, Vol. 39**,** pp. 508-515, accessed 11 May 2014 from Elsevier ScienceDirect, doi:10.1016/S0022-1031(03)00018-0

Storbeck, J & Clore, GL 2008, The Affective Regulation of Cognitive Priming, *Emotion*, Vol. 8**,** No. 2, pp. 208-214, doi:10.1037/1528-3542.8.2.208

Styles, EA 2005, *Attention, Perception and Memory: An Integrated Introduction*, Psychology Press, Hove

Sutherland, S 1992/2007, *Irrationality*, Pinter & Martin Ltd, London

Taleb, NN 2004, *Fooled by Randomness*, Penguin Books, London

——— 2007, *The Black Swan: The Impact of the Highly Improbable*, Kindle edn, Penguin Books, London

Tan, LH *et al.* 2008, Language affects patterns of brain activation associated with perceptual decision, *PNAS*, Vol. 105**,** No. 10, pp. 4004-4009

Thibodeau, R & Aronson, E 1992, Taking a closer look: Reasserting the role of the self-concept in dissonance theory, *Personality and Social Psychology Bulletin*, Vol. 18**,** No. 5, pp. 591-602, accessed 8 May 2014 from Sage Journals, doi:10.1177/0146167292185010

Thompson, PF 2007, Submission to upgrade to PhD

——— 2005, Mental Imagery: Sentences or Pictures? Propositional versus iconic representation

Tibboel, H, Van Bockstaele, B & De Houwer, J 2011, Is the emotional modulation of the attentional blink driven by response bias?, *Cognition and Emotion*, Vol. 25**,** No. 7, pp. 1176-1183

Trope, Y, Ferguson, M & Raghunathan, R 2001, Mood as a resource in processing self-relevant information, in *Handbook of affect and social cognition*, (ed, Forgas, JP), Taylor & Francis e-Library 2009, Lawrence Erlbaum Associates, Inc., Mahwah, NJ, pp. 258-276, chapter 12, accessed 24 April 2014 from www.amazon.com

Tversky, A & Kahneman, D 1974, Judgment under Uncertainty: Heuristics and Biases, *Science*, Vol. 185**,** No. 4157, pp. 1124-1131

——— 1982, Judgment under uncertainty: Heuristics and biases, in *Judgment under uncertainty: Heuristics and biases*, (eds, Kahneman, D, Slovic, P & Tversky, A), Cambridge University Press, Cambridge, pp. 3-20

——— 2002, Extensional versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment, in *Heuristics and Biases: The Psychology of Intuitive Judgment*, (eds, Gilovich, T, Griffin, D & Kahneman, D), Cambridge University Press, Cambridge, pp. 19-48

Uttl, B, Ohta, N & Siegenthaler, AL (Eds.) 2006a, *Memory and Emotion*, Blackwell Publishing, Malden, MA

Uttl, B, Siegenthaler, AL & Ohta, N 2006b, Memory and Emotion from Interdisciplinary Perspectives, in *Memory and Emotion*, (eds, Uttl, B, Ohta, N & Siegenthaler, AL), Blackwell Publishing, Malden, MA, pp. 1-12

Vitevitch, MS 2003, Change Deafness: The Inability to Detect Changes Between Two Voices, *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 29**,** No. 2, pp. 333-342

Vuilleumier, P & Huang, Y-M 2009, Emotional Attention : Uncovering the Mechanisms of Affective Biases in Perception, *Current Directions in Psychological Science*, Vol. 18**,** pp. 148-152, doi:10.1111/j.1467-8721.2009.01626.x

Vuilleumier, P & Schwartz, S 2001, Beware and be aware: Capture of spatial attention by fear-related stimuli in neglect, *NeuroReport*, Vol. 12**,** No. 6, pp. 1119-1122

Westbury, C & Dennett, DC 2000, Mining the Past to Construct the Future: Memory and Belief as Forms of Knowledge, in *Memory, Brain and Belief*, (eds, Schacter, DL & Scarry, E), Harvard University Press, Cambridge, Massachusetts, pp. 11-32

Whalen, PJ 1998, Fear, Vigilance, and Ambiguity: Initial Neuroimaging Studies of the Human Amygdala, *Current Directions in Psychological Science*, Vol. 7**,** No. 6, pp. 177-188

Wheatley, T 2009, Everyday Confabulation, in *Confabulation*, (ed, Hirstein, W), Oxford University Press, Oxford, pp. 203-221

Winston, JS & Dolan, RJ 2004, Feeling States in Emotion: Functional Imaging Evidence, in *Feelings and Emotions*, (eds, Manstead, ASR, Frijda, N & Fischer, A), Cambridge University Press, Cambridge, pp. 204-220

Wittmann, M 2009, The Inner Experience of Time, *Philosophical Transactions of the Royal Society: Biological Sciences*, Vol. 364**,** No. 1525, pp. 1955-1967

Wolpert, DM & Flanagan, JR 2001, Motor prediction, *Current Biology*, Vol. 11**,** No. 18, pp. R729-R732, accessed 25 April 2014 from Elsevier ScienceDirect

Wolpert, DM, Ghahramani, Z & Flanagan, JR 2001, Perspectives and problems in motor learning, *TRENDS in Cognitive Science*, Vol. 5**,** No. 11, pp. 487-494, accessed 25 April 2014 from Elsevier ScienceDirect

Zajonc, RB 1980, Feeling and Thinking: Preferences Need No Inferences, *American Psychologist*, Vol. 35**,** No. 2, pp. 151-175

——— 1984, On the Primacy of Affect, *American Psychologist*, Vol. 39**,** No. 2, pp. 117-123