1 Mental Mechanisms

Emotional Cognition

Emotional cognition is thinking that is influenced by emotional factors such as particular emotions, moods, or motivations. Here are some examples of situations where people are strongly influenced by their emotions, for good or bad:

A jury member reviews the evidence that an accused man is guilty of murder, but discounts much of it because he seems like a nice guy.
A scientist is told by advisors that her planned research is pointless, but she pursues it because she finds it exciting, and goes on to win a Nobel prize.
An entrepreneur decides to start a novel kind of business because his gut instinct tells him that it will be highly successful, and it is.
A voter wavers over what political candidate to choose, then opts for one with enormous charisma.
A person overcome with anxiety and guilt because of family disasters finds comfort in becoming a born-again Christian.

The primary descriptive aim of this book is to increase understanding of how emotional cognition works in such situations. The secondary normative aim is to suggest ways of improving thinking by appreciating the difference between cases where emotions foster good decisions and those where emotions get in the way. Reasoning is a particular kind of thinking in which decisions are made or beliefs are acquired as the result of the comparative evaluation of different options with respect to various kinds of evidence. Contrary to standard philosophical assumptions, reasoning is often an emotional process, and improving it requires identifying and assessing the impact of emotions.

Accomplishing both the descriptive and normative aims requires an appreciation of the mental mechanisms that underlie mental cognition. Part I of

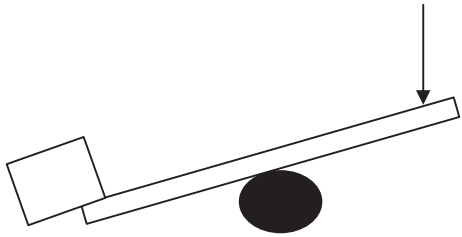

Figure 1.1
A simple machine, the lever.

this book describes some of the cognitive, social, neural, and molecular mechanisms that are crucial for emotional thinking. These mechanisms provide explanations of how emotions often influence people's decisions and other inferences. Part II shows how such mechanisms operate in emotional reasoning in law, science, and religion, and discusses the difference between desirable and undesirable kinds of emotional thinking. Before getting into details, however, it is necessary to say what mechanisms are and how they contribute to explanation.

Machines and Mechanisms

Humans invented simple machines such as levers and wedges long before the beginning of recorded history. Consider the basic lever shown in figure 1.1. It consists of only two parts, a stick and a rock. But levers are very powerful and enabled people to build huge structures such as the Egyptian pyramids. In general, a machine is an assemblage of components that transmit force, motion, and energy to each other to accomplish some task. To describe a machine and explain its operation, we need to specify its components, their properties, and their relations to other parts. Most important, we need to describe how changes to the properties and relations of the parts with respect to force, motion, and energy enable the machine to accomplish its tasks. The lever in figure 1.1 operates by virtue of the fact that the stick is rigid and is on top of the rock, which is solid. Applying force to the top of the stick makes the bottom of the stick move and lift the block, thus accomplishing the machine's task.

Ancient philosophers such as Epicurus and Lucretius realized that natural events can be explained in the same way that people explain the operations of machines. They postulated that all matter consists of atoms whose interactions determine the behavior of the objects that they constitute. According to

Lucretius (1969), even human thought is the result of the interactions of atoms. Natural phenomena are thus explained by *mechanisms*, which are like machines in that their changes result from forces and motions applied to their parts. Unlike machines, however, a natural mechanism is not constructed by humans and has no human-contrived task, although it may have a biological function. Nevertheless, the form of explanation is the same for both natural and constructed mechanisms: specify components, their properties and relations, and describe how changes in force, motion, and energy propagate through the system.

Mechanistic explanation has been fabulously successful in modern science, with triumphs such as Newton's theory of motion, the atomic theory of matter, the germ theory of disease, and the biological theories of evolution and genetics. But many philosophers, from Plato to Descartes, Leibniz, and even some modern thinkers, have resisted the extension of mechanistic explanation to mental phenomena. Plato's forms, Descartes's soul, Leibniz's monads, and twentieth-century phenomenology and hermeneutics are all attempts to understand the human mind in terms other than its material parts. Granted, early mechanistic attempts to explain mental operations were unsatisfactory, as Lucretius' atoms, Hume's associations of ideas, and modern behaviorist stimulus–response theories were much too crude to explain the richness of human thought. But I will now review a range of powerful mechanisms that have become available in the past fifty years for explaining thought in general and emotional reasoning in particular.

Kinds of Mental Mechanisms

Current cognitive science explains human thinking using a confluence of cognitive, neural, molecular, and social mechanisms. Those mechanisms most familiar are cognitive ones, which describe the mind as operating by the application of computational procedures on mental representations (for a survey, see Thagard 2005). Mental representations are cognitive structures, such as concepts, rules, and images, that are manipulated by algorithmic processes such as spreading activation of concepts, chaining of rules, and rotation of images. Connectionist explanations of mental phenomena describe them as arising from representations constructed by the activity of simple neuronlike units and processes that include spreading activation and modification of links between units.

Neural mechanisms are much closer to the operation of the brain in two key respects. First, they involve artificial neurons that are much more neurologically realistic than the highly simplified units found in connectionist

Table 1.1
Constituents of mental mechanisms

Mechanisms	Components	Relations	Interactions	Changes
Social	Persons and social groups	Association, membership	Communication	Influence, group decisions
Cognitive	Mental representations such as concepts	Constituents, associations, implication	Computational processes	Inferences
Neural	Neurons, neural groups	Synaptic connections	Excitation, inhibition	Brain activity
Molecular	Molecules such as neurotransmitters and proteins	Constituents, physical connection	Biochemical reactions	Transformation of molecules

explanations. For example, real neurons have temporal properties such as firing in coordination with each other that are crucial to their ability to support complex mental activity. Second, whereas connectionist models usually employ small numbers of interconnected units, brain mechanisms involve billions of neurons organized into functional areas such as the hippocampus and various parts of the cortex. Neural mechanisms depend on molecular mechanisms, such as the movement of ions within neurons, the flow of neurotransmitters in the synaptic connections between them, and the circulation of hormones through the brain's blood supply.

Finally, because human thought often involves interaction with other people, we need to attend to the social mechanisms that allow one person's thinking to influence another's. Social mechanisms involve verbal and other kinds of communication, including ones that make possible the transfer of emotions as well as other kinds of information. Table 1.1 summarizes some of the ingredients of the four kinds of mechanisms useful for different levels of mental explanation. Chapters 2 through 7 provide much more detailed examples of the mechanisms at each level that are relevant for understanding emotional thought.

The Nature of Complex Mechanisms

My approach to emotions is in keeping with the mechanism-based view of explanation espoused by such philosophers of science as Salmon (1984b), Bechtel and Richardson (1993), and Bechtel and Abrahamsen (2005). Machamer, Darden, and Craver (2000, p. 3) characterize mechanisms as

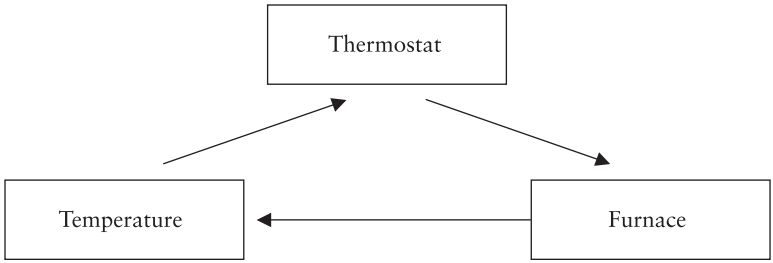

Figure 1.2
Feedback mechanism that regulates temperature.

"entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions." I prefer the term "component" to "entity" because it indicates that the objects in a mechanism are part of a larger system; and I prefer "interaction" and "change" to "activity" because they sound less anthropomorphic. More important, I find the reference to start and finish conditions misleading, because the mechanisms needed to explain emotional thinking involve ongoing feedback processes rather than unidirectional changes.

For a simple example of feedback, consider the system shown in figure 1.2. A thermostat is a device that detects when the temperature drops below a given threshold; it then signals a furnace to start up and heat the air. Taken separately, the thermostat and the furnace are unidirectional machines describable by simple rules: if the temperature is less than *X*, then signal the furnace to start; if a signal to start is received, then start the furnace. Taken together, however, they constitute a feedback mechanism that functions to maintain the temperature within a stable range. There are no start and finish conditions, but rather an ongoing process of temperature regulation. The human body contains many such feedback mechanisms, for example, ones that maintain blood pressure and cholesterol levels. As we will see in the next four chapters, those mechanisms most relevant to explaining emotions—cognitive, neural, molecular, and social—are feedback mechanisms rather than unidirectional ones with start and finish conditions.

Machamer, Darden, and Craver (2000) observe that mechanisms occur as nested hierarchies and that their description is frequently multilevel. They argue that higher-level entities and activities are essential to the intelligibility of those at lower levels and vice versa, and that mental phenomena are best explained by the integration of different levels. I fully agree, and will describe in detail how an understanding of emotional thinking needs to integrate

the cognitive, neural, molecular, and social levels. This integration is much more complex than simply reducing the social to the cognitive, the cognitive to the neural, and the neural to the molecular. Some emotional events require an explanation that involves all levels simultaneously, for example, when a social interaction between two people generates fear that increases cortisol (stress hormone) levels that then change the social interaction. An insult leads to cognitive and physiological changes that incite a fight. Thus causality runs down as well as up the nested hierarchy of social–cognitive–neural–molecular mechanisms.

The nested hierarchy of mechanisms is a hierarchy of components as well as interactive changes. A social group consists of people who have brains that contain neurons made up of molecules; changes in the social group produce and are produced by changes in individual brains that arise from neural and molecular changes. Hence I reject single-level explanatory strategies, both reductive ones that privilege the lowest-level strategies, and intermediate ones that insist that only one level such as the cognitive one are appropriate for psychological explanation. Like McCauley and Bechtel (2001), I favor *explanatory pluralism*, which rejects both extreme reductionism and antireductionism. I hope to show that pluralistic, multilevel explanations based on multiple mechanisms are the best way to understand emotional thinking. It is impossible to argue for this position a priori, but the chapters to follow will provide ample experimental and theoretical support for a multilevel approach to explaining emotional thinking. (See chap. 16 for further discussion.)

At each level of the hierarchy, explanation of aspects of emotional thought will specify mechanisms consisting of components and their interactive changes. These aspects can include both positive cases of emotional thinking such as the scientist guided by excitement toward the Nobel prize, and negative cases such as the voter swayed by charisma rather than political substance. Just as knowing how machines work enables us both to understand and to fix them, so multilevel mechanisms will suggest a hierarchy of ways to explain and to improve emotional thinking.

Many of the chapters in this book employ computational models of various levels of thinking. A model is a description of a mechanism, that is, a specification of the components, relations, interactions, and changes that can be used to explain phenomena. Cognitive science routinely uses computational models for two reasons. The first reason, as seen in areas as diverse as physics and economics, is that computers provide a powerful way of describing complex interactions and discovering their consequences. The second reason, more specific to cognitive science, is that thinking itself is hypothe-

sized to be a computational process, and that the brain is a special kind of computer. For a defense of this perspective, see Thagard 2005.

My main purpose in discussing computational mechanisms such as those described in part I is to provide explanations of important aspects of human thinking. Mechanisms do not provide explanations by describing what always happens or even what probably happens. Rather, they explain by describing the systematic interactions of the key factors that cause something to happen. In computational mechanisms, these factors include mental representations at the cognitive level and neurons at the neural level.

Guide to the Book

The rest of part I of this book describes mechanisms at different levels of analysis that are responsible for the various ways in which emotions affect thought. Chapters 2 through 4 describe the cognitive level and use a connectionist model of emotional coherence to illuminate the nature of decision making and other kinds of emotional reasoning. Chapter 2 provides a gentle introduction to decision making from the perspective of emotional coherence, and chapter 3 applies this perspective to the role of emotions in analogical thinking. Chapter 4 considers emotional thinking in terms of dynamic systems.

Chapter 5 discusses social mechanisms of emotional communication that contribute to the achievement of consensus in group decision making. Like individual decisions, group decisions are often highly emotional, and the extended HOTCO model provides computer simulations of how consensus can arise even in groups that initially diverge in their emotional reactions to different options. This chapter shows how a social model of consensus can be built on top of a cognitive model of emotional decisions.

The HOTCO models described in chapters 2 through 5 are not neurologically realistic: they employ artificial neurons that represent whole concepts and propositions; and these neurons are not organized into functional brain areas. That is why I have categorized them as employing cognitive rather than neural mechanisms. In contrast, the GAGE model presented in chapter 5 is much more neurologically realistic in that it uses spiking neurons, distributed representations, and distinct brain areas. This realism enables the model to account for the behavior of people with a particular kind of brain damage, as well as for how emotion and cognition can affect each other generally via neural mechanisms.

Chapter 6 moves down another notch to the molecular level, arguing that a full theory of thinking, particularly emotional thinking, will need to take

into account mental computations accomplished by molecular mechanisms such as chemical reactions that operate within and between neurons and other cells.

The rest of the book is concerned with applying knowledge of emotional mechanisms to illuminate thinking in the domains of law, science, and religion. Chapter 8 provides an explanation of why the jury in the O. J. Simpson trial did not convict him, arguing that their decision was based on a combination of explanatory and emotional coherence. More generally, chapter 9 discusses the nature of doubt and reasonable doubt in legal and other contexts. Chapters 10 through 12 are concerned with both the positive and negative roles of emotions in scientific thinking. Chapter 10 provides a general overview of the ways in which emotions affect the pursuit, discovery, and evaluation of scientific ideas, and chapter 11 looks at how they contributed to the development of research in a particular case. Chapter 12 provides a distillation of advice on how to be a successful scientist, including valuable emotional habits. Chapters 13 and 14 discuss the application of emotional coherence to religious thinking, the former by looking at a particular case of the self-deception of a minister, the latter considering more generally the emotional content of religious beliefs.

Implicit in all these discussions of the applications of emotional reason is a set of judgments about appropriate standards for when the incursion of emotional elements into cognitive deliberation is legitimate. Chapter 15 begins the discussion of this normative question, which is continued in the concluding chapter 16. I reject the traditional assumption that emotion is inimical to good reason, but also argue against the romantic view that emotion always contributes to good thinking. Chapter 16 also sketches a number of other areas of ongoing investigation of hot thought, including its relevance to conflicts of interest, generation of explanatory hypotheses, and neuroeconomics.

We shall see that many aspects of human thinking and reasoning are influenced by emotions. An understanding of how emotions affect thought both positively and negatively requires the specification of cognitive, social, neural, and molecular mechanisms. Such mechanisms consist of components that undergo interactive changes that may involve feedback processes. Our best strategy for explaining emotional thinking is to look at multiple integrated levels of causal mechanisms.