11 Feelings: Chocolate, Lust, and Coherence

Nature has placed mankind under the governance of two sovereign Masters, pain and pleasure.

—Jeremy Bentham

The postulate that affective processes have an objective existence and that they intervene between stimulus and response has great utility. . . . Experimental findings practically demand such a hypothesis.

—Paul Thomas Young

Almost every feeling of physical pleasure or pain felt by your forebrain has climbed its way there through the brain stem.

—Kent Berridge

At this juncture in our story, we are approaching a dilemma. The model of mind that we have been developing to explain organismic behavior postulates multiple qualitatively distinct systems based on different principles and mechanisms. The autonomic mind is based on homeostasis and reflex arcs. The instinctive mind is explained as consisting of fixed action patterns triggered by (metaphorical) action-specific energy reservoirs. The associative mind is about learning through reinforcement. The reasoning mind is about coherence relations between propositional attitudes. Despite multiple systems, behavior consists of a single blended response. This implies a global integration function that takes input from each system and determines the blended response. The conundrum is to explain how there can be integration across the levels in the absence of a common language. The solution to this problem requires a bold, speculative conjecture: what is common to each system are *feelings*. Feelings provide the common currency allowing for communication across levels and for calculation of the overall response to a given situation.

This is a pivotal chapter in our story. I want to make the controversial case for the central role of feelings in the behavior of all mammals, perhaps even all vertebrates. I will characterize feelings, differentiate them from emotions, and argue that feelings are generated in old brain stem, diencephalon, and subcortical systems that have been widely conserved across large parts of the phylogenetic tree. This means that they are available to each of our four minds. I will propose that feelings are the solution evolution has come up with to solve the two critical problems of *selecting* behavior and *initiating* behavior. Feelings serve these critical roles in the internal operations of each of the four different kinds of minds we have been discussing and provide the common currency for the integration of a single behavioral response.

Who's Afraid of Feelings?

Feelings span the range of sensations associated with the kink in my neck, the warmth of sunshine on my face, the pain in my knee, my heart racing after sprinting, hunger pangs, bowel distention, the taste of chocolate cake, sexual arousal, the desire to see a loved one, fear of lions, pangs of jealousy, anger and remorse, and my very sense of being. Feelings are undoubtedly the most vexing and shunned topic in psychology and neuroscience.

The problem with feelings is that they are *feelings*. They have a first-person ontology, meaning they are inherently subjective. I'm certain of the feelings associated with the kink in my neck, the taste of chocolate cake, and my fear of heights through my direct first-person experience and only through my direct first-person experience. If this is the case, how can feelings be studied objectively? Modern empirical psychology began in the 1870s as an exploration of conscious feelings. Wilhelm Wundt, the founder of the first empirical psychology lab, is attributed to have said that "when we study a living system from the outside, we are doing physiology; when we study it from the inside, we are doing psychology." Since the only method for directly accessing feelings is through introspection, and introspection does not generally permit repetition and verification of experiments, the hallmarks of the scientific method, this initiative was short-lived and was soon overwhelmed by behavioral psychology.[1]

Since then, many serious, accomplished psychologists and neuroscientists have argued that there is no independent, objective evidence for the reality and causal efficacy of feelings. They note that what can be observed and measured are behaviors and their underlying physiology; let us stick to these measures in our theorizing (LeDoux, 2012; Skinner, 1953). In fact, we have learned a great deal about living organisms through studies of behaviors and physiology, without any appeal to feelings. If feelings do exist, perhaps they are epiphenomenal. That is, they may be like the sound generated by hammering a nail into a piece of wood. It is a natural consequence of metal striking metal but has no part to play in the causal story of driving the nail into the wood and thus has no part in scientific explanation.

Indeed, physics has made enormous progress by redefining intuitive concepts to eliminate feelings. Consider the notion of heat. Heat was initially defined as the feeling of the sunshine on your face or the feeling of putting your hand in a pot of boiling water. This did not turn out to be particularly useful for understanding the world. As physics progressed, it redefined heat as the transfer of mean kinetic energy of the object's component particles. This redefinition carved off the subjective feeling component and provided an objective, measurable concept that actually deepened our understanding of the world. Skeptical psychologists and neuroscientists have taken a similar approach to feelings. They may be real, but they are not causally relevant in explaining the world and therefore not a subject matter in and of themselves. What needs to be studied is the accompanying behavior. The danger with this position is that if feelings turn out to be essential to the subject matter of psychology and neuroscience, carving them away also means carving away our subject matter (Searle, 1992). To confront this issue is to confront the problem of consciousness. The philosophical problem of consciousness is beyond my paygrade. Nonetheless, feelings are central to my story. By feelings I'm referring to the pain sensation of an electric shock resulting in withdrawal and avoidance behavior (both in myself and in the rat) and the pleasurable taste of chocolate cake resulting in consumption behavior (again, both in myself and in the rat).

Fortunately, a number of psychologists and neuroscientists are gravitating away from this skepticism and are willing to acknowledge feelings both in their informal conversation and formal theory development (Berridge & Kringelbach, 2013, 2015; Bindra, 1974; Craig, 2009; Damasio & Carvalho, 2013; Leknes & Tracey, 2008; Panksepp, 2011; Young, 1959). The main reasons for this radical shift are a combination of behavioral and neurophysiological data and some basic tenets of the theory of evolution. These scientists are beginning to accept that one can explain more data more coherently by *positing* feelings rather than ignoring them. However, sharp internal divisions remain as to the source and nature of feelings (Barrett, 2006; LeDoux, 2012; Lindquist, Wager, Kober, Bliss-Moreau, & Barrett, 2012; Panksepp, 2007). I agree that feelings are not just phenomenologically real but also causally efficacious in behavior. In fact, they may be the primal essential feature of all mammalian, perhaps even all vertebrate, life. The role that they play may be that of initiating and guiding behavior.

Characterizing Feelings

Feelings are sensations. Pioneering neuroscientist Charles Sherrington (1952) grouped sensations into the following five categories: teloreceptive (vision and hearing), proprioceptive (limb position), exteroceptive (touch, temperature, and pain), chemoreceptive (smell and taste), and interoceptive (visceral). More recently, some neuroscientists have been suggesting reorganization of the interoceptive category to include all aspects of the physiological condition of the body, not just the viscera (Craig, 2002). For our purposes, it may be adequate to group feelings into two broad categories: interoceptive, referring to all internally generated homeostatic and visceral bodily sensations,[2] and exteroceptive, referring to the sensations emanating from the impact of the external environment on the five senses. It may also be useful to recognize, as a third category, the "feeling of effort," or what John Searle (1983) called "intention in action," associated with volitional action.

Feelings are typically characterized along two orthogonal dimensions: valence and arousal. Valence constitutes a scale along a positive/negative or pleasant/unpleasant dimension. For many, eating chocolate cake is at the positive end of the scale and eating bitter melon is at the negative end of the scale. The second dimension is arousal. It ranges from high to low intensity. Sexual orgasm is at the high end of the intensity scale and the positive end of the valence scale. Experiencing an electric shock is also at the high end of the intensity scale but at the negative end of the valence scale. Feeling depressed implies both low arousal and negative valence. Feeling calm and content implies low arousal and positive valence. Some researchers argue that each and every feeling, whether it be the kink in my neck or the feeling of disappointment in the peer reviews of my last manuscript, can be captured along the two dimensions of valence and arousal, but others disagree (Colibazzi et al., 2010; Panksepp, 2007; Panksepp & Biven, 2012; Russell, 2003).[3] Either way, these two dimensions are a useful organizing tool. Feelings also have a third dimension, temporal duration, which is often ignored. Feelings have a beginning and an end. They can wax and wane. The pleasurable taste of chocolate cake commences with the first bite (or perhaps even first aroma or sight) and subsides postconsumption.

From Feelings to Emotions

In addition to these basic feelings associated with internal bodily regulation and exterior sensory inputs, there is another set of feelings or affective states that all humans are familiar with: *emotions*.[4] Our emotional states encompass feelings, but feelings are not emotions. Examples of common human emotions are fear, rage, disgust, hope, and jealousy. Like interoceptive and exteroceptive feelings, emotions have valence, physiological arousal, and duration components, but unlike interoceptive and exteroceptive feelings, they also have intentional objects, action tendencies, physiological expressions, and cognitive antecedents associated with them (Elster, 1998).

To say that emotions have intentional objects is to say that emotional states are representational states. That is, they are directed at, refer to, or represent objects, individuals, and states of affairs in the world.[5] Hunger in itself is not a directed state; I can feel hungry without wanting to eat anything in particular. My craving for chocolate cake, though, is a directed state. Recall the discussion about the nature of contents of intentional states from chapter 6. I made a distinction between the directedness of the intentional states of a cat tracking a mouse and human intentional states that have sophisticated propositional contents. We need to maintain this distinction here to allow for emotional states in animals. My directedness toward an immediately present object, such as the slice of chocolate cake on my plate, may be similar to the type of directedness that a hungry dog has to that same slice of cake, but I am also capable of more sophisticated forms of directedness that require propositional content:

> I pushed my fork through the top layer of creamy frosting, then all three layers of the cake. Keeping my eyes down, I put the fork to my mouth. He'd used good chocolate, I knew, and after a moment, I picked up a note of coffee, which only intensified the flavor of the chocolate. The frosting was decadent and smooth, but not cloying. In fact, the entire bite struck the precise balance of sass and sweet. (Stuart, 2017)

Both types of mental states directed at the chocolate cake qualify as referential or directed mental states, but they are qualitatively different. The former type of directedness should be widely available on the evolutionary tree. The latter type requires mental states with propositional content, presumably only available to humans, and allows emotions to enter the reasoning mind.

Physiological expressions such as bodily posture, pitch of voice, blushing, smiling, baring of teeth, laughing, frowning, weeping, and crying are usually associated with emotions. There are various degrees of conscious control over these expressions. No one can blush on command, but some people can cry on cue. Actors can learn to imitate many of these facial expressions and body postures. It has been argued that specific facial expressions are associated with specific emotions across all human cultures and even in nonhuman animals, providing evidence for a basic set of emotions (Berridge & Kringelbach, 2015; Ekman, 1993).

Emotions are also associated with action tendencies. Action tendencies are not unlike the "fixed action patterns" discussed in the context of instinctual behaviors (chapter 4). Nico Frijda, quoted in Elster (1998, p. 51), described them as "states of readiness to execute a given kind of action. . . . Action tendencies have the character of urges or impulses." Such "urges and impulses" also appeared in Lorenz's instinct model as pent-up energy reservoirs. For example, fear may lead to fight or flight, lust may lead to actions to possess the object of sexual desire, shame may lead one to hide or disappear, guilt may lead to atonement or confession, envy and malice may lead one to destroy, love may lead to approaching and touching the other person, and anger may lead to hurting the person who has hurt you. But don't fixed action tendencies violate the "gap" that plays a critical role in our conception of rationality (chapter 6)? This need not be the case if no *specific* stimuli are necessary and sufficient to trigger a *specific* emotion and if no *specific* action patterns are associated with each emotion. This is illustrated in the example below from *King Lear*.

Many human emotional states are triggered by other intentional states, typically beliefs and desires. Why did Shakespeare's King Lear become angry with his youngest, favorite daughter, Cordelia? When asked to profess her love for him alongside her two sisters, Cordelia has no words to compete with the insincere flattery offered by Goneril and Regan and, when pressed, replies as follows:

> Good my lord,
> You have begot me, bred me, loved me: I
> Return those duties back as are right fit,
> Obey you, love you, and most honour you.
> Why have my sisters husbands, if they say
> They love you all? Haply, when I shall wed,
> That lord whose hand must take my plight shall carry
> Half my love with him, half my care and duty:
> Sure, I shall never marry like my sisters,
> To love my father all.

This is not enough for Lear. He desires more obsequious displays of her love. Failing to receive them, he comes to believe that she does not really love him, and being particularly wounded because she is his favorite, he flies into an angry rage and disowns her and divides his kingdom among her sisters:

> Let it be so; thy truth, then, be thy dower: . . .
> Here I disclaim all my paternal care,
> Propinquity and property of blood,
> . . . thou my sometime daughter.

This illustrates not only the triggering of the human emotion (anger) by beliefs and desires but also the triggering of accompanying action tendencies. Notice that there is nothing that *compels* Lear to be angered by any *specific* antecedent belief and desire. However, once the anger is triggered, the action tendency unfolds. It is not, however, stereotyped as in the case of instincts. There are numerous actions, ranging from a verbal expression of disappointment to execution, for Lear to express his anger. The rationality gap is intact. The duration component of emotions is also illustrated when Lear belatedly regrets his anger and actions:

> O Lear, Lear, Lear!
> Beat at this gate that let thy folly in
> And thy dear judgement out!

This characterization of emotions in terms of intentional or directed states, associated with action tendencies, physiological expressions, and the fact that they are usually triggered by other intentional states (in humans), differentiates them from sensations associated with interoceptive and exteroceptive systems and volitional motor actions. The status of emotional states is a highly contentious and debated issue in the literature. Some researchers believe that emotions are high-level cognitive constructs computed or inferred from core interoceptive biofeedback signals by neocortical structures (Barrett, 2006; Barrett, Quigley, & Hamilton, 2016; LeDoux & Brown, 2017; Lindquist et al., 2012; Ortony, Clore, & Collins, 1988; Seth, 2013). As such, they are available only to humans and perhaps some other primates. Other researchers believe that some primal emotions are internally generated in specific deep brain stem, diencephalon, and subcortical regions of mammalian brains (perhaps even vertebrate brains generally) and thus are a common heritage of large parts of the evolutionary tree (Berridge & Kringelbach, 2015; Damasio & Carvalho, 2013; Ekman, 1993; Kringelbach & Berridge, 2009; Panksepp, 2007; Panksepp, Lane, Solms, & Smith, 2017; Toronchuk & Ellis, 2013). I believe that the bulk of the evidence supports the latter position, and I also adopt it. However, there is also a potential for confusion here that needs to be preempted. When talking about emotions in nonhuman animals, we are referring to emotions *without* propositional contents. I will shortly introduce some vocabulary to distinguish full-blown human emotions, which have propositional content and participate in the reasoning mind, from nonhuman emotions. We will also see that the latter have a characterization similar to instincts.

Origins of Feelings

You are sure that you have feelings. Based on the observation that I'm physically very much like you, and belong to the same species, you are probably prepared to accept that I have feelings. But what is the evidence that nonhuman animals also experience feelings? How can you know whether your dog feels pain or is capable of loving you?

The first-person ontology of feelings makes this a difficult question to answer, but, in reality, we can answer the question of feelings in nonhuman animals in the same way we answer it in fellow humans: through behavioral and anatomical/physiological observations. We have already noted some of the pitfalls of relying purely on behavioral observations (chapter 7). However, combining behavioral with anatomical/physiological observations and noting the basic engineering principle that structure is not unrelated to function allows us to make some headway. In biology, many functional homologies can be mapped onto structural homologies. Similarities in motor function or visual function across species are underwritten by similar neural architecture. If we can identify the source or generators of feelings in the brain, we can then see how widely these structures are available on the evolutionary tree.

Where feelings first appeared on the phylogenetic tree is relevant to understanding their potential functions. If feelings are generated in the neocortex, then only humans, and perhaps some other primates with well-developed neocortices, will have access to them. In that case, their role may be confined largely to cognitive systems. However, if feelings are generated in the brain stem, diencephalon, and subcortical structures—which appeared very early and have been conserved, certainly in mammals and perhaps in all vertebrates—feelings may be more widely available and have much more basic functions affecting many survival systems. There is evidence from electrical stimulation studies, decorticate (removal of cortex) studies, chemical stimulation studies, and conditioning studies to suggest that feelings are indeed generated in deep brain stem, diencephalon, and subcortical structures, though they can be represented in higher-level cortical structures for various purposes and combined into sophisticated, complex emotions, available only to humans (Berridge & Kringelbach, 2013; Damasio & Carvalho, 2013; Kringelbach & Berridge, 2009; Panksepp, 2007, 2011; Panksepp & Biven, 2012; Pfaff, Martin, & Faber, 2012; Venkatraman, Edlow, & Immordino-Yang, 2017).

The study of feeling systems in the brain began with a classic study from the 1950s where electrodes were placed in the septal region in the brains of rats (and other animals) and the animals were placed in a Skinner box with a lever that, when pressed, generated an electrical stimulation in the electrode (figure 11.1a). The rats soon learned to press the lever and then use it extensively to obtain continuous stimulation (Olds & Milner, 1954).[6] Not only would animals work for the stimulation, the stimulation could be used as a substitute for food as a reward in classical conditioning and operant conditioning experiments (Ross et al., 1965; White & Milner, 1992). As electrical stimulation serves no biological homeostatic function (unlike food), animals are expending effort to receive it presumably because they find it pleasurable or rewarding in itself (Olds & Milner, 1954). Deep electrode stimulation of the same septal brain region in humans (figure 11.1b) was found to be associated with sexual arousal in two patients (Heath, 1972). More recent studies note that these electrodes were placed very close to the nucleus accumbens and suggest that it, not the septal region, is the source of the rewarding arousal. There is also some current reconsideration as to whether the studies have activated pleasure centers or motivation centers (Berridge & Kringelbach, 2015). We will return to this important distinction.

Electric brain stimulation in the septal region elicits these rewarding feelings not only in intact rats but also in decorticate rats that have had their neocortex surgically removed (Panksepp, 2007). These animals still work to receive electrical stimulation, and they continue to engage in many pleasurable behaviors, such as play (Panksepp, Normansell, Cox, & Siviy, 1994) and sexual lordosis (Carter, Witt, Kolb, & Whishaw, 1982). Another strong source of evidence that the generators for these feelings are in the brain stem, diencephalon, and subcortical regions comes from clinical cases of children born without a neocortex. These children are still capable of conscious experience and emotional reactions (Merker, 2007; Shewmon, Holmes, & Byrne, 1999).

Rats are known to emit frequency-modulated 50 kHz calls during positively valenced appetitive behaviors such as sex and play. Electrical brain stimulation in deep brain structures such as the lateral preoptic area, lateral hypothalamus, and ventral tegmental area elicits the same 50 kHz vocalizations, suggesting that the animals find the stimulation equally pleasing

This is a portion of the eBook at doi:10.7551/mitpress/12811.001.0001

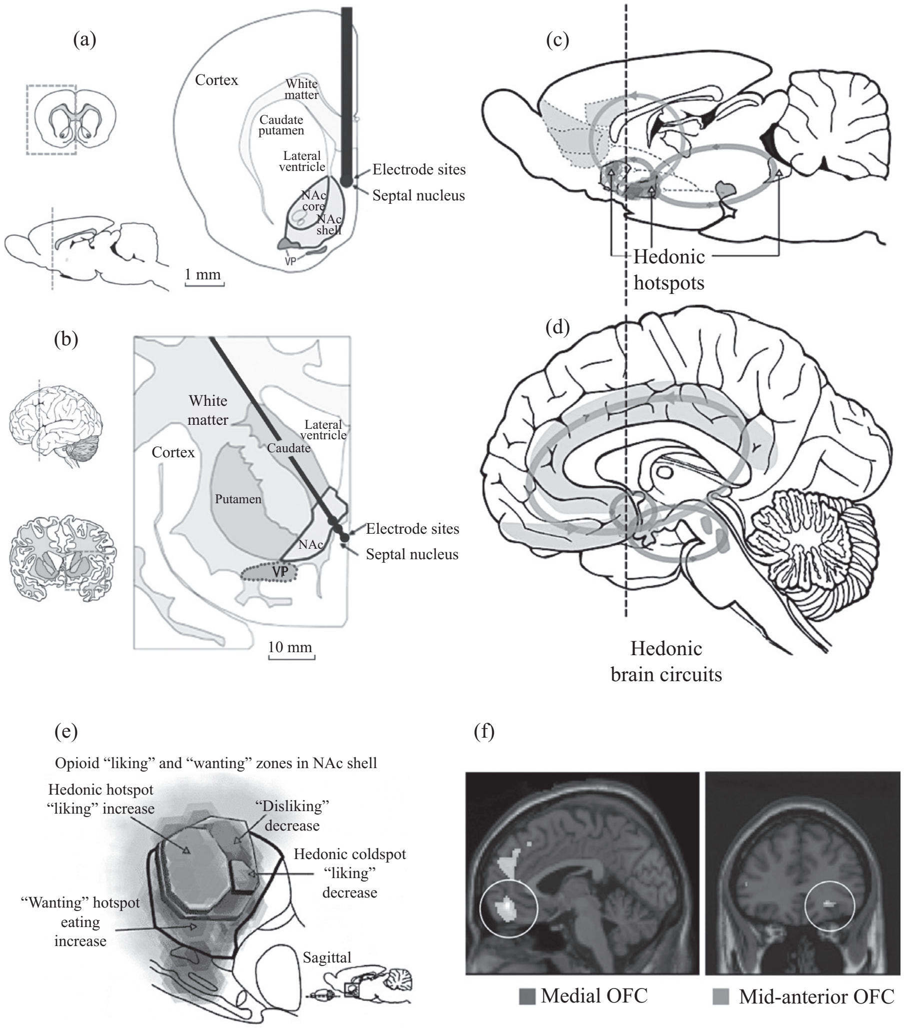

Figure 11.1
(a) The location of the hedonic "hotspot" discovered by Olds and Milner (1954) in the septal nucleus of rats has been reconstructed and found to be very close to the nucleus accumbens. (b) Placement of electrodes in or near the nucleus accumbens in one patient reported by Heath (1972) resulted in feelings of sexual arousal. (c), (d) Hedonic reward systems in both rodents and humans involve similar interlinked brain stem, diencephalon, and subcortical networks, with some cortical representation, particularly in humans. The systems involved include the periaqueductal gray (PAG), the ventral tegmental area (VTA), ventral pallidum, nucleus accumbens, amygdala, hypothalamus, insular cortex, cingulate cortex, and orbital frontal cortex. (e) Shows the identification of distinct "liking"/"disliking" and "wanting" hotspots in the nucleus accumbens of a rat brain. (f) Reward regions are represented in the human orbital frontal cortex. Figure reproduced (with some reorganization) from Kringelbach and Berridge (2010) with permission of the authors.

(Burgdorf, Wood, Kroes, Moskal, & Panksepp, 2007; Burgdorf et al., 2008, Burgdorf, Panksepp, & Moskal, 2011).

Opiate drugs such as morphine and heroin mediate sensory pleasure and positive social bonding systems in the brain (Panksepp, 1981; Panksepp & Biven, 2012). Like humans, animals will seek out and work for these drugs (Ikemoto, 2010), presumably for similar reasons: consuming them is pleasurable. When these drugs are infused directly into animal brains, the animals prefer morphine infusions into primitive brain stem regions such as the periaqueductal gray and the ventral tegmental area over infusions into other regions, even though these other regions also have abundant opiate receptors. This preference suggests that the preferred brain stem and subcortical regions for infusion may be the main source of the pleasurable affect generation associated with the drug (Olmstead & Franklin, 1997).

Finally, it is well known that animals develop preferences for places where they have had positive experiences such as food and sex (conditioned place preference) and avoid places where they have had negative experiences such as an electric shock or the odor of a predator (conditioned place aversion). They develop the same place preferences and aversions, respectively, to artificial electrical and chemical stimulation of the relevant brain systems, suggesting that they find the brain stimulations equally as pleasing as food and sex or as displeasing as an electric shock (Olmstead & Franklin, 1997; Panksepp & Biven, 2012; Pfaus et al., 2012).

Since these pioneering studies, neuroscientists have made considerable progress in identifying and mapping out reward (and aversion) systems in both nonhuman and human brains. There is considerable consensus that hedonic systems in human and nonhuman animals involve overlapping hierarchically organized neural nets in the periaqueductal gray (PAG), ventral tegmental area (VTA), ventral pallidum, nucleus accumbens, amygdala, hypothalamus, insular cortex, cingulate cortex, and orbital frontal cortex (figure 11.1c, d). Many different types of rewarding stimuli (food, sex, addictive drugs, even art and music) activate this same common system (Berridge & Kringelbach, 2015).

As noted in chapter 10, brain stem, diencephalon, and subcortical structures are highly conserved across mammals and even across vertebrates. The data reviewed in chapter 10 and the present chapter suggest that they serve similar essential functions across species. If one of these functions is the generation of feelings in humans, it is reasonable to hypothesize that homologous regions serve homologous functions in large parts of the phylogenetic tree.

This is a hypothesis. Do not bet the house on it. But it is a reasonably robust hypothesis; I would bet my car on the grounds that we can make sense of more behavioral and neurophysiological data with this hypothesis than without it. It is also important to note that these considerations do not speak to the question of whether your dog can actually love you. They only indicate that dogs have an affective life. The particulars of that life will undoubtedly be shaped by the evolutionary niche of each species.

But why should these feeling circuits be conserved? Nature is not prodigal. What critical role do feelings play in enhancing survival and reproduction (i.e., fitness)? Why should brains, which are very expensive to maintain—comprising 2% of body weight but consuming 25% of energy (in humans)—conserve structures and processes required to generate feelings? Why do we need feelings? The answer may lie in the fact that survival and propagation of organisms depends on the selection and initiation of appropriate behaviors or actions in response to environmental (internal and external) change. Feelings may be the solution that evolution has converged on to detect certain changes and select and initiate actions. This hypothesis is particularly robust if feelings are generated in phylogenetically old brain stem, diencephalon, and subcortical structures, as the data indicate.

Function of Feelings: Motivate and Guide

I propose that feelings evolved to allow organisms to detect changes in their environments and to select and initiate appropriate actions.[7] Consider the oral sensory system of taste discrimination as an illustration (figure 11.2). It begins with a chemical reaction activating taste buds, proceeds via sensory neurons to brain stem nuclei, to the thalamus, and then to the insular cortex (Matsumoto, 2013). The pathway is very similar for rodents and primates, except it bypasses the parabrachial nucleus in primates and humans. It allows for differentiation between the sweet buttery taste of chocolate cake and the salty taste of crackers and the sour taste of lemons.

Not only can I differentiate between different tastes, I also have different preferences for them. I like certain tastes more than others. I will go out and purchase and consume chocolate cake more frequently than lemons. The taste differentiation system alone cannot explain this. A notion of reward or pleasant affect is needed to account for preference. I prefer chocolate cake to lemons because it tastes better. We can *feel* this preference in ourselves and we can see it behaviorally in others. Similar facial liking and disgust reactions to sweet and bitter tastes can be elicited from children on

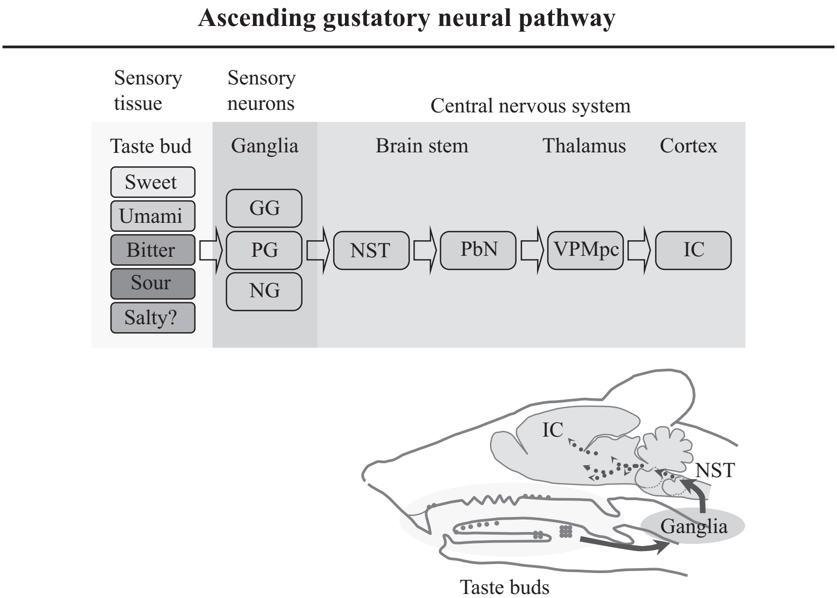

Figure 11.2
The rat's system for taste discrimination. Chemical reactions in the oral cavity between food and the taste buds are recognized as distinct tastes in the insular cortex after processing in the brain stem and thalamus. GG = geniculate ganglia; IC = insular cortex; NG = nodose ganglia; NST = nucleus of the solitary tract; PbN = parabrachial nucleus; PG = petrosal ganglia; VPMpc = ventral posterior medial nucleus of the thalamus. Figure reproduced with permission from Matsumoto (2013).

the first postnatal day and are homologous across humans, primates, and even rodents (Berridge & Kringelbach, 2015).

Neuroscientists Kent Berridge of the University of Michigan and Morton Kringelbach of Oxford University, along with many colleagues, have spent decades studying the neural basis of the brain's reward system. They propose that it can be subdivided into two distinct (but interrelated) components, the "wanting" system and the "liking" system. They characterize *wanting* as the incentive salience or motivational magnet component of reward. *Liking*, by contrast, is a hedonic reaction (i.e., feels pleasurable) and is detectable both behaviorally and in neural signals generated by subcortical brain structures (Berridge & Kringelbach, 2015).

At the neural level, taste sensations are generated by a small set of discrete "hotspots" (liking) and "coldspots" (disgust) located in the nucleus accumbens (figure 11.1e), ventral pallidum, parabrachial nucleus in the brain stem, and perhaps also in the orbital frontal cortex and insular cortex, at least in humans (figure 11.1f). Activation of hotspots for liking in the nucleus accumbens amplifies pleasurable reactions, while activation of the coldspots dampens pleasure and initiates disgust reactions. Lesions in the ventral pallidum result in loss of hedonic response to taste (Berridge & Kringelbach, 2015).

The liking or hedonic system seems to have two functions. The first function is selection. It is not an accident that omnivores such as humans innately respond positively to sweetness and fat. Sweetness signals fast-releasing carbohydrate energy sources, such as in ripe fruit, and motivates consumption. Herbivores such as sheep, cattle, and rabbits will eat more forage grasses cut later in the day, when sugar content is highest, than cut earlier in the morning. Carnivores, by contrast, are indifferent to sweet tastes. The taste and texture of fat signals high-density energy sources. Similarly, it is not an accident that humans generally find bitter tastes unpleasant and objectionable. Bitterness signals the presence of noxious toxins and poisons. Taste is the interface (and guardian) between the external environment of potential foods and our internal bodily environment. It maximizes an organism's chances of survival (Prescott, 2012).

The second function of the liking or hedonic system is modulation of the duration and intensity of an activity. As we will see here and in chapter 12, how hard and long mice and men work at an activity is a direct function of how pleasurable they find the reward (Yeomans, 1996; Young, 1959).

The other component of the reward system is the wanting system. I may find chocolate cake pleasant if it is placed in my mouth, but why should I get up and make the effort to acquire it and place it in my mouth? Motivating me to do so is the job of the wanting or incentive-salience system. It is the motivational magnet for action. The brain systems for wanting are more broadly distributed and involve opioid- and dopamine-sensitive sites in the brain stem including the ventral tegmental area, with mesolimbic projections to ventral pallidum, nucleus accumbens, and amygdala and extend into the orbital frontal cortex, anterior cingulate cortex, and insular cortex, particularly in humans (see figure 11.1c, d). But the evidence suggests that the actual *generators* of these feelings reside in the brain stem, diencephalon, and subcortical structures rather than in the neocortex (Berridge, 2009; Berridge & Kringelbach, 2015). To serve its motivational function, the wanting system must also be able to trigger certain action tendencies (discussed shortly).

In normal cases, liking and wanting act together and constitute the reward system of the brain, which compels the organism to act, and act appropriately. But they seem to have distinct neural bases, and in certain pathological conditions they can become dissociated. Rewarding (liking plus wanting) experiences naturally lead to Pavlovian and operant learning (as in conditioned place preference and conditioned place aversion reported earlier) (Berridge, 2009; Berridge, Ho, Richard, & DiFeliceantonio, 2010; Berridge & Kringelbach, 2015).

There is also a third related system—the general *arousal* system—which has been studied by Donald Pfaff and his colleagues (Pfaff et al., 2012, p. 468), and characterized as "the most powerful and essential force in the nervous system for activating behavior." A more highly aroused animal is more responsive to sensory stimuli, displays greater voluntary motor activity, and is more emotionally reactive. One can imagine the general arousal system modulating the intensity of the liking and wanting systems and serving as the neural basis of the "energy reservoir" in Lorenz's model of instincts (figure 4.1). It has its origins in medullary brain stem structures so it is widely available across branches of the phylogenetic tree.

Having discussed feelings, differentiated them from emotions, identified their origins in phylogenetically old, widely conserved brain stem, diencephalon, and subcortical structures, and proposed that their function is to guide and motivate behavior, I now want to illustrate how feelings are instantiated and operate within each of the four types of minds that account for human behavior.

Feelings and the Autonomic Mind

The autonomic mind is usually considered to be beyond volitional control. This is largely true. While I generally cannot consciously modulate autonomic processes, I can make behavioral and environmental changes that will affect autonomic systems. For example, it was noted earlier that biochemical reactions are critical for maintaining homeostasis in the digestive system (chapter 3). When blood glucose levels drop below a set point, the pancreas will secrete glucagon into the bloodstream, signaling the liver to start converting the stored glycogen into glucose and releasing it into the bloodstream. There is no conscious awareness (feelings) associated with this process. (The same is true of reflexes and many other autonomic processes.) It just happens. I cannot will it or unwill it.[8] If all is going well, I am not, and do not need to be, aware of it.

However, at a certain point, the stored energy reserves will be insufficient to maintain energy requirements so intervention will be required. A meal will need to be ingested. Without this intervention, the system will eventually break down. This intervention will require engagement of one or more of the systems that evolved to interact with the external environment. For nonhuman animals, these will be the instinctive and associative systems. For humans, it will be these plus the reasoning system. But how do I know *when to eat*, *what to eat*, and why should I make the *effort* to do so? Why should I bother, especially if it requires effort and may expose me to predation?

The solution evolution has converged on—perhaps with mammalian brains, perhaps earlier—is the utilization of *feelings* of reward (wanting and liking) and aversion (disgust) as intervening variables between stimulus and response. In the case of energy management, a homeostatic system signals it is time to eat by generating hunger pangs. This is a restless, unpleasant, agitating feeling, which the organism wants to get rid of. Its function is to make us care about initiating or stopping an action. It does so by being directly connected to certain behaviors controlled by instinctive, associative, or reasoning systems. These feelings are drivers, motivators, and inhibitors. They activate certain action tendencies, which result in the organism eventually undertaking the actions to procure and ingest food. Feelings will also guide (via taste) food selection. The control and operation of this energy management system is considered in some detail in chapter 12.

Feelings and the Instinctive Mind

Instinctive behavior is one important way in which organisms satisfy biologically critical goals through interaction with the external environment. These goals include acquisition of food, water, a sexual partner, a nest, or a home territory, securing well-being of offspring, predator avoidance, avoiding injurious levels of hot and cold, and even cheater detection and punishment, among many others. Some of these goals will be species specific, others are available across species.

It is one thing for the autonomic system to signal hunger or sexual arousal but another thing for the organism to get up and actually acquire food or sex. The ethologists struggled with an explanation of why an animal would do this. The initial models of robot-like chain reflex arcs were one solution to the problem, but they did not have the degrees of freedom necessary to accommodate intention actions and vacuum activities, where behavior remains incomplete or is initiated in the absence of the stimuli, respectively (chapter 4). As already noted, Wallace Craig's insight that "an element of appetite, or aversion, or both" (i.e., feelings) between a stimulus and the instinctive behavior was instrumental in the development of the Lorenz and Tinbergen model of instincts (figure 4.1). We can now redescribe this model in neuroscientific terms.

In Berridge and Kringelbach's vocabulary, we would describe Lorenz's action-specific energy reservoir as a specific appetitive (wanting) or aversive (avoiding) state (figure 4.1). An appetite has a positive valence (e.g., sexual arousal), an aversion a negative valence (e.g., hunger pangs). The volume of pent-up energy (i.e., "built-up pressure") in the reservoir corresponds to the level of arousal. Pfaff's general arousal system would serve this function (Pfaff et al., 2012). Separate arousal systems may also be associated with each specific instinctive system, or there may be a generalized arousal system that when coupled with specific feeling systems leads to specific motivated behavior (Garey et al., 2003). The presence of the appropriate stimuli initiates the wanting system (releases the action-specific energy), resulting in execution of the fixed action pattern. The consummatory response (associated with increase of liking or decrease of disliking) is modulated by the level of arousal (volume of pent-up energy) and the properties of the stimuli. The greater the level of arousal, the greater the urge to act (the need to relieve the pressure) by engaging in the consummatory behavior. An animal will actively seek environments in which the behavior can be discharged. The execution provides increased pleasure (or relieves the agitation) and returns the animal to a state of equilibrium.

Not only does this redescription of instincts convert a metaphorical model into a plausible biological model while preserving the critical insights, it also highlights that the model of instincts overlaps with the model of emotions described earlier in that both involve directedness, valence and arousal, physiological expressions, such as relaxed facial muscles and licking of the lips in response to sweet taste (Berridge & Kringelbach, 2015), and triggering of action tendencies. (It differs from human emotions in not involving propositional attitudes.) Neuroscientist Jaak Panksepp has made this connection very explicit by proposing a model where specific appetitive wanting systems (Lorenz's action-specific energy reservoirs) correspond to discrete "primordial emotion" systems (Panksepp, 2011). To minimize confusion between Panksepp's primordial emotion systems in rats and human emotions, I will use the hyphenated form: primordial-emotion. The activation of a primordial-emotion state by the presence of appropriate environmental stimuli, in conjunction with sufficient levels of arousal, will release the fixed action behavioral patterns, resulting in the consummation of the action pattern and the associated feelings of pleasure and relief.

Panksepp identified the following seven systems in rats and proposed that they constitute a basic blueprint applicable to all mammals, including humans: SEEKING, RAGE, FEAR, LUST, CARE, PANIC, and PLAY (Panksepp, 2011). The terms are capitalized, following Panksepp's convention, to indicate that they are being used to label *specific brain systems associated with specific neural networks and neurochemicals* rather than human emotions that might be similarly labeled. Judith Toronchuk and George Ellis (2013) subsequently amended these basic primordial-emotion systems to comprise nine systems by adding DISGUST and POWER systems and relabeling the PANIC system as the NEED/ATTACHMENT system. Table 11.1 lists these systems, grouped into categories of basic functioning, survival, reproduction, social bonding and interaction, and group conflict regulation. The functions of each system are highlighted along with how they interact with other systems to achieve various behaviors. The neuroanatomy and neurochemistry underlying each system are also specified.

What is important for our purposes is not whether there are seven systems, nine systems, or 90 systems. The interesting point is that these systems delineated along neurophysiological lines correspond to the types of things identified as instincts by ethologists and evolutionary psychologists. There is no harm in referring to these states in rats as primordial-emotions so long as my earlier admonition is heeded: any reference to any type of emotion in nonhuman animals is a reference to a directed mental state but one that lacks propositional content.[9] Let's examine one such primordial-emotion or instinct.

The LUST System

I will use the LUST system to illustrate the workings of instinctive systems and the central role of feelings.[10] Since there are behavioral, neuronal, and hormonal gender differences in the operation of the LUST system, the discussion will be confined to males. In males, the appetitive phase commences with wanting and leads to a feeling of sexual arousal, easily recognizable in ourselves.[11] Nonimpotent, postpubescent human males experience specific powerful and pleasurable feelings associated with sexual arousal as a prelude to copulatory behavior. The consummatory phase (copulation) is dominated by liking, building up to orgasm and ejaculation, followed by satiety or restoration of equilibrium. Figure 11.3 graphs these three phases in human males in relation to reported pleasure, along with the associated brain area activations.

Not only is the pleasurable arousal a "prelude" to courtship and copulation behavior, it imparts a degree of urgent compulsion to the initiation of the behavior. The level of desire will differ among individuals as a function of the level of arousal and the quality of the stimuli. The slightest thought, visual perception, even a picture or a dream of nubile females can result in sexual arousal and erection in postpubescent males. Men reportedly think about sex on average 19 times per day (Fisher, Moore, & Pittenger, 2012) and experience sexual desire 37 times per week (Regan & Atkins, 2006).[12] Sometimes it seems that men's whole world revolves around seeking out and discharging these feelings. Once aroused, the only relief comes through ejaculation, which can occur through sexual intercourse or, in the absence of a partner, through masturbation. Sometimes no cost seems too high to incur for this experience. Men are willing to not only expend considerable resources but also bypass social and legal prohibitions at great personal risk to experience and consummate these feelings, as indicated by the John Edwards example in chapter 1 and the following excerpt from *Lolita* (Nabokov, 1991, p. 285):

> I recall certain moments, let us call them icebergs in paradise, when after having had my fill of her—after fabulous, insane exertions that left me limp and azure-barred—I would gather her in my arms with, at last, a mute moan of human tenderness (her skin glistening in the neon light coming from the paved court through the slits in the blind, her soot-black lashes matted, her grave gray eyes more vacant than ever—for all the world a little patient still in the confusion of a drug after a major operation)—and the tenderness would deepen to shame and despair, and I would lull and rock my lone light Lolita in my marble arms, and moan in her warm hair, and caress her at random and mutely ask her blessing, and at the peak of this human agonized selfless tenderness (with my soul actually hanging around her naked body and ready to repent), all at once, ironically, horribly, lust would swell again—and "oh, no," Lolita would say with a sigh to heaven, and the next moment the tenderness and the azure—all would be shattered.

Great novelists, playwrights, and poets make the most perceptive psychologists, but in line with scientific methodology, we must ask to see the hard evidence for the objective existence of such feelings and their causal role in behavior. The same conclusion has been reached less eloquently in behavioral and neuroscience animal research.

Behaviorally, in lab animals the most common and reliable measure of sexual arousal in males is penile erection (tumescence of tissue) (Sachs, 2007). Levels of arousal are measured in various ways, including the time it takes the male to begin mounting an estrus female, the time between intromissions, the time from first intromission to ejaculation, and the time to resume copulation after ejaculation (Clark, 2013). Males in many species display erections in response to remote sexual stimuli: visual erotica for humans, inaccessible estrus females for rhesus monkeys, estrus odors for rats. Humans and rhesus monkeys will masturbate after arousal if a copulatory partner is unavailable (Slimp, Hart, & Goy, 1978). Male rats will mount

Table 11.1

Panksepp's primordial emotions (or instincts)

Evolutionary needs met	Primordial-emotion systems (instincts)	Putative neurochemicals	Putative key components of neural networks	Works with:	Functions
Individual Needs					
Basic functioning	E1: SEEKING system (hedonic appraisal, "liking" component	endorphins (+), GABA (+,–), enkephalins, DA (?), endocannabinoids (+)	nucleus accumbens, ventral pallidum, brain stem nuclei	E2–9	situation evaluation, hedonic appraisal, learning
	E1: SEEKING system (incentive motivation "wanting" component)	DA (+), glutamate, Ach, CCK (+,–), neurotensin, endorphins	nucleus accumbens, ventral pallidum, lateral hypothalamus, and VTA to PAG		incentive salience
Basic survival	E2: DISGUST system (repulsion, avoidance)	serotonin (+), substance P (+)? endocannabinoids (–)	anterior insular cortex, putamen, lower brain stem (area postrema, NTS)		avoiding harmful foods, substances, environments
	E3: RAGE system	substance P (+), Ach (+), glutamate (+), vasopressin (+)	medial amygdala, BNST, medial perifornical hypothalamus, dorsal PAG	E4,E9	defense: protection of organism, resources, and conspecifics, limiting restraint on movement
	E4: FEAR system	glutamate (+), DBI, CRH (+), CCK (+), α-MSH, NPY	lateral and central amygdala, medial and anterior hypothalamus to dorsal PAG and pontine nuclei	E3, E9	defense: flight, limiting of tissue damage
Social Needs					
Reproduction	E5: LUST system (sexual desire)	steroids (+), vasopressin (+), LH-RH (+), DA (+)	basal forebrain, amygdala, BNST, anterior cingulate, medial preoptic, and VMH to ventral PAG	E6, E7	ensuring procreation, enhancement of bonding

	Sexual satisfaction	opioids (+), oxytocin (+)	septum, medial preoptic (VMH in ♂?), VTA to PAG		
Group cohesion: bonding and development	E6: NEED/ATTACHMENT (separation distress)	opioids (–,+), oxytocin (–,+), prolactin (–,+), CRH	anterior cingulate, BNST, POA, VTA to PAG	E5, E7	protection of vulnerable individuals; creates bonding through need for others
	E7: CARE system	oxytocin (+), prolactin (+), DA, opioids (±), glutamate (+)	anterior cingulate, BNST, preoptic hypothalamus to VTA and PAG	E5, E6	caring for others, particularly offspring
	E8: PLAY system	opioids (+,–), DA (+), Ach (?)	dorsomedial diencephalon (thalamic nuclei) to ventral PAG	E6, E7	bonding with conspecifics, development of basic adaptive and social skills, creativity
Group function: regulating conflict	E9: POWER/dominance system (rank, status, submission)	serotonin (±), DA(±), testosterone (±), vasopressin (±), CCK, CRH (±)	medial prefrontal cortex, ventral pallidum and other basal ganglia, hypothalamic nuclei to PAG	E3, E4, E5	limiting aggression in social groups: allocating resources, especially sexual ones

Note: The nonspecific effects of serotonin and norepinephrine are omitted, as are higher cortical areas. Key: Ach = acetylcholine; BNST = bed nucleus of the stria terminalis; CCK = cholecystokinin; CRH = corticotrophin releasing hormone; DA = dopamine; DBI = diazepam binding inhibitor; GABA = gamma-aminobutyric acid; LH-RH = luteinizing hormone, releasing hormone; MSH = melanocyte stimulating hormone; NPY = neuropeptide Y; NTS = nucleus tractus solitarius; PAG = periaqueductal gray; POA = preoptic areas; VMH = ventromedial hypothalamus; VTA = ventral tegmental area.

Source: Reproduced with slight modifications from Toronchuk and Ellis (2013).

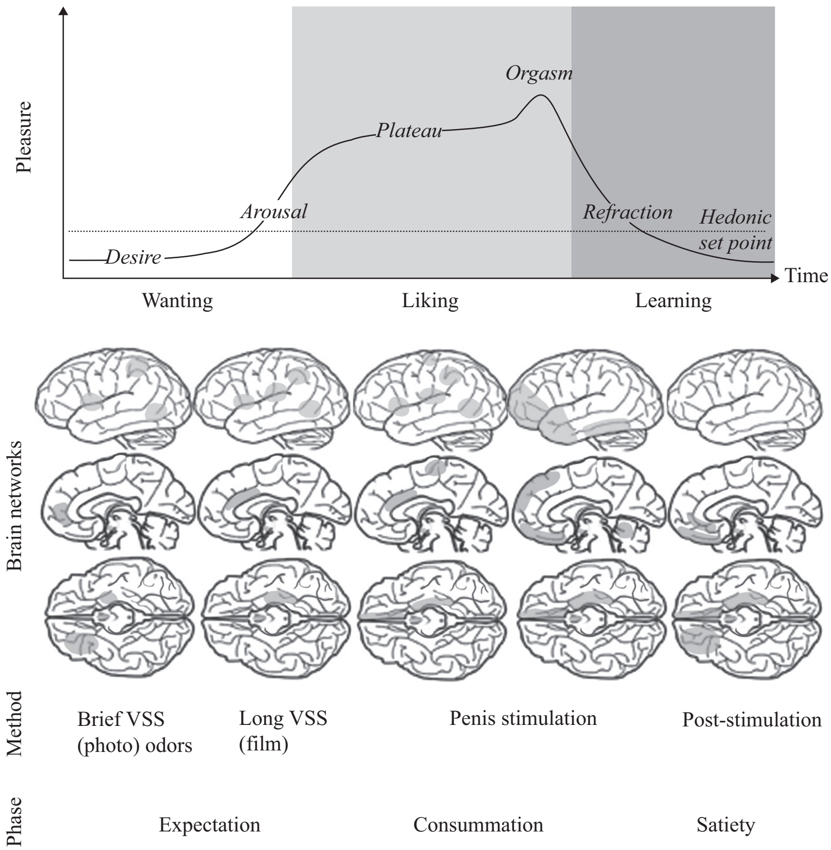

Figure 11.3

The top graph tracks the pleasure modulation associated with the appetitive (wanting), consummatory (liking), and satiety (equilibrium) phases of the LUST system in humans. The brain images show the activation of brain regions as a function of three phases of pleasure. Both subcortical and cortical regions are involved. Given the ubiquity of the LUST system in large segments of the phylogenetic tree, the generators of the feelings and behaviors will be in the brain stem, diencephalon, and subcortical regions, while their representations also involve cortical regions. Reproduced with permission from Georgiadis and Kringelbach (2012).

a female much sooner after noncontact exposure to estrus females than rats not preexposed to the odor of estrus females (Sachs, 2007). The compulsive desire for consummation after arousal can be measured in experiments such as the following. If a partition with a hole is placed between a male rat and an estrus female, the male will become agitated and sniff, chew, and nose poke the partition in continuous efforts to access the female. Its blood testosterone levels will increase. It will linger 12 times longer in the vicinity of the hole when an estrus female is present than in the control condition (Amstislavskaya & Popova, 2004).

There are similarities in hormonal and neural systems and pathways in rats and humans, and sex differences in both. Similar drugs delivered to similar deep brain sites will stimulate sexual desire in both rats and humans (Pfaus, 2009). In human males, the epicenter of sexual arousal seems to be located in the medial anterior hypothalamus, with some variability across species. In rats, this region is the preoptic area.

In adolescent males, sexual maturity commences with the production of testosterone by the testicles. The production of testosterone activates a number of neuropeptides, including vasopressin, which in animal models initiates sexual arousal and courtship. Males produce twice as much vasopressin as females. Testosterone also activates a gaseous nitric oxide transmitter in the brain, which is thought to enhance sexual arousal and aggressiveness. Key areas of the anterior hypothalamus contain such testosterone receptors. Testosterone produces a greater effect in males than in females because male brains have larger areas of testosterone receptors in the anterior hypothalamus than female brains. Thus, unsurprisingly, male rats will work to have testosterone injected into their preoptic area (POA) (Georgiadis & Kringelbach, 2012; Panksepp & Biven, 2012; Pfaus et al., 2012).

In animal models, lesions to the testicles and to the key anterior hypothalamus regions produce similar effects: both weaken sexual urges and abilities. Interestingly, there is an inverted effect of sexual maturity. Preoptic area lesions in sexually naive male animals weaken sexual arousal more than lesions in sexually experienced animals. The latter will continue to work to access estrus females, but their consummatory behavior will be sluggish. One explanation for this is that while sexual urges and responses are being generated in deep brain regions, in the experienced animals they have also become encoded in higher-level cortical regions and can continue in spite of lesions to the subcortical generators.[13]

While we cannot directly access the feelings of sexual arousal in nonhuman animals (or indeed other humans), the behavioral, hormonal, and neuroanatomical homologies reviewed here suggest that feelings are pivotal in driving behavior in both human and nonhuman animals. It is certainly a plausible working hypothesis.

I chose LUST as a paradigmatic example of how hedonic reward systems initiate action because it is something most people have experienced; it has powerful positive valence and arousal components that motivate, even compel, subsequent behavior (mounting and intromission in male rats and lordosis in female rats); and we understand a great deal about the anatomical pathways and neurotransmitter systems involved in each of the anticipatory motivational (wanting), rewarding (liking), and satiety phases of the behavior (Pfaff, 2009). Not only is LUST relevant to explaining the Edwards example from chapter 1, it provides a readily comprehensible blueprint for *all* instinctive behaviors. In particular, it highlights the fact that all instinctive behaviors will be initiated by the reward system and driven by specific interoceptive and exteroceptive feelings.

Similar appetitive feelings drive the compulsion in teenagers to groom, the compulsion in parents to alleviate a child's suffering, the compulsion to punish a cheater, the compulsive attraction to alpha male cues, and so on. The feelings will vary in valence and arousal, may have specific flavors associated with them, and will trigger different action tendencies. The consummatory behavior will be accompanied by feelings of liking and return the human or nonhuman animal to equilibrium.

It is also worth mentioning Panksepp's SEEKING system here. The SEEKING system is the general-purpose wanting, seeking, exploring, searching, motivating, and interest system. It is associated with positive, even euphoric, hedonic value. It is this system that energizes me to get up and pursue everything from chocolate cake to PhDs. The pleasure associated with the SEEKING system is not the sensory pleasure of consummating the act (e.g., eating the chocolate cake) but rather the excitement of pursuit and anticipation (Hamburg, 1971). It interacts with and facilitates the operation of all other primordial-emotion (instinctive) systems (table 11.1) and drives behavior at all levels. It seems very similar to Berridge and Kringelbach's wanting system and will need to contain components of Pfaff's general arousal system. Panksepp suggests that the "generalized reward" system discovered by Olds and Milner was actually the SEEKING system. Anatomically, the SEEKING system involves connections running from the ventral tegmental area to the medial forebrain bundle and lateral hypothalamus, nucleus accumbens, and medial prefrontal cortex. The main neurotransmitter involved is dopamine (Panksepp, 2011; Panksepp & Biven, 2012). The SEEKING system will reemerge at the cognitive level as *desire*.

Feelings and the Associative Mind

How are associative behaviors learned and initiated? Associations are formed through positive and negative reinforcement and punishment. The reader will recall that the behaviorists defined reinforcement and punishment without appealing to feelings. Positive reinforcement occurs when some stimulus event (e.g., picking up and comforting a baby) increases the probability of repetition of the behavior that precedes it (baby crying). Negative reinforcement occurs when the termination of a stimulus event (e.g., baby crying) increases the probability of the repetition of the behavior that follows the stimulus event (e.g., picking up and comforting the baby). A punishment event occurs when the presentation of a stimulus following a behavior results in a decreased probability of that behavior reoccurring. This formulation defines reinforcement and punishment as probability occurrences. It tells us nothing about the reinforcing event itself.

This vacuous construal persisted despite the fact that the reinforcers utilized were pleasing or discomforting biologically important stimuli, such as food, water, sexual partner, harmful levels of heat or cold, odor of a predator, or distress call of an offspring. The natural way to describe an animal's reaction to such stimuli is in terms of appetitive and aversive motivation. Positive reinforcers produce approach behaviors (appetitive motivation) associated with "pleasant hedonic impact," and negative reinforcers and punishers produce withdrawal behavior associated with aversive, unpleasant feelings. Yet subjective experience of the reward or punishment was not allowed to play any role in the theoretical accounts of the behaviorists (Bozarth, 1994).

Despite the failure to acknowledge the rewarding properties of reinforcement, behavioral psychologists ironically continued to measure the amount of reinforcement in grams or count them in pellets, in the number of fluid drops, or as the concentration of sugar in solution. The rats continued to consume the reinforcement! Such vocabulary only makes sense if there is an underlying understanding that the reinforcer is a food object that the animal finds rewarding (Young, 1959). The same is true for a punishing reinforcer.

Several behaviorist psychologists did notice the shortcomings in this approach to reinforcement and suggested that the purpose of reinforcement is "not response strengthening, but the creation of a motivational state that influences a wide variety of subsequent behavior of the animal" (Bindra, 1974, p. 200). Two psychologists who attempted to rectify this shortcoming were Paul Thomas Young, a student of Edward Titchener (himself a student of Wilhelm Wundt), and Dalbir Bindra. The former provided behavioral evidence for the role of feelings and reinforcement (Young, 1959), and the latter developed a brain-based theoretical model of how such a system might work (Bindra, 1969, 1974), not unlike the model developed by Tinbergen and Lorenz already discussed. We will focus on the behavioral data.

Paul Young and his colleagues carried out carefully controlled behavioral experiments in rats to distinguish between food palatability (preference or liking) and appetite (quantity eaten) and sensory intensity and hedonic intensity (e.g., the distinction between detecting the concentration of sugar or salt in a water solution versus preference for the solution), and then used these distinctions to carry out further studies to demonstrate that the hedonic value or valence (pleasant or unpleasant) and arousal or intensity of the stimuli (solution) affect performance (Young & Falk, 1956; Young & Greene, 1953). These experiments provided objective behavioral measures for the causal efficacy of what in humans are referred to as subjective feelings.

In subsequent experiments, rats were offered a choice between a 1% salt solution and sugar solutions of different fixed concentrations ranging from 2% (very weakly sweet) to 54% (very sweet). All rats developed a preference for the sugar solution. However, the speed at which they learned to discriminate between the sugar and salt solutions was a function of the concentration of the sugar solution. Rats in the 54% sugar solution condition required only 17 trials to discriminate between the salt and the sugar solutions, rats in the 18% sugar solution condition required 38 trials, rats in the 6% sugar solution condition required 66 trials, and rats in the 2% sugar solution condition required 122 trials to discriminate (Young & Asdourian, 1957).

In another experiment, rats were trained to run down a runway to a circular platform around the circumference of which were placed five evenly spaced cups. During pretraining, all animals received one drop of 10% sugar solution in each of the five cups. In the first phase of the experiment, one cup was baited with the sugar solution (the others were empty) and the animals learned to run to the baited cup. In the second phase, the animals were divided into four groups, and animals in each group learned to run to a baited cup containing either a 20%, 10%, 5%, or 0% sugar solution. The rate of learning was a function of sugar concentration in the baited cups. The 20% sugar solution resulted in better performance than the 10% solution, which resulted in better performance than the 5% solution, which in turn resulted in better performance than the 0% solution. In fact, the 0% sugar solution resulted in an extinction of the behavior learned in the original training period (Dufort & Kimble, 1956). Both these experiments demonstrate that preference discrimination learning is a function of not only practice (i.e., number of trials run) but also concentration of sugar solution (i.e., preference or palatability of the stimuli).

It is not only the valence or palatability of the reinforcer that is relevant to behavior; the intensity and duration also matter. In another experiment, food pellets were used to train rats to run back and forth from testers containing sucrose, wheat, and casein solutions. Independent tests determined that the rats preferred sucrose to wheat and wheat to casein solutions. Two interesting results were reported: learning was a function of the amount of practice in running the pattern, not the palatability of the incentive solutions. However, the rats ran *faster* for the preferred incentive solution (Young, 1947).

In a follow-up experiment, the reinforcing stimuli controlled for affective intensity (via concentration of sugar solutions), affective duration (by varying the number of seconds in contact with solutions), and frequency of affective arousal (by varying the frequency of access). These factors modulated the rate of running for the food incentive. That is, the intensity, duration, and frequency of affective arousal determined how hard the animals tried (Young & Shuford, 1954). Similar conclusions should apply to other reinforcers, such as sexual behavior, play, and exploration.

As in instinctive behaviors, feelings (appetite for the pleasant feelings of food and sex and aversion to toxins and electric shock) are critical in forming and triggering associative behaviors. The discovery that artificial stimulation of the medial forebrain bundle and lateral hypothalamus regions in rats has an effect on learning similar to that of positive reinforcement suggests that the SEEKING system may be of particular importance in motivating and driving reinforcement behavior (Burgdorf, Knutson, & Panksepp, 2000; Panksepp & Biven, 2012). All these data suggest that, as in the autonomic and instinctive minds, feelings are also the internal currency of the associative mind. What about the reasoning mind?

Feelings and the Reasoning Mind

Rationality is about recruiting reason in the service of a goal or desire. The machinery of reason consists of propositional attitudes and the coherence relation (chapter 6). Is there a role for feelings in the reasoning mind? Can we *feel* the coherency?

Cognitive scientists are happy to commit to representational or intentional mental states but have been as adamant as the behaviorists in avoiding any commitment to affects or feelings. There are no chapters on feelings in cognitive science textbooks. The rationale is that, in the cognitive account, it is the representational content of our mental states that is causally efficacious. Feelings, if they exist, are superfluous.

The most obvious entry point for feelings into the reasoning mind is via emotions. As already noted, emotions are a subset of propositional attitudes such as hope, jealousy, love, fear, shame, surprise, pride, regret, happiness, anger, disgust, contempt, sadness, guilt, and resentment. To say that they are emotional states is to say not only that they are referential or directed, but that they also have valence, arousal, duration, physiological expressions, cognitive antecedents, and action tendencies associated with them.

Emotions, broadly construed, are not unique to the reasoning mind. We have already encountered precursors and noted they are cognized versions of Panksepp's primordial-emotions or instincts. The cognization occurs through the appearance of propositional contents. Once propositional contents are involved, we are largely confined to the hominina or homo branch of the phylogenetic tree. The finer-grained distinctions made possible by propositional contents allow for multiplication of emotions in humans through cognitive construction.

The cognitive strategy with respect to emotions has been to either ignore them or strip off valence, arousal, and duration components and redefine them in terms of complex series of beliefs and desires. For example, my *hope* that P can be redefined as a *desire* for P and a *belief* that the probability of P is very low. My *surprise* that P can be redefined as a *belief* (up to now) that P is not the case and the *belief* (now) that P actually is the case. The issue of affect is not usually raised in the context of beliefs and desires. I am, of course, challenging this approach.

Desires should be another obvious entry point for feelings into the reasoning system. My desire to complete this book has valence and arousal components associated with it. It *feels* like something. It motivates and drives me daily. It is a source of considerable pleasure and satisfaction and occasional frustration. This seems obvious, but cognitive science, restricted to mere information processing, does not have the theoretical machinery to deal with it. In fact, in their information processing theory of problem solving, Allen Newell and Herbert Simon (1972) situated desires or goals *outside* the organism, as part of the task environment! We can remedy this shortcoming by recruiting Panksepp's SEEKING system. This is exactly what desires are. Human desires differ from nonhuman animal desires in that they have sophisticated propositional content that allows them to be directed at an unlimited number of states of affairs in the world, but ultimately, they are a sophisticated cognized variant of the SEEKING system.

What about beliefs? Does my belief that "all apples are fruit" have a feeling associated with it? There is a tradition within philosophy, and certainly within cognitive science, that assumes the answer is "no" (Horgan & Tienson, 2002; Kriegel, 2003; Searle, 1992), yet there may be reasons to reconsider this answer. In 1980, UC Berkeley philosopher (and my mentor) John R. Searle offered the world an argument that came to be known as the Chinese room argument. It was intended to show that a rule-based computational system could not be considered intelligent. There was more to intelligence than following rules. The argument became infamous. It is probably fair to say that no other argument or claim so exercised and consumed the cognitive and artificial intelligence communities during the decade as the Chinese room argument. At the heart of Searle's argument was the following thought experiment (Searle, 1980).

Imagine that you are locked in a room. (It is important that it is you.) You do not know any Chinese, written or spoken. Chinese characters are just meaningless squiggles for you. Inside the room is a box full of cards with Chinese characters written on them. There is also a book of instructions inside the room that gives you rules (in English, which you understand) on how to correlate one set of Chinese symbols with another set. There is a slot through which you can receive and pass out cards. You are now handed a card from outside the room containing a set of Chinese symbols. You consult your rulebook and find in it the symbols appearing on the card. The rulebook tells you the symbols that are correlated with the input symbols. You find the card containing the correlated symbols in your box of symbols and hand it back to the person outside the room. (With practice, you may become so proficient that you have memorized all the rules in the book, so you may not even need to consult the rulebook to generate the correct answer.) To the person outside the room, who speaks Chinese, the card he is handing you contains a question, and the card that you hand back contains the answer to the question. Given that the answers are correct or sensible, that person may well draw the inference that you understand Chinese. Your behavior is certainly consistent with understanding Chinese.

But do you really understand Chinese (or even the symbols written on the cards that you are manipulating)? This is the key question of the thought experiment. Searle concludes that when he places himself in the room he does not understand the "questions" he is being asked and the "answers" that he is providing. He does not understand Chinese. If he (Searle) does not understand, then a rule-based computer program will not understand either.

My concern is not with claims regarding computers and artificial intelligence techniques. It is with why Searle concludes that he does not understand

Chinese. An essential part of the thought experiment that is often overlooked by critics is that Searle himself is in the room (and asks you to put yourself in the room). Searle concludes that he does not understand Chinese because from his first-person perspective it does not *feel* like he understands. If you put yourself in the room, from your first-person perspective, you will most likely draw a similar conclusion. I certainly do. All the vociferous objections to the argument, and Searle's conclusion, ignored first-person experience and relied on behavioral or functional accounts of semantics to argue that Searle did indeed understand Chinese.[14] This thought experiment opens up the possibility that perhaps feelings permeate all propositional attitudes, not just desires and emotional states.[15]

The final component of the reasoning mind is the coherence relation. Can we *feel* coherence? I think so. In introducing the coherence relation, I noted that, ultimately, it is an intuited relation. An argument or a set of beliefs is coherent if it *feels* right. Coherence feels right (positive valence); incoherence or inconsistency feels wrong (negative valence). For example, if Mary is taller than George and George is taller than Michael, it feels right to say that Mary is taller than Michael. It feels wrong to say that Michael is taller than Mary. As with this example, the bathtub analogy from chapter 1, and all self-evident postulates, we cannot prove the rightness of the answer. It is a feeling that presumably all humans with normal cognitive capacity will share. Even in more complex situations, where we appeal to formal, normative rules, these rules are accepted because we can break them down into simple components and test them for the feeling of rightness against our intuitions. As logician Clarence Irving Lewis reportedly noted, when a point of logic is in question, the only thing we can do is appeal to intuition.

Why should this be? Why should there be a feeling associated with coherence relations? As in the case of taste, the feeling is fitness enhancing. Coherency feels good because representations that are internally consistent and veridical will enhance survival. Incoherency feels unpleasant because it can be harmful.

We are creatures whose behavior is a function of our beliefs about the world, rather than the world itself. If I have the belief that there is a tiger under my desk, then I am asserting a certain state of affairs is the case in the world (viz. that there is a tiger under my desk). The source of this belief can be direct perception or an inference based upon perception and/or other beliefs. Irrespective of source, to be useful in the facilitation of my survival and thriving, beliefs need to be veridical. In the case where beliefs are formed by direct perception, there is considerable sensory machinery devoted to getting this largely right, most of the time. In the case of inference, consistency will facilitate veridicality (assuming veridicality of perceptions and existing beliefs).

The importance of veridicality is largely self-evident.[16] Where the beliefs are based on inferences from perceptions and/or other beliefs, the consistency of these inferences becomes critical for guaranteeing veridicality and appropriate actions. For example, if my inferences lead to the belief "tigers are extremely dangerous" and also the belief "tigers are not extremely dangerous," what is it that I believe? More importantly, what do I do when confronted by a tiger: approach or run away? Two different actions are mandated; one will lead to survival, the other to death. For creatures with propositional attitudes, the ability to distinguish between coherency and incoherency is as important as distinguishing between sweetness and bitterness, and for similar reasons. Furthermore, we do it with the same common currency: feelings.

While cognitive scientists studying reasoning have thus far ignored the crucial role of feelings, they do seem to play a central role in Festinger's very influential theory of cognitive dissonance (Festinger, 1957; Harmon-Jones & Mills, 2019). The key idea here is that the presence of discordant or inconsistent beliefs and desires results in cognitive dissonance or discomfort that can sometimes be minimized by changing beliefs. For instance, I believe that I'm overweight and I believe that eating an extra slice of chocolate cake is detrimental to my health. I desire to maintain good health. I also desire an extra slice of chocolate cake. The combination of these beliefs and desires results in cognitive discord or dissonance. I can of course simply choose not to eat the slice of chocolate cake. This follows rationally from my beliefs and my desire to maintain good health. In this case, the theory has nothing to say. Alternatively, I can eat the cake and reduce the cognitive dissonance by telling myself that this cake is made with artificial sweetener and half the regular amount of butter, so it has far fewer calories and will have minimal negative health consequences. This allows me to eat the cake.

There are two possible interpretations of the role of cognitive dissipation of the dissonance in the situation where I eat the cake. First, the introduction of the new beliefs that the cake is made with artificial sweetener and has fewer calories allows me to reason away (reformulate) my previous belief that eating the cake is detrimental to my health, and this allows me to rationally pursue my desire to eat the cake. Second, it may be a cognitive rationalization occurring after the fact. It may make me feel better but provides no explanation for my eating of the cake. In the first interpretation, the behavior is still driven by the reasoning mind. In the second interpretation, the behavior is left unexplained. So while cognitive dissonance theory may recognize the critical role of feelings in reasoning, the actual machinery and control structures underlying behavior are very different from those envisioned by tethered rationality (chapter 12).

Finally, unlike in the case of autonomic, instinctive, and associative systems, we know next to nothing about the brain systems involved in generating the feelings associated with propositional attitudes and coherence relations. There is some evidence for the involvement of the right lateral prefrontal cortex in the detection of inconsistency or incoherency (Goel et al., 2000; Goel & Dolan, 2003; Stollstorff ,Vartanian, & Goel, 2012; Tsujii, Masuda, Akiyama, & Watanabe, 2010; Tsujii, Sakatani, Masuda, Akiyama, & Watanabe, 2011). Making progress along these lines will not be easy because unlike the work on reward systems, primordial-emotions, and interoceptive and exteroceptive affects, animal models cannot be utilized in the study of reason, and the animal-testing techniques cannot be adopted for human participants, for obvious ethical reasons. But because propositional attitudes and coherence relations belong to the reasoning mind, we would expect the corresponding feelings to be constructed in higher cortical and neocortical systems, with input from older subcortical systems (Lieberman & Eisenberger, 2009).

* * *

Feelings permeate all levels of behavior—autonomic,[17] instinctive, associative, and rational. They are not only integral to the operation of each type of mind but also provide a common currency for the *interaction* of the different levels. Accepting such an account paints a very different picture of human choices and decisions than postulated by standard theories. Tethered rationality views behavioral responses to be a blend of the responses generated by the different systems available to organisms. This requires some sort of common currency that can be used for global integration. The suggestion in this chapter is that feelings provide this common currency and allow for global integration of responses of each behavioral system. But we still do not know how the system is controlled. Who is in charge of the tethered mind? This question is addressed in chapter 12.

This is a section of doi:10.7551/mitpress/12811.001.0001

Reason and Less

Pursuing Food, Sex, and Politics

By: Vinod Goel

Citation:

Reason and Less: Pursuing Food, Sex, and Politics
By: Vinod Goel
DOI: 10.7551/mitpress/12811.001.0001
ISBN (electronic): 9780262369701
Publisher: The MIT Press
Published: 2022

The open access edition of this book was made possible by generous funding and support from MIT Press Direct to Open

The MIT Press

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Stone Serif and Stone Sans by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Goel, Vinod, author.
Title: Reason and less : pursuing food, sex, and politics / Vinod Goel.
Description: Cambridge, Massachusetts : The MIT Press, [2022] |
Includes bibliographical references and index.
Identifiers: LCCN 2021017752 | ISBN 9780262045476 (paperback)
Subjects: LCSH: Decision making. | Reasoning. | Logic. | Cognitive neuroscience.
Classification: LCC BF448 .G64 2022 | DDC 153.4/3—dc23
LC record available at https://lccn.loc.gov/2021017752