PART IV

Bound to Know

Chapter Nine

Feeling Feelings

Feeling Feelings

I began this book by describing an obstacle: emotions cannot be known to the subject having them before there is consciousness. Now, after presenting my views on the nature of consciousness, it is time to explain how we can know an emotion. Beginning at the beginning: We know that we have an emotion when the sense of a feeling self is created in our minds. Until there is the sense of a feeling self, in both evolutionary terms as well as in a developing individual, there exist well-orchestrated responses, which constitute an emotion, and ensuing brain representations, which constitute a feeling. But we only know that we feel an emotion when we sense that emotion is sensed as happening in our organism.

The sense of "happening in the organism" comes from representing the proto-self and its changes in second-order structures. The sense of the "emotion as object" comes from representing, in structures

subserving second-order representations, the activity pattern in the induction sites of emotion. Following what was outlined for other objects, I propose that: (1) the inaugural proto-self is represented at second-order level; (2) the "object" that is about to change the proto-self (the neural-activity pattern in emotion-induction sites) is represented at second-order level; (3) the ensuing changes in proto-self (enacted by "body loop" or "as if body loop" mechanisms) are also represented at second-order level.

Feeling an emotion is a simple matter. It consists of having mental images arising from the neural patterns which represent the changes in body and brain that make up an emotion. But knowing that we have that feeling, "feeling" that feeling, occurs only *after* we build the second-order representations necessary for core consciousness. As previously discussed, they are representations of the relationship between the organism and the object (which in this case is an emotion), and of the causal effect of that object on the organism.

The process that I am outlining is precisely the same we discussed for an external object, but it is difficult to envision when the object in question is an emotion, because emotion occurs within the organism, rather than outside of it. The process can only be understood when we keep in mind some of the notions introduced in the chapters on emotion (chapter 2) and on the organism (chapter 5), namely: (1) that there are several brain sites whose activity pattern induces the cortege of actions that become an emotion, and (2) that the activity pattern can be represented within second-order brain structures. Examples of emotion induction sites include nuclei in the hypothalamus, brain stem, basal forebrain, amygdala, and ventromedial prefrontal cortices. Examples of second-order structures include thalamus and cingulate cortices.

It may sound strange, at first, that feelings of emotion—which are steeped in the representation of body states, only come to be known after *other* representations of body state have been integrated to give rise to a proto-self. And it sounds strange, for certain, that the means to know a feeling is another feeling. The situation becomes understandable, however, when we realize that the proto-self, feelings of

emotion, and the feelings of knowing feelings emerged at different points in evolution and to this day emerge at different stages of individual development. Proto-self precedes basic feeling and both precede the feeling of knowing that constitutes core consciousness.

The Substrate for Feelings of Emotion

The collection of neural patterns which constitute the substrate of a feeling arise in two classes of biological changes: changes related to body state and changes related to cognitive state. The changes related to body state can be achieved by two mechanisms.[1] One mechanism involves what I call the "body loop." It uses both humoral signals (chemical messages conveyed via the bloodstream) and neural signals (electrochemical messages conveyed via nerve pathways). As a result of both types of signal, the body landscape is changed and is subsequently represented in somatosensory structures of the central nervous system, from the brain stem on up. The change in the representation of the body landscape can be partly achieved by another mechanism, which involves the "as if body loop." In this alternate mechanism, the representation of body-related changes is created directly in sensory body maps, under the control of other neural sites, for instance, in the prefrontal cortices. It is "as if" the body had really been changed, but it has not. The "as if body loop" mechanism bypasses the body proper, partially or entirely, and I have suggested that bypassing the body saves both time and energy, something that may be helpful in certain circumstances. The "as if" mechanisms are not only important for emotion and feeling, but also for a class of cognitive processes one might designate as "internal simulation."[2]

The changes related to cognitive state are generated when the process of emotion leads to the secretion of certain chemical substances in nuclei of the basal forebrain, hypothalamus, and brain stem, and to the subsequent delivery of those substances to several other brain regions. When these nuclei release neuromodulators in the cerebral cortex, thalamus, and basal ganglia, they cause a host of

significant alterations of brain function. The most important alterations I envision include (1) the induction of particular behaviors (such as bonding and nurturing, playing and exploring); (2) a change in the ongoing processing of body states (as an example, body signals may be filtered or allowed to pass, selectively inhibited or enhanced, and their pleasant or unpleasant quality altered); and (3) a change in the mode of cognitive processing (an example of the latter, in relation to auditory or visual images, would be a change from a slow to a fast rate of image production or a change from sharply focused to vaguely focused images, a change which is an integral part of emotions as disparate as those of sadness or elation).

I suspect all three kinds of change are present in humans and in numerous nonhuman species. It is possible, however, that the third kind of change—the change in the mode of cognitive processing—is only made conscious in humans because it requires an especially high-level representation of neural events: the sort of metarepresentation of aspects of brain processing that only prefrontal cortices are likely to support.

In short, emotional states are defined by myriad changes in the body's chemical profile; by changes in the state of viscera; and by changes in the degree of contraction of varied striated muscles of the face, throat, trunk, and limbs. But they are also defined by changes in the collection of neural structures which cause those changes to occur in the first place and which also cause other significant changes in the state of several neural circuits within the brain itself.

To the simple definition of emotion as a specifically caused transient change of the organism state corresponds a simple definition for feeling an emotion: It is the representation of that transient change in organism state in terms of neural patterns and ensuing images. When those images are accompanied, one instant later, by a sense of self in the act of knowing, and when they are enhanced, they become conscious. They are, in the true sense, feelings of feelings.

There is nothing vague, elusive, or nonspecific about emotional responses, and there is nothing vague, elusive, or nonspecific about the representations which can become feelings of emotions. The substrate

for emotional feelings is a very concrete set of neural patterns in maps of selected structures.

From Emotion to Conscious Feeling

In summary, the complete course of events, from emotion to feeling to feeling of feeling, may be partitioned along five steps, the first three of which were outlined in the chapter on emotion.

1. Engagement of the organism by an inducer of emotion, for instance, a particular object processed visually, resulting in visual representations of the object. The object may be made conscious or not, and may be recognized or not, because neither consciousness of the object nor recognition of the object are necessary for the continuation of the cycle.
2. Signals consequent to the processing of the image of the object activate neural sites that are preset to respond to the particular class of inducer to which the object belongs (emotion-induction sites).
3. The emotion-induction sites trigger a number of responses toward the body and toward other brain sites, and unleash the full range of body and brain responses that constitute emotion.
4. First-order neural maps in both subcortical and cortical regions represent changes in body state, regardless of whether they were achieved via "body loop," "as if body loop," or combined mechanisms. Feelings emerge.
5. The pattern of neural activity at the emotion-induction sites is mapped in second-order neural structures. The proto-self is altered because of these events. The changes in proto-self are also mapped in second-order neural structures. An account of the foregoing events, depicting a relationship between the "emotion object" (the activity at the emotion-induction sites) and the proto-self, is thus organized in second-order structures.

This perspective on emotion, feeling, and knowing is unorthodox. First, I am suggesting that there is no central feeling state before the respective emotion occurs, that expression (emotion) precedes feeling.

Second, I am suggesting that "having a feeling" is not the same as "knowing a feeling," that reflection on feeling is yet another step up. Overall, this curious situation reminds me of E. M. Forster's words: "How can I know what I think before I say it?"

The inescapable and remarkable fact about these three phenomena—emotion, feeling, consciousness—is their body relatedness. We begin with an organism made up of body proper and brain, equipped with certain forms of brain response to certain stimuli and with the ability to represent the internal states caused by reacting to stimuli and engaging repertoires of preset response. As the representations of the body grow in complexity and coordination, they come to constitute an integrated representation of the organism, a proto-self. Once that happens, it becomes possible to engender representations of the proto-self as it is affected by interactions with a given environment. It is only then that consciousness begins, and only thereafter that an organism that is responding beautifully to its environment begins to discover that *it* is responding beautifully to its environment. But all of these processes—emotion, feeling, and consciousness—depend for their execution on representations of the organism. Their shared essence is the body.

What Are Feelings For?

It might be argued that emotions without feelings would be a sufficient mechanism to regulate life and promote survival. It might be argued that signaling the results of that regulatory mechanism would hardly be necessary for survival. But that is simply not the case. Having feelings is of extraordinary value in the orchestration of survival. Emotions are useful in themselves, but the process of feeling begins to alert the organism to the problem that emotion has begun to solve. The simple process of feeling begins to give the organism *incentive* to heed the results of emoting (suffering begins with feelings, although it is enhanced by knowing, and the same can be said for joy). The availability of feeling is also the stepping stone for the next development—*the feeling of knowing that we have feelings.* In turn, knowing is the stepping stone for the process of planning specific and nonstereotyped responses which can either complement an emotion or guarantee that the immediate gains brought by emotion can be maintained over time, or both. In other words, "feeling" feelings extends the reach of emotions by facilitating the planning of novel and customized forms of adaptive response.

Now consider this: Knowing a feeling requires a knower subject. In looking for a good reason for the endurance of consciousness in evolution, one might do worse than say that consciousness endured because organisms so endowed could "feel" their feelings. I am suggesting that the mechanisms which permit consciousness may have prevailed because it was useful for organisms to know of their emotions. And as consciousness prevailed as a biological trait, it became applicable not just to the emotions but to the many stimuli which brought them into action. Eventually consciousness became applicable to the entire range of possible sensory events.

A Note on Background Feelings

What little attention has been paid to the neuroscience of emotion in the twentieth century has been concentrated on the core types of emotion studied by Darwin. Fear, anger, sadness, disgust, surprise, and happiness have been found to be universal emotions in terms of their facial expression and recognizability, as shown in the work of Ekman and others. As a result, the feelings that are most often considered are those which constitute the conscious readout of those major emotions. This would be all well and good if it would not have distracted us from the fact that we continuously have emotional feelings although those feelings are not necessarily part of the set of six "universal feelings" that hail from the six universal emotions. Most of the time we do not experience any of the six emotions, which is certainly a blessing given that four of them are unpleasant. Nor do we experience any of the so-called secondary or social emotions, a good

thing, too, since they hardly fare any better in terms of pleasantness. But we do experience other kinds of emotion, sometimes low grade, sometimes quite intense, and we do sense the general physical tone of our being. I have called the readout of this background perturbation "background feelings," a term I first used in *Descartes' Error,* because these feelings are not in the foreground of our mind. Sometimes we become keenly aware of them and can attend to them specifically. Sometimes we do not and attend, instead, to other mental contents. In one way or another, however, background feelings help define our mental state and color our lives. Background feelings arise from background emotions, and these emotions, although more internally than externally directed, are observable to others in myriad ways: body postures, the speed and design of our movements, and even the tone of our voices and the prosody in our speech as we communicate thoughts that may have little to do with the background emotion. For this reason, I believe it is important to broaden our notion of the source of feelings.

Prominent background feelings include: fatigue; energy; excitement; wellness; sickness; tension; relaxation; surging; dragging; stability; instability; balance; imbalance; harmony; discord. The relation between background feelings and drives and motivations is intimate: drives express themselves directly in background emotions and we eventually become aware of their existence by means of background feelings. The relation between background feelings and moods is also close. Moods are made up of modulated and sustained background feelings as well as modulated and sustained feelings of primary emotions—sadness, in the case of depression. Finally, the relation between background feelings and consciousness is just as close: background feelings and core consciousness are so closely tied that they are not easily separable.

It is probably correct to say that background feelings are a faithful index of momentary parameters of inner organism state. The core ingredients of that index are (1) the temporal and spatial shape of the

operations of the smooth musculature in blood vessels and varied organs, and of the striated muscle of heart and chest; (2) the chemical profile of the milieu close to all those muscle fibers; and (3) the presence or absence of a chemical profile signifying either a threat to the integrity of living tissues or conditions of optimal homeostasis.[3]

Thus, even a phenomenon as simple as background feelings depends on many levels of representation. For instance, some background feelings that have to do with internal milieu and viscera must depend on signals occurring as early as the substantia gelatinosa and intermediate zone of each segment of the spinal cord, and the pars caudalis of the trigeminal nerve nucleus. Other background feelings have to do with the cyclical operations of striated muscle in cardiac function and with patterns of contraction and dilation in smooth muscle which require representations in specific brain-stem nuclei such as the nucleus tractus solitarus and the parabrachial nucleus.

My notion of background feelings is similar to the notion of vitality affects presented by the developmental psychologist Daniel Stern, a notion he uses in his work with infants. That notion was first hinted at by the remarkable but unsung American philosopher Susanne Langer, a disciple of Alfred North Whitehead.[4]

The Obligate Body-Relatedness of Feeling

Regardless of the mechanism by which emotions are induced, the body is the main stage for emotions, either directly or via its representation in somatosensory structures of the brain. But you may have heard that this idea is not correct, that in essence this was the idea proposed by William James—in brief, James proposed that during an emotion the brain causes the body to change, and that the feeling of emotion is the result of perceiving the body's change—and that time has cast the idea aside. First, there is more to my proposal than what was advanced by James. Second, the attack against James, which held sway throughout most of this century and still lingers, is just

not valid, although his proposal on emotion is neither flawless nor complete.

The mechanisms I have outlined to enact emotion and produce a substrate for feelings are compatible with William James's original formulation on this theme but include many features absent in James's text. None of the features I have added undermines or violates the basic idea that feelings are largely a reflection of body-state changes, which is William James's seminal contribution to this subject. The new features I proposed add a new dimension to these phenomena, however. Even in the most typical course of events, the emotional responses target *both* body proper and brain. The brain produces major changes in neural processing that constitute a substantial part of what is perceived as feeling. The body is no longer the exclusive theater for emotions and consequently the body is not the only source for feelings, as James would have wished. Moreover, the body source may be virtual, as it were, it may be the representation of the body "as if" rather than the body "as is." I must add that I did not develop additional features or mechanisms for emotion as a means to circumvent the attacks on William James's idea, although some of my proposals do precisely that. I developed my proposals before I understood what the attackers were attacking.

One might say that there is no need to respond to the critics of William James since his seminal idea is so plausible, but that would be a mistake for several reasons. First, the account offered by William James was understandably incomplete and it must be extended in modern scientific terms. Second, part of the account that was complete was not correct in the detail. For instance, James relied exclusively on representations arising in the viscera, gave short shrift to skeletal muscles as a source for the representation of feelings, and made no mention of the internal milieu. The current evidence suggests that most feelings probably rely on all sources—skeletal and visceral changes as well as changes in internal milieu. The third reason is that the misconceptions that are part of the critique and that

are still cited stand in the way of a comprehensive understanding of emotion and feeling.

Emotion and Feeling after Spinal Cord Transection

The idea that inputs from the body are not relevant to feelings is often based on the false notion that patients with spinal cord transection caused by injury should not be able to emote or feel. The problem, say the critics, is that they seem to be able to emote and to feel. Yet, only a part of the body input most relevant for feelings travels in the spinal cord. First, a considerable part of the relevant information actually travels in nerves such as the vagus, which exit and enter the brain at the level of the brain stem, well above the highest level of the spinal cord possibly damaged by an accident. Likewise, only a part of the enactment of emotions depends on the spinal cord: a large proportion of the process is mediated by cranial nerves at brain-stem level (which can act on the face and on viscera) and by other brain-stem nuclei (which can act directly on the brain above their level).

Second, a significant part of body input actually does not travel by nerves but by way of the bloodstream, again reaching the central nervous system at the level of brain stem, for instance at the area postrema, or higher.

Third, all the surveys of patients with spinal cord damage, including those that seem biased to discover an impairment of feeling and those that were biased to discover that feelings were intact, have revealed some degree of impaired feeling, as one should have expected given that the spinal cord is a *partial* conduit for relevant body input.[5] Moreover, one undisputed fact emerged in those studies: the higher the placement of damage in the spinal cord, the more impaired feeling is. This is important because the higher the section made in the spinal cord, the less input from the body will reach the brain. Higher sections should correlate with less feeling, lower sections with more. The finding would be difficult to explain were it not that some body input is, in fact, precluded by spinal cord damage. (Although it might

be argued, not very credibly, that higher cord lesions by causing greater defects in movement would be accompanied by greater psychological defects and thus less feeling.)

Fourth, spinal cord transections are hardly ever complete, thus allowing for escape pathways into the central nervous system.

Fifth, some of the critics seem to conceive of the body as that part of the organism that is below the neck, the head being just forgotten. As it turns out, the face and skull, as well as the oral cavity, tongue, pharynx, and larynx—whose combination constitutes the upper portion of the respiratory and digestive tracts as well as most of the vocal system—provide a massive input into the brain. This input penetrates the brain at brain-stem level, again at a level higher than that of any spinal cord injury. Since most of the emotions express themselves prominently in changes of the facial musculature, in changes of the musculature of the throat, and in autonomic changes of the skin in the face and scalp, the representation of the related changes in the brain does not need the spinal cord for anything whatsoever and remains available as a base for feelings, even in patients with the most *complete* forms of spinal cord transection.

In conclusion, in normal circumstances we use the spinal cord both to enact a *part* of some emotions and to bring back to the brain signals about *part* of the enactment of those emotions. Accordingly, even the most complete section of the spinal cord fails to disrupt the two-way flow of signals required for emotion and feeling. The fact that any defect is found at all in spinal cord injury supports the notion that body input is relevant to the experience of emotion and feeling; such a defect can hardly be used to argue the opposite. But no one should expect Christopher Reeve not to have emotions and feelings after his accident. The fact that he has both is not evidence against the paramount role of the body in emotion and feeling.

Evidence from the Section of Vagus Nerve and Spinal Cord

The evidence from the section of the vagus nerve or of the vagus nerve and spinal cord has also been misinterpreted ever since W. Cannon turned C. S. Sherrington's experiments in dogs and his own experiments in cats into the centerpiece of his 1927 attack on James.[6] Cannon's argument is an example of the confusions that result from not distinguishing that which is external, such as an emotion, from that which is internal, such as a feeling. Why should a dog or cat, in whom the vagus nerve and spinal cord have been severed, have a complete loss of emotional display, as Cannon predicted? It should not. Severing the vagus nerve *and* the spinal cord does not impede the pathways for the responses that alter the face of the animal, such that it will display rage, fear, or peaceable cooperation with the examiner. Those responses come from the brain stem and are mediated by cranial nerves which were not compromised in Sherrington's or Cannon's experiments. Those facial expressions remained intact after combined sections of the vagus and of the spinal cord, as they should. Dogs responded angrily when shown cats and vice versa, even if they could not move their bodies, which were paralyzed below the neck. (Incidentally, if those animals had been stimulated electrically in the appropriate brain sites, they would have shown the phenomenon known as "sham rage," a display of unmotivated expressions of anger.)

But what about the animal's feelings? They certainly could not be tested, but based on the ideas I have proposed, those feelings were probably altered in part—the animals would receive signals from their facial expressions and would have intact signaling from brain-stem nuclei, both of which would be a base for feeling, but they would not receive visceral input which would have been based on signals from the vagus nerve and the spinal cord. At this point, Cannon threw caution to the winds and wondered if feelings could possibly be far when there was so much of an emotional display. He took the presence of emotion as a sure sign for the presence of feeling. The error rests entirely with the failure of making a principled distinction between emotion and feeling and of recognizing the sequential, unidirectional enchainment of the process—from inducer, to automated emotion, to representation of emotional changes, to feeling.

Lessons from Locked-In Syndrome

One of the most intriguing, albeit indirect, lines of evidence for the importance of body input in the generation of feelings comes from locked-in syndrome. As discussed in chapter 8, locked-in occurs when a part of the brain stem such as the pons or midbrain is damaged anteriorly, in its ventral aspect, rather than posteriorly, in its dorsal aspect. The motor pathways which convey signals to the skeletal muscles are destroyed, and only one pathway for vertical movement of the eyes is spared, sometimes not completely. The lesions that cause locked-in are placed directly in front of the area whose lesions cause coma or persistent vegetative state, yet locked-in patients have an intact consciousness. They cannot move any muscle in their face, limbs, or trunk, and their communication ability is usually limited to vertical movements of the eyes, sometimes one eye only. But they remain awake, alert, and conscious of their mental activity. The voluntary blinking of these patients is their sole means of communicating with the outside world. Using a blink to signify a letter of the alphabet is the laborious technique with which locked-in patients compose words, sentences, and even books, slowly dictated—one should say blinked—to an attentive note taker.

A remarkable aspect of this tragic condition and one that has been neglected to date is that although patients are plunged, fully conscious, from a state of human freedom to one of nearly complete mechanical imprisonment, they do not experience the anguish and turmoil that their horrifying situation would lead observers to expect. They have a considerable range of feelings, from sadness to, yes, joy. And yet, from accounts now published in book form, the patients may even experience a strange tranquillity that is new to their lives. They are fully aware of the tragedy of their situation, and they can report an intellectual sense of sadness or frustration with their virtual imprisonment. But they do not report the terror that one imagines would arise in their horrible circumstances. They do not seem to have anything like the acute fear experienced by so many perfectly healthy and mobile individuals inside a magnetic resonance scanner, not to mention a crowded elevator.[7]

My way of explaining this surprising finding is as follows: Blinking and vertical eye movements aside, the damage in locked-in precludes any motion, either voluntary or enacted by emotional responses, of any part of the body. Facial expression and bodily gestures in response to a deliberate intention or an emotion are precluded (there is only a partial exception—tears can be produced although the motor accompaniments of crying are missing). Under the circumstances, any mental process which would normally induce an emotion fails to do so through the "body loop" mechanism we have discussed. The brain is deprived of the body as a theater for emotional realization. Nonetheless, the brain can still activate emotion-induction sites in the basal forebrain, hypothalamus, and brain stem, and generate some of the internal brain changes on which feelings depend. Moreover, since most signaling systems from body to brain are free and clear, the brain can get direct neural and chemical signaling from organism profiles that fit background emotions. Those profiles are related to basic regulatory aspects of the internal milieu and are largely uncoupled from the patient's mental state because of brain-stem damage (only the bloodstream chemical routes remain open both ways). I suspect that some of the internal-milieu states are perceived as calm and harmonious. Support for this idea comes from the fact that when these patients have a condition which ought to produce pain or discomfort, they can still register the presence of that condition. For instance, they feel stiff and cramped when they are not moved by others for a long time. Curiously, the suffering that usually follows pain seems to be blunted, perhaps because suffering is caused by emotion, and emotion can no longer be produced in the body theater: it is restricted to "as if body" mechanisms.

Another line of evidence corroborating this interpretation comes from patients undergoing surgery who receive an injection of curare, a substance that blocks the activity of skeletal muscles by acting on

the nicotinic receptors of acetylcholine. If curare acts before the proper induction of anesthesia suspends consciousness, the patients become aware of their paralysis. Like patients with locked-in, curarized patients are able to hear the conversations of those around. Based on reports obtained after the event, these patients are less calm than patients with locked-in and closer to what one might expect if one imagines being in the same situation. There may be a clue to explain the difference. Curare blocks the nicotinic receptors of acetylcholine, the transmitter that is necessary for nerve impulses to contract muscular fibers. Since the skeletal muscles throughout our face, limbs, and trunk are of the striated type and have such nicotinic receptors, curare blocks neurochemical impulses at the site of all those neuromuscular junctions and causes paralysis. However, the nerve impulses that lead smooth muscles to respond under the autonomic control of emotions use muscarinic receptors that are *not* blocked by curare. Under the circumstances, it is possible for one part of the emotional responses, that which depends on pure autonomic signals, to be enacted in the body theater and be represented back in neural structures.

As a whole, this evidence suggests that the "body loop" mechanism of emotion and feeling is of greater importance for real experience of feelings than the "as if body loop" mechanism that I have proposed as an alternate and complement.

Learning from Emotion with the Help of the Body

A recent series of learning experiments also provides evidence for the role of the body in emotion. It has been demonstrated, in both rats and humans, that recall of new facts is enhanced by the presence of certain degrees of emotion during learning. James McGaugh and his colleagues have led these studies whose results are now well confirmed.[8] For instance, if you are told two stories of comparable length that have a comparable number of facts, differing only because in one of them the facts have a high emotional content, you will remember far more detail from the emotional story than from the other. You

may be pleased to know that we have this in common with rats when they are placed in an equivalent situation. They, too, have better success in a standard learning situation when a certain amount of emotion happens at the right time. Now, after the vagus nerves of the rats are severed, emotion no longer helps their performance. Why so? Well, without the vagus, the rats are also deprived of substantial visceral input to the brain. It must be the case that the particular visceral input now missing is vital for the sort of emotion that assists learning.