emotion review

Emotion Review
Vol. 5, No. 3 (July 2013) 241–247

ISSN 1754-0739
DOI: 10.1177/1754073913477514
er.sagepub.com

What Are Emotional States, and Why Do We Have Them?

Edmund T. Rolls
Oxford Centre for Computational Neuroscience, UK
Department of Computer Science, University of Warwick, UK

Abstract

An approach to emotion is described in which emotions are defined as states elicited by instrumental reinforcers, that is, by stimuli that are the goals for action. This leads to a theory of the evolutionary adaptive value of emotions, which is that different genes specify different goals in their own self-interest, and any actions can then be learned and performed by instrumental learning to obtain the goals. The brain mechanisms for emotion in brain regions such as the orbitofrontal and anterior cingulate cortex provide a foundation for understanding the neural basis of emotion. Classically conditioned effects may modulate the actions initiated by this system. In addition to this instrumental learning system, some stimuli may elicit responses, for example approach, withdrawal, or fixed action patterns, but intervening states are not required for this type of adaptive response. In addition, a rational thought system involved in multistep planning can allow gene-specified goals to be deferred or avoided in order to achieve longer-term types of goals that may be more advantageous to the individual than to the genes.

Keywords
amygdala, approach, avoidance, classical conditioning, instrumental reinforcer, orbitofrontal cortex, punisher, reward

Emotions as States Elicited by Instrumental Reinforcers

Emotions can usefully be defined (operationally) as states elicited by rewards and punishers which have particular functions (Rolls, 1999, 2005b, in press). The functions are defined in what follows, and include working to obtain or avoid the rewards and punishers, respectively. A reward is anything for which an animal (which includes humans) will work. A punisher is anything that an animal will escape from or avoid. An example of an emotion might thus be the happiness produced by being given a particular reward, such as a pleasant touch, praise, or winning a large sum of money. Another example of an emotion might be fear produced by the sound of a rapidly approaching bus, or the sight of an angry expression on someone's face. We will work to avoid such stimuli, which are punishing. Another example would be frustration, anger, or sadness produced by the omission of an expected reward, or the termination of a reward such as the death of a loved one. Another example would be relief, produced by the omission or termination of a punishing stimulus such as the removal of a painful stimulus, or sailing out of danger. These examples indicate how emotions can be produced by the delivery, omission, or termination of rewarding or punishing stimuli, and go some way to indicating how different emotions could be produced and classified in terms of the rewards and punishers received, omitted, or terminated. A diagram summarizing some of the emotions associated with the delivery of a reward or punisher or a stimulus associated with them, or with the omission of a reward or punisher, is shown in Figure 1.

Before accepting this approach, we should consider whether there are any exceptions to the proposed rule. Are any emotions caused by stimuli, events, or remembered events that are not rewarding or punishing? Do any rewarding or punishing stimuli not cause emotions? We will consider these questions in more detail in the following lines. The point is that if there are no major exceptions, or if any exceptions can be clearly encapsulated, then we may have a good working definition at least of what causes emotions. Moreover, it is worth pointing out that many approaches to or theories of emotion have in common that part of the process involves "appraisal" (Frijda, 1986; Lazarus, 1991; Oatley & Jenkins, 1996; Scherer, 2009). In all these theories the concept of appraisal presumably involves assessing whether something is rewarding or punishing. The description in terms of reward or punishment adopted here seems more tightly and operationally specified.

I consider elsewhere a slightly more formal definition than rewards or punishers, in which the concept of reinforcers is introduced, and it is shown that emotions can be usefully seen as states produced by instrumental reinforcing stimuli (Rolls, 2005b). Instrumental reinforcers are stimuli which, if their occurrence, termination, or omission is made contingent upon

Corresponding author: Edmund T. Rolls, Department of Computer Science, University of Warwick, Warwick CV4 7AL, UK. *Email*: Edmund.Rolls@oxcns.org

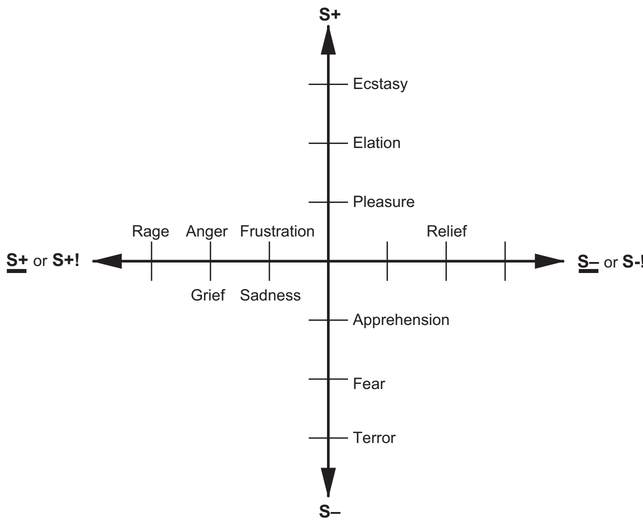

Figure 1. Some of the emotions associated with different reinforcement contingencies are indicated. Intensity increases away from the centre of the diagram, on a continuous scale. The classification scheme created by the different reinforcement contingencies consists of (a) the presentation of a positive reinforcer (S+), (b) the presentation of a negative reinforcer (S–), (c) the omission of a positive reinforcer (S+) or the termination of a positive reinforcer (S+!), and (d) the omission of a negative reinforcer (S–) or the termination of a negative reinforcer (S–!). It should be understood that each different reinforcer will produce different emotional states: this diagram just summarizes the types of emotion that may be elicited by different contingencies, but the actual emotions will be different for each reinforcer.

the making of a response, alter the probability of the future emission of that response. Some stimuli are unlearned reinforcers (e.g., the taste of food if the animal is hungry, or pain); while others may become reinforcing by associative learning because of their association with such primary reinforcers, thereby becoming "secondary reinforcers."

This foundation has been developed (Rolls, 2005b) to show how a very wide range of emotions can be accounted for, as a result of the operation of a number of factors, including the following:

1. The *reinforcement contingency* (e.g., whether reward or punishment is given, or withheld; see Figure 1).
2. The *intensity* of the reinforcer (see Figure 1).
3. Any environmental stimulus might have a *number of different reinforcement associations*. (For example, a stimulus might be associated with the presentation of both a reward and a punisher, allowing states such as conflict and guilt to arise.)
4. Emotions elicited by stimuli associated with *different primary reinforcers* will be different.
5. Emotions elicited by *different secondary reinforcing stimuli* will be different from each other (even if the primary reinforcer is similar).
6. The emotion elicited can depend on whether an *active or passive behavioural response* is possible. (For example, if an active behavioural response can occur to the omission of a positive reinforcer, then anger might be produced, but if only passive behaviour is possible, then sadness, depression or grief might occur.)

By combining these six factors, it is possible to account for a very wide range of emotions (Rolls, 2005b, in press). It is also worth noting that emotions can be produced just as much by the recall of reinforcing events as by external reinforcing stimuli; that cognitive processing (whether conscious or not) is important in many emotions, for very complex cognitive processing may be required to determine whether or not environmental events are reinforcing. Indeed, emotions normally consist of cognitive processing which analyzes the stimulus, then determines its reinforcing valence; and then an elicited mood change if the valence is positive or negative. I note that a mood or affective state may occur in the absence of an external stimulus, as in some types of depression, but that normally the mood or affective state is produced by an external stimulus, with the whole process of stimulus representation, evaluation in terms of reward or punishment, and the resulting mood or affect being referred to as emotion (Rolls, in press).

The Functions of Emotion

The functions of emotion also provide insight into the nature of emotion. These functions, described more fully elsewhere (Rolls, in press), can be summarized as follows:

1. The *elicitation of autonomic responses* (e.g., a change in heart rate) and *endocrine responses* (e.g., the release of adrenaline). These prepare the body for action, and are responses (not instrumental actions) produced by stimuli that produce emotions, and can be classically conditioned.
2. *Flexibility of behavioural responses to reinforcing stimuli*. Emotional (and motivational) states allow a simple interface between sensory inputs and action systems. The essence of this idea is that goals for behaviour are specified by reward and punishment evaluation. When an environmental stimulus has been decoded as a primary reward or punishment, or (after previous stimulus–reinforcer association learning) a secondary rewarding or punishing stimulus, then it becomes a goal for action. The human can then perform any action to obtain the reward, or to avoid the punisher. (Instrumental learning typically allows any action to be learned, though some actions may be more easily learned than others [Lieberman, 2000; Pearce, 2008].) Thus there is flexibility of action, and this is in contrast with stimulus–response or habit learning in which a particular response to a particular stimulus is learned. The emotional route to action is flexible not only because any action can be performed to obtain the reward or avoid the punishment, but also because the human can learn in as little as one trial that a reward or punishment is associated with a particular stimulus, in what is termed "stimulus–reinforcer association learning."

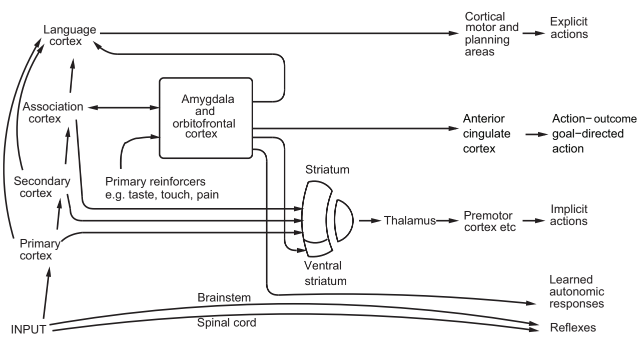

Figure 2. Dual routes to the initiation of action in response to rewarding and punishing stimuli. The inputs from different sensory systems to brain structures such as the orbitofrontal cortex and amygdala allow these brain structures to evaluate the reward- or punishment-related value of incoming stimuli, or of remembered stimuli. The different sensory inputs enable evaluations within the orbitofrontal cortex and amygdala based mainly on the primary (unlearned) reinforcement value for taste, touch, and olfactory stimuli, and on the secondary (learned) reinforcement value for visual and auditory stimuli. In the case of vision, the "association cortex" which outputs representations of objects to the amygdala and orbitofrontal cortex is the inferior temporal visual cortex. One route for the outputs from these evaluative brain structures is via projections directly to structures such as the basal ganglia (including the striatum and ventral striatum) to enable implicit, direct behavioural responses based on the reward- or punishment-related evaluation of the stimuli to be made. The second route is via the language systems of the brain, which allow explicit decisions involving multistep syntactic planning to be implemented.

To summarize and formalize, two processes are involved in emotional behaviour. The first is stimulus–reinforcer association learning; emotional states are produced as a result (Rolls, in press). This process is implemented in structures such as the orbitofrontal cortex and amygdala (see Figures 2 and 3; Grabenhorst & Rolls, 2011; Rolls, in press; Rolls & Grabenhorst, 2008). The second is instrumental learning of an action made to approach and obtain the reward or to avoid or escape from the punisher. This is action–outcome learning, and involves brain regions such as the cingulate cortex when the actions are being guided by the goals, and the striatum and rest of the basal ganglia when the behaviour becomes automatic and habit-based, that is, uses stimulus–response connections (see Figures 2 and 3; Rolls, 2005b, 2009, in press; Rushworth, Noonan, Boorman, Walton, & Behrens, 2011). Emotion is an integral part of this, for it is the state elicited in the first stage by stimuli which are decoded as rewards or punishers, and this state has the property that it is motivating. The motivation is to obtain the reward or avoid the punisher (the goals for the action), and animals must be built to obtain certain rewards and avoid certain punishers. Indeed, primary or unlearned rewards and punishers are specified by genes which effectively determine the goals for action. This is the solution that natural selection has found for how genes can influence behaviour to promote their fitness (as measured by reproductive success), and for how the brain could interface sensory systems to action systems, and is an important part of Rolls' theory of emotion (2005b, in press).

Selecting between available rewards with their associated costs, and avoiding punishers with their associated costs, is a process that can take place both implicitly (unconsciously) and explicitly using a language system to enable long-term plans to be made (Rolls, 2005b, 2008b). These many different brain systems, some involving implicit evaluation of rewards and others explicit, verbal, conscious, evaluation of rewards and planned long-term goals, must all enter into the selector of behaviour (see Figure 2).

The implication is that operation by animals (including humans) using reward and punishment systems tuned to dimensions of the environment that increase fitness provides a mode of operation that can work in organisms that evolve by natural selection. It is clearly a natural outcome of Darwinian evolution to operate using reward and punishment systems tuned to fitness-related dimensions of the environment, if arbitrary actions are to be made by the animals, rather than just preprogrammed movements such as tropisms, taxes, reflexes, and fixed action patterns. This view of brain design in terms of reward and punishment systems built by genes that gain their adaptive value by being tuned to a goal for action offers, I believe, a deep insight into how natural selection has shaped many brain systems, and is a fascinating outcome of Darwinian thought (Rolls, 2005b, 2011b, in press).

3. Emotion is *motivating*, as just described. For example, fear learned by stimulus-reinforcement association provides the motivation for actions performed to avoid noxious stimuli.
4. *Communication*. Monkeys, for example, may communicate their emotional state to others, for instance by making face expressions (such as an open-mouth threat to indicate the extent to which they are willing to compete for resources), and this may influence the behaviour of other animals. This aspect of emotion was emphasized by Darwin (1872), and has been studied more recently by Ekman (1993) and, in terms of the brain mechanisms, by Rolls (2005b, in press).
5. *Social bonding*. Examples of this are the emotions associated with the attachment of the parents to their young, and the attachment of the young to their parents.
6. The current mood state can affect the *cognitive evaluation of events or memories* (Blaney, 1986; Oatley & Jenkins, 1996). For example, happy memories are more likely to be recalled when happy. Another example is that when people are in a depressed mood, they tend to recall memories that were stored when they were depressed. The recall of depressing memories when depressed can have the effect of perpetuating the depression, and this may be a factor with relevance to the aetiology and treatment of depression. The interactions between mood and memory systems using neural networks that capture the effects of backprojection connectivity from emotional to perceptual and cognitive

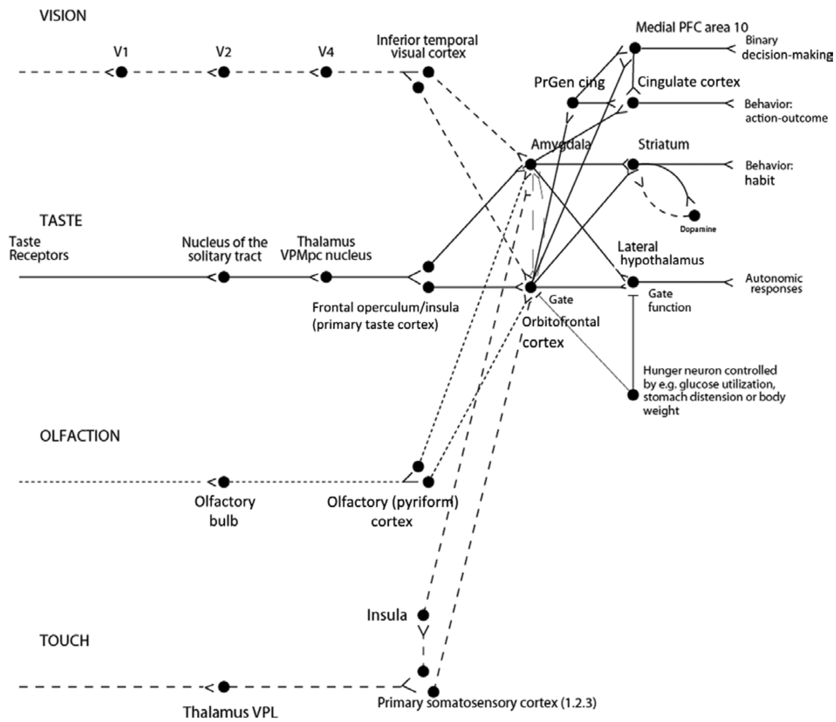

Figure 3. Schematic diagram showing some of the gustatory, olfactory, visual, and somatosensory pathways to the orbitofrontal cortex, and some of the outputs of the orbitofrontal cortex, in primates. The secondary taste cortex and the secondary olfactory cortex are within the orbitofrontal cortex. V1: primary visual cortex; V4: visual cortical area V4; PreGen Cing: pregenual cingulate cortex. "Gate" refers to the finding that inputs such as the taste, smell, and sight of food in some brain regions only produce effects when hunger is present (Rolls, 2005b). The column of brain regions including and below the inferior temporal visual cortex represents brain regions in which what stimulus is present is made explicit in the neuronal representation, but not its reward or affective value which are represented in the next tier of brain regions, the orbitofrontal cortex and amygdala, and in the anterior cingulate cortex. In areas beyond these, such as medial prefrontal cortex area 10, choices or decisions about reward value are taken, with the mechanisms described elsewhere (Rolls, 2008b, in press; Rolls & Deco, 2010). Medial PFC area 10: medial prefrontal cortex area 10; VPMpc: ventralposteromedial thalamic nucleus.

systems have been analysed by Rolls and Stringer (Rolls, 2008b; Rolls & Stringer, 2001).

7. Emotion may facilitate the *storage of memories*. One way this occurs is that episodic memory (i.e., one's memory of particular episodes) is facilitated by emotional states (Rolls, 2005b, 2008b). A second way in which emotion may affect the storage of memories is that the current emotional state may be stored with episodic memories, providing a mechanism for the current emotional state to affect which memories are recalled. A third way is by guiding the cerebral cortex in the representations of the world that are set up, using backprojections (Rolls, 2008b).
8. Another function of emotion is that by enduring for minutes or longer after a reinforcing stimulus has occurred, it may help to produce *persistent and continuing motivation and direction of behaviour*, to help achieve a goal or goals.
9. Emotion may trigger the *recall of memories* stored in neocortical representations. Amygdala backprojections to the cortex could perform this for emotion in a way analogous to that in which the hippocampus could implement the retrieval in the neocortex of recent (episodic) memories (Rolls, 2008b; Rolls & Stringer, 2001).

Different Systems for Emotional Learning and Memory

When stimuli are paired with primary reinforcers, associations that perform many types of function are formed. Some are as follows, and are described in more detail in *Emotion and Decision-Making Explained* (Rolls, in press). The importance of this is that many processes take place during emotion, and they can all contribute to the richness and sometimes the inconsistency of what happens during emotional behaviour. Understanding the diversity of these processes provides a foundation for analyses and descriptions of emotional behaviour (including those found in literature).

First, as shown in Figure 4, Pavlovian (classical) conditioning (in which a stimulus is paired with another stimulus or response, and where the actions have no influence on the pairing) has the potential to create multiple associative representations in the brain, as described next (Cardinal, Parkinson, Hall, & Everitt, 2002; Rolls, in press).

1. *Stimulus–response (S–R) association*. The conditioned stimulus (CS) may become directly associated with the unconditioned response (UR), a simple stimulus–response association that carries no information about the identity of the unconditioned stimulus (US; Pathway 1 in

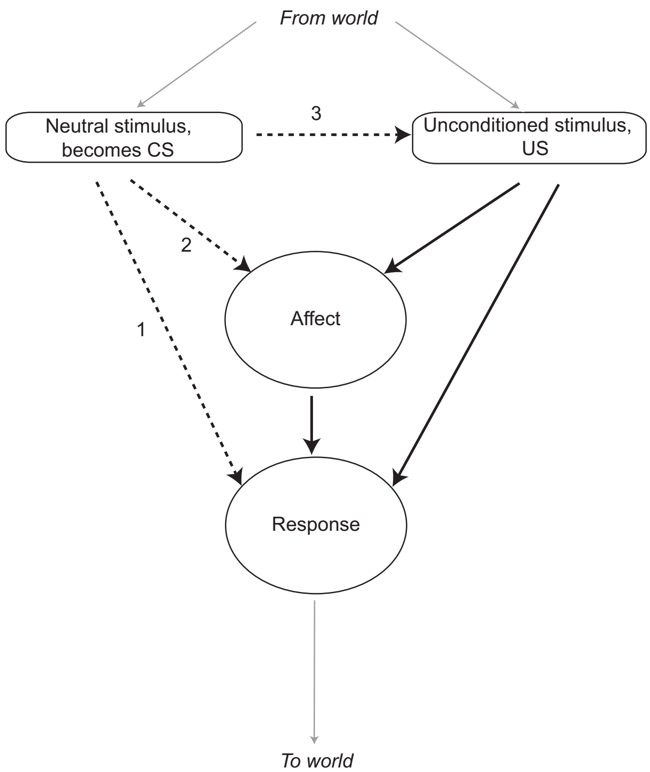

Figure 4. Pavlovian conditioning has the potential to create associations between a conditioned stimulus (CS) and representations of the unconditioned stimulus (US), central affective or emotional states such as fear, and unconditioned responses. Dashed lines represent associatively learned links. Several different types of response may be involved, including preparatory responses which are not specific to the type of US involved (e.g., orienting to a stimulus, or increased arousal), and "consummatory" responses which are specific to the US such as salivation to food, or blinking to an air puff applied to the eye (after Cardinal et al., 2002).

Figure 4). Such US-elicited responses include preparatory responses that are not specific to the type of US, involved (e.g., orienting to a stimulus, or increased arousal), and "consummatory" responses which are specific to the US such as salivation to food, or blinking to an air puff applied to the eye, or approach to a food. A single US may elicit both preparatory and consummatory responses, and thus the CS may enter into simple S–R associations with several types of response.

2. *A representation of affect, that is, an emotional state*. The CS can evoke a representation of affect, that is, an emotional state, such as fear or the expectation of reward (Pathway 2 in Figure 4). It is demonstrated operationally by the phenomenon of transreinforcer blocking (Cardinal et al., 2002). However, I note that, at least in humans, affective states normally have content; that is, they are about particular reinforcers (such as feeling happy because I am seeing a friend, or feeling happy because I am receiving a gift), and these states are better described by the third type of association, detailed next.
3. *Conditioned-stimulus (CS)–unconditioned stimulus (US) associations*. The CS can become associated with the specific sensory properties of the US including its visual appearance, sound, and smell as well as its "consummatory" (primary reinforcing) properties such as its taste, nutritive value, and feel (Pathway 3 in Figure 4). This is the process involved in stimulus–reinforcer association learning, and in the brain involves structures such as the orbitofrontal cortex and amygdala (Grabenhorst & Rolls, 2011; Rolls, in press; Rolls & Grabenhorst, 2008).

   Different pathways in the brain are involved in the Pavlovian learned autonomic and skeletal responses to a CS, and in the affective representation or state (e.g., fear), which may itself enter into associations and influence choice (Rolls, in press).

Second, in instrumental learning, there is a contingency between the behaviour and the reinforcing outcome. A number of different learning processes may operate during this procedure, which, it turns out, may have somewhat different brain implementations (Cardinal et al., 2002; Rolls, in press). One key process is action–outcome learning. The outcome is represented as reward or affective value, such as that implemented by the firing of orbitofrontal cortex neurons that respond to the taste of food only if hunger is present. Other processes influence instrumental learning including Pavlovian processes that can facilitate performance (as in Pavlovian-instrumental transfer). Further, approach to a food may be under Pavlovian rather than instrumental control.

We must be aware of the fact that after overtraining, responses may become inflexibly linked to stimuli, and that the goals, and the reward value of the goals, may no longer be directly influencing behaviour in an ongoing way. My theory is that normally we want because we like. Indeed, that is inherent in my theory, for the genes that make a stimulus (such as a sweet taste) rewarding (i.e., wanted, a goal for action) also make the stimulus liked (i.e., accepted, with a subjective correlate of pleasure, pleasantness, and affective liking). If behaviour becomes overlearned and a habit or stimulus–response connection is built up by another brain system, then animals may make automatic responses that are not goal directed. There has been confusion in the literature caused by overlooking this point (Berridge, Robinson, & Aldridge, 2009). The fact that behaviour can become stimulus–response and no longer under the control of the goal need not surprise us. Normally, and certainly during learning before habits set in, we want a goal, and when we get the goal we like it: goal stimuli normally specify what is wanted, and what is liked (Rolls, 2005b, in press).

The impact of this analysis in the present context of approach and avoidance is as follows. There are many different brain processes involved in learning emotional responses, different brain regions are involved in the different types of learning, and what is learnt in each system may be somewhat independent of what is learned in the other systems (Rolls, 2005b, in press). We should thus not assume that emotion is a single unified process. There may be many different underlying processes that take place, and they are not always consistent with each other. My hypothesis is that it is the states elicited by instrumental reinforcers that are emotional states. Stimuli that produce such states may have other effects, eliciting perhaps autonomic responses, approach, fixed action patterns, and learned habits, and while these responses are adaptive, they do not require the intervening states involved in instrumental actions made to obtain goals, and are therefore not crucial to emotional states. The emotional states produced by such instrumental reinforcers are states of the brain, the firing of neurons, that action systems in the brain seek to obtain or avoid by producing instrumental actions. These emotional states may or may not be conscious: my approach to consciousness suggests that emotional states may gain access to consciousness especially when we must perform reasoning that involves these states, and correcting errors in such reasoning (Rolls, 2007a, 2007b, 2008a).

A Separate, Rational, Reasoning, Conscious System for Identifying Emotional Goals

I have put forward a position elsewhere that, in addition to the gene-based goal system for emotion described before, there is a separate rational, that is reasoning, system that can plan ahead and work for what are sometimes different, long-term, goals (Rolls, 1997b, 2003, 2004, 2005a, 2005b, 2007a, 2007b, 2008a, 2011a, 2012). This type of processing involves multistep trains of thought, as might be required to formulate a plan with many steps. Each step has its own symbols (e.g., a word to represent a person), and so syntactic linking (binding) is needed between the symbols within each step, and some syntactic (relational) links must be made between symbols in different steps. I have argued that when we correct such multistep plans or trains of thought, we need to think about these first-order thoughts, and the system that does this is thus a higher-order thought system (in that it is thinking about first-order thoughts).

There is a fundamentally important distinction here: Working for a gene-specified reward, as in many emotions, is performed for the interests of the "selfish" genes. Working for rationally planned rewards may be performed in the interest of the particular individual (e.g., the person), and not in the interests of the genotype (Rolls, 2011a).

It is suggested that this arbitrary symbol manipulation, using important aspects of language processing and used for planning, but not for initiating, all types of behaviour, is close to what consciousness is about. In particular, consciousness may *be* the state which arises in a system that can think about (or reflect on) its own (or other people's) thoughts, that is in a system capable of second or higher order thoughts (Carruthers, 1996; Dennett, 1991; Gennaro, 2004; Rolls, 1995, 1997a, 1997b, 1999, 2004, 2005b, 2007a; Rosenthal, 1986, 1990, 1993, 2004, 2005).

It is of great interest to comment on how the evolution of a system for flexible planning might affect emotions. Consider grief which may occur when a reward is terminated and no immediate action is possible (see Rolls, 1990, 2005b). It may be adaptive by leading to a cessation of the formerly rewarded behaviour and thus facilitating the possible identification of other positive reinforcers in the environment. In humans, grief may be particularly and especially potent because it becomes represented in a system which can plan ahead, and understand the enduring implications of the loss.

The question then arises of how decisions are made in animals such as humans that have both the implicit, direct reward-based, and the explicit, rational, planning systems (see Figure 2; Rolls, 2008b). One particular situation in which the first, implicit, system may be especially important is when rapid reactions to stimuli with reward or punishment value must be made, for then structures such as the orbitofrontal cortex may be especially important (Rolls, 2005b, in press). Another is when there may be too many factors to be taken into account easily by the explicit, rational, planning system, when the implicit system may be used to guide action. In contrast, when the implicit system continually makes errors, it would then be beneficial for the organism to switch from automatic habit, or from action–outcome goal-directed behaviour, to the explicit conscious control system which can evaluate with its long-term planning algorithms what action should be performed next. Indeed, it would be adaptive for the explicit system to regularly be assessing performance by the more automatic system, and to switch itself in to control behaviour quite frequently, as otherwise the adaptive value of having the explicit system would be less than optimal.

It may be expected that there is often a conflict between these systems, in that the first, implicit, system is able to guide behaviour particularly to obtain the greatest immediate reinforcement, whereas the explicit system can potentially enable immediate rewards to be deferred, and longer-term, multistep, plans to be formed that may be in the interests of the individual not the genes. For example, an individual might decide not to have children, but instead to devote himself or herself to being a creative individual, or to enjoying opera, et cetera. This type of conflict will occur in animals with a syntactic planning ability, that is, in humans and any other animals that have the ability to process a series of "if . . . then" stages of planning. This is a property of the human language system, and the extent to which it is a property of nonhuman primates is not yet fully clear. In any case, such conflict may be an important aspect of the operation of at least the human mind, because it is so essential for humans to correctly decide, at every moment, whether to invest in a relationship or a group that may offer long-term benefits, or whether to directly pursue immediate benefits (Rolls, 2005b, 2008b, 2011a).

Decision-Making Mechanisms in the Brain, and How They Are Influenced by Noise in the Brain

Recently a theoretical foundation for understanding decision-making in the brain has been emerging (Deco, Rolls, Albantakis, & Romo, 2012; Deco, Rolls, & Romo, 2009; Rolls, 2008b, in press; Rolls & Deco, 2010; Wang, 2002). A fundamental part of the architecture is a neural network that has positive internal feedback between its neurons, and that can fall into one of a number of states, each one of which corresponds to a decision, and consists of one winning population of neurons that is firing at a high rate and inhibits the other populations. When the decision process starts, if the inputs are relatively equal, the state that is reached is influenced by the noisy, that is random, spike timings of the firings of the neurons in the different populations. This type of noise in decision-making processes may operate at many different stages of brain processing, and may even influence the way in which decisions are influenced on different occasions between the unconscious emotional system and rational decision-making processes (Rolls, 2004, 2005a, 2007a, 2007b, 2008a, 2008b; Rolls & Deco, 2010). I emphasize that by rational I mean here "reasoned." In this way, noise in the brain may influence what behavioural responses or actions are made to emotional stimuli, including, for example, whether actions are based on activity in the emotional or reasoning brain systems.

Conclusions

In this article, I have outlined a theory of emotion. This provides an account of approach and avoidance in terms of first, associative learning to define which stimuli are goals (this involves associative learning between stimuli that are primary and secondary reinforcers), and second, of instrumental or action–outcome (i.e., action–reinforcer) association learning.

References

Berridge, K. C., Robinson, T. E., & Aldridge, J. W. (2009). Dissecting components of reward: "Liking," "wanting," and learning. *Current Opinion in Pharmacology*, *9*, 65–73.

Blaney, P. H. (1986). Affect and memory: A review. *Psychological Bulletin*, *99*, 229–246.

Cardinal, N., Parkinson, J. A., Hall, J., & Everitt, B. J. (2002). Emotion and motivation: The role of the amygdala, ventral striatum, and prefrontal cortex. *Neuroscience and Biobehavioural Reviews*, *26*, 321–352.

Carruthers, P. (1996). *Language, thought and consciousness*. Cambridge, UK: Cambridge University Press.

Darwin, C. (1872). *The expression of the emotions in man and animals* (3rd ed.). Chicago, IL: University of Chicago Press.

Deco, G., Rolls, E. T., Albantakis, L., & Romo, R. (2012). Brain mechanisms for perceptual and reward-related decision-making. *Progress in Neurobiology*. Advance online publication. doi: 10.1016/j.pneurobio.2012.01.010

Deco, G., Rolls, E. T., & Romo, R. (2009). Stochastic dynamics as a principle of brain function. *Progress in Neurobiology*, *88*, 1–16.

Dennett, D. C. (1991). *Consciousness explained*. London, UK: Penguin.

Ekman, P. (1993). Facial expression and emotion. *American Psychologist*, *48*, 384–392.

Frijda, N. H. (1986). *The emotions*. Cambridge, UK: Cambridge University Press.

Gennaro, R. J. (Ed.). (2004). *Higher order theories of consciousness*. Amsterdam, The Netherlands: John Benjamins.

Grabenhorst, F., & Rolls, E. T. (2011). Value, pleasure, and choice in the ventral prefrontal cortex. *Trends in Cognitive Sciences*, *15*, 56–67.

Lazarus, R. S. (1991). *Emotion and adaptation*. New York, NY: Oxford University Press.

Lieberman, D. A. (2000). *Learning*. Belmont, CA: Wadsworth.

Oatley, K., & Jenkins, J. M. (1996). *Understanding emotions*. Oxford, UK: Blackwell.

Pearce, J. M. (2008). *Animal learning and cognition* (3rd ed.). Hove, UK: Psychology Press.

Rolls, E. T. (1990). A theory of emotion, and its application to understanding the neural basis of emotion. *Cognition & Emotion*, *4*, 161–190.

Rolls, E. T. (1995). A theory of emotion and consciousness, and its application to understanding the neural basis of emotion. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 1091–1106). Cambridge, MA: MIT Press.

Rolls, E. T. (1997a). Brain mechanisms of vision, memory, and consciousness. In M. Ito, Y. Miyashita, & E. T. Rolls (Eds.), *Cognition, computation, and consciousness* (pp. 81–120). Oxford, UK: Oxford University Press.

Rolls, E. T. (1997b). Consciousness in neural networks? *Neural Networks*, *10*, 1227–1240.

Rolls, E. T. (1999). *The brain and emotion*. Oxford, UK: Oxford University Press.

Rolls, E. T. (2003). Consciousness absent and present: A neurophysiological exploration. *Progress in Brain Research*, *144*, 95–106.

Rolls, E. T. (2004). A higher order syntactic thought (HOST) theory of consciousness. In R. J. Gennaro (Ed.), *Higher-order theories of consciousness: An anthology* (pp. 137–172). Amsterdam, The Netherlands: John Benjamins.

Rolls, E. T. (2005a). Consciousness absent or present: A neurophysiological exploration of masking. In H. Ogmen & B. G. Breitmeyer (Eds.), *The first half second: The microgenesis and temporal dynamics of unconscious and conscious visual processes* (pp. 89–108). Cambridge, MA: MIT Press.

Rolls, E. T. (2005b). *Emotion explained*. Oxford, UK: Oxford University Press.

Rolls, E. T. (2007a). The affective neuroscience of consciousness: Higher order linguistic thoughts, dual routes to emotion and action, and consciousness. In P. Zelazo, M. Moscovitch, & E. Thompson (Eds.), *Cambridge handbook of consciousness* (pp. 831–859). Cambridge, UK: Cambridge University Press.

Rolls, E. T. (2007b). A computational neuroscience approach to consciousness. *Neural Networks*, *20*, 962–982.

Rolls, E. T. (2008a). Emotion, higher order syntactic thoughts, and consciousness. In L. Weiskrantz & M. K. Davies (Eds.), *Frontiers of consciousness* (pp. 131–167). Oxford, UK: Oxford University Press.

Rolls, E. T. (2008b). *Memory, attention, and decision-making: A unifying computational neuroscience ppproach*. Oxford, UK: Oxford University Press.

Rolls, E. T. (2009). The anterior and midcingulate cortices and reward. In B. A. Vogt (Ed.), *Cingulate neurobiology and disease* (pp. 191–206). Oxford, UK: Oxford University Press.

Rolls, E. T. (2011a). Consciousness, decision-making, and neural computation. In V. Cutsuridis, A. Hussain, & J. G. Taylor (Eds.), *Perception–action cycle: Models, algorithms and systems* (pp. 287–333). Berlin, Germany: Springer.

Rolls, E. T. (2011b). A neurobiological basis for affective feelings and aesthetics. In E. Schellekens & P. Goldie (Eds.), *The aesthetic mind: Philosophy and psychology* (pp. 116–165). Oxford, UK: Oxford University Press.

Rolls, E. T. (2012). *Neuroculture: On the implications of brain science*. Oxford, UK: Oxford University Press.

Rolls, E. T. (in press). *Emotion and decision-making explained*. Oxford, UK: Oxford University Press.

Rolls, E. T., & Deco, G. (2010). *The noisy brain: Stochastic dynamics as a principle of brain function*. Oxford, UK: Oxford University Press.

Rolls, E. T., & Grabenhorst, F. (2008). The orbitofrontal cortex and beyond: From affect to decision-making. *Progress in Neurobiology*, *86*, 216–244.

Rolls, E. T., & Stringer, S. M. (2001). A model of the interaction between mood and memory. *Network: Computation in Neural Systems*, *12*, 111–129.

Rosenthal, D. M. (1986). Two concepts of consciousness. *Philosophical Studies*, *49*, 329–359.

Rosenthal, D. M. (1990). A theory of consciousness. ZiF Technical Report no. 40, Bielefeld University, Germany: Zentrum für Interdisziplinaire Forschung.

Rosenthal, D. M. (1993). Thinking that one thinks. In M. Davies & G. W. Humphreys (Eds.), *Consciousness* (pp. 197–223). Oxford, UK: Blackwell.

Rosenthal, D. M. (2004). Varieties of higher-order theory. In R. J. Gennaro (Ed.), *Higher order theories of consciousness* (pp. 17–44). Amsterdam, The Netherlands: John Benjamins.

Rosenthal, D. M. (2005). *Consciousness and mind*. Oxford, UK: Oxford University Press.

Rushworth, M. F., Noonan, M. P., Boorman, E. D., Walton, M. E., & Behrens, T. E. (2011). Frontal cortex and reward-guided learning and decision-making. *Neuron*, *70*, 1054–1069. doi: S0896–6273(11)00395–3 [pii] 10.1016/j.neuron.2011.05.014

Scherer, K. S. (2009). The dynamic architecture of emotion: Evidence for the component process model. *Cognition & Emotion*, *23*, 1307–1351.

Wang, X. J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. *Neuron*, *36*, 955–968.