L&PS - Logic and Philosophy of Science
Vol. 1 · No. 1 · 2003

Complexity as a new framework for emotion theories

Giovanna Colombetti
School of Cognitive and Computing Sciences, University of Sussex, UK
giovc@cogs.susx.ac.uk

Abstract

In this paper I suggest that several problems in the study of emotion depend on a lack of adequate analytical tools, in particular on the tendency of viewing the organism as a modular and hierarchical system whose activity is mainly constituted by strictly sequential causal events. I argue that theories and models based on this view are inadequate to account for the complex reciprocal influences of the many ingredients that constitute emotions. Cognitive processes, feelings and bodily states are so subtly intertwined that it is not possible to determine which one "comes first" in a causal chain. The dynamical systems approach in cognitive science, I suggest, provides a more appropriate framework for the study of emotion. In particular, the notion of circular causation and collective action help depict the organism as a self-organising system in which emotion emerges as a function of its global activity. Among others, this dynamical perspective allows revising the popular notion of appraisal in a way that can dissolve some of the questions that have taunted emotion theorists thus far.

Keywords: emotion, amygdala, appraisal, circular causation, somatic marker hypothesis, A-not-B error.

1. Introduction: perennial problems in the study of emotions

In 1970 Magda Arnold listed what she saw as perennial problems in the field of emotion:

> There are some perennial problems, apparently, that still require solution: first, the question of how emotion is related to action. Does it have a dynamic component, is it inevitably connected with an instinctive impulse that provides the dynamics, or is it separable from any kind of activity, remaining either purely a mental state … or purely a physiological upset? Secondly, how is emotion aroused; directly, by way of perception which produces "connate adaptation of the nervous system" resulting in activity, including physiological changes which are then sensed; or is it the personal reaction to

> a particular situation … ; or, finally, is it the matrix of all experience and action … ? Thirdly, what is the difference between an emotion that accompanies goal-directed striving, and an emotion that interferes with it? And, finally, how are the physiological changes that go with emotion really produced? (Arnold, 1970, pp.172-173)

From Aristotle on, indeed, the proliferation of very different theories of emotion has raised more questions than answers. The relation between emotions and physiological changes, behaviour, feelings, evaluations, drives, beliefs, desires, pleasures and instincts has been accounted for in several, often incompatible ways, and the general sensation is that emotion theory is still finding its way across a maze of different explanatory frameworks.

Emotion theorists agree, at least, that most difficulties come from the fact that emotion is a complex and multifaceted phenomenon. I suggest here that, indeed, the slippery character of emotion needs new and more sophisticated analytical tools than those provided so far. In particular, I criticise good old fashioned frameworks based on *modular* and *hierarchical* perspectives of the mind, which try to explain the elicitation of emotion by positing a *strictly sequential* causal chain of mental and/or physical events.

The dynamical systems approach to the mind (Thelen and Smith, 1994; Port and van Gelder, 1995; Kelso, 1995) seems to offer more adequate conceptual tools to account for the subtle interactivity of the components of emotion. Particularly useful ones are the positive explanatory role assigned to the notion of *circular causation* and the ability to deal with complex phenomena that emerge on the *collective action* of micro-components. In cognitive science these features have suggested new ways of looking at traditional dichotomies, such as perception vs. action (Freeman, 1991; Churchland et al., 1994) or knowing vs. doing (Thelen and Smith, 1994; Thelen et al., 2001). Similarly, I believe, they can blur the dichotomy between emotion and cognition often implicit in the study of emotion, offering a framework that allows to see them as related in a complex way. The attempt to side-step hierarchical processing and to treat mental and physical processes at the same level of complex, symbiotic influences is a crucial step towards this new framework.

In what follows I develop my arguments by evaluating some models used in emotion theory. I first show that their analytical tools are inadequate to capture what goes on during the "Bechara gambling test" - an experiment designed to track the interaction between emotion and cognition in decision-making (Damasio, 1994; Churchland, 1996; Tranel et al., 2000). What happens during the experiment, I argue, is best understood as a continuous process of reciprocal causation taking place among several components; a *simple* notion of "interaction" or "feedback" is not enough to capture such complex relations. I then compare the gambling experiment with the so-called "A-not-B error" experiment, which has recently been modelled by a strongly formalised dynamic field theory (Thelen et al., 2001). I draw some analogies between the two tests, suggesting that the former could be captured by a similar dynamical model. My aim is to illustrate the kind of *explanation* offered by this dynamical theory and to argue that a similar approach should be adopted by emotion theorists. This paper thus provides ideas along which to conduct and interpret future research, rather than a detailed theory of emotion. Finally, I show how the proposed framework provides a new understanding of the notion of *appraisal* (a popular concept in emotion theories) and how this helps reformulate some

important questions, dissolving the perennial problems listed by Arnold.

2. The gambling experiment

Antonio Damasio and colleagues have been studying subjects with lesions in the ventromedial prefrontal cortex. Among them, Elliot's story resembles the one of the famous Phineas Gage. As it is known, Gage was a capable construction foreman who had a terrible accident at work; an iron bar literally traversed his brain, entering the front from the cheek and exiting from the top of the head. Despite that, Gage survived and his intellectual capabilities remained intact. However, his everyday attitudes changed dramatically and transformed him from a diligent, faithful and reliable man into an irresponsible and untrustworthy one. Damasio (1994) reports recent studies on Gage's skull according to which the accident damaged both his prefrontal cortices in the ventral and inner surfaces, preserving the external or lateral ones. The removal of a brain tumour damaged the same areas in Elliot who, as a consequence, experienced a similar change in personality. The lesions did not affect any of his cognitive performances; he kept scoring well on various tests for intelligence, learning, memory, attention, etc., but every-day observations revealed that something in him had changed.

A first striking feature was his inability to make long-term decisions. For example, when asked to set the date of a meeting, he could go on for hours considering all the possible dates, and all the possible impediments, and all the possible alternatives and relative consequences of each choice, without producing any answer. He was, roughly speaking, similar to a computer running down all its search trees without heuristics. His inability to plan was disastrous for his life. He lost his job and went through a series of divorces and sudden marriages; in general, he seemed to have become indifferent towards risky situations. In addition, Damasio noticed that his general attitude was "pleasant and intriguing, thoroughly charming but emotionally contained. He had a respectful, diplomatic composure … he was cool, detached, unperturbed even by potentially embarrassing discussion of personal events" (Damasio, 1994, pp.34-35).

These observations induced Antoine Bechara, one of Damasio's colleagues, to build the gambling experiment in order to test Elliot's emotional responses to risky situations. In this test, subjects are presented with four decks of cards; each time they turn one card from one of them, they win or lose some money (play bills, but looking like the real thing). The experimenter tells them to turn the cards (for an unspecified number of times) and try to make as much profit as possible. After several card turns (the experiment usually lasts 100 card turns, although the subjects do not know this when they start playing), controls understand that in the long-term it is better to play decks C and D because although decks A and B pay more, they also contain higher penalty cards. Unlike controls, ventromedial subjects usually show a preference for the risky decks and hence end with a loss, even after repeated testing.

Further research on skin conductance during the game provided other interesting results. During the first card turns, neither controls nor ventromedial subjects showed a skin response (1) to card selection. After about round 20,

controls started to show an *anticipatory* skin response when reaching for decks A and B (the bad decks), unlike ventromedial subjects whose response remained neutral. However, when asked, controls reported that they were making their choices randomly. At about round 50, controls continued to show anticipatory skin responses and, when asked, said that A and B seemed less favourable; once they reached the 100th round, controls were able to report which are the winning decks and still showed strong anticipatory skin responses. Unlike them, subjects like Elliot never showed any anticipatory skin response and never reported an intermediate stage of awareness (a "hunch period") in which the losing decks look unfavourable. However that both control and ventromedial subjects showed a skin response *after* turning the card (from both favourable and unfavourable decks). This is important, because it shows that ventromedial subjects do not completely lack the capacity to produce skin responses. In other words, they do not *completely* lack emotion; unlike healthy subjects, they rather lack the capacity to *anticipate* the future disadvantageous consequences of their actions (in this sense, they are more like alcohol abusers, who seem indifferent towards future situations; this is sometimes referred to as "*alkohol myopia*" and Damasio uses the term "short-sightedness" to characterise Elliot's affective condition).

3. Traditional frameworks and why they are inadequate

What does this tell us about how cognition and emotion interact to lead towards a successful gambling strategy? Let us review some explanatory frameworks.

Damasio has formulated the so-called *somatic marker hypothesis* in order to explain what goes on in the gamblers and in Elliot at the neurophysiological level (Damasio, 1994; Tranel et al., 2000). According to this hypothesis, the successful cooperativity of emotion and cognition is implemented, in the normal-functioning brain, as the *association* between the amygdala (2) and parts of the cortex related to the elaboration of complex stimuli. The association is orchestrated by the ventromedial prefrontal cortices, which thus play the role of *convergence zones* (Damasio and Damasio, 1994). In Elliot, these zones are damaged, which prevents the association between the amygdala and the other parts of the cortex. Crudely put, Elliot's thoughts do not get *marked* by the activation of the amygdala, which prevents the construction of a successful gambling strategy.

Notice that the somatic marker hypothesis implicitly assumes that, somehow, the amygdala is the "centre" of emotion, while the cortex is the "centre" of cognition. As such, it is compatible with modular and hierarchical models of emotion that have been proposed in cognitive science (by neurologists, psychologists and artificial intelligence (AI) students). In particular, it is compatible with the idea that emotions can be distinguished between *primary* and *secondary* ones (Damasio, 1994).

Many psychologists and neurologists believe that there is a set of basic, pan-cultural and evolutionarily old emotions whose seat is the limbic system (3), of which the amygdala is an important component. Although there is no agreement on how many basic emotions there are, nor on which ones they are, it is generally acknowledged that they are *automatically* elicited. Joseph

LeDoux (1996), for example, has studied how primary fear may be implemented in the brain. He suggests that the ancestral fear that makes primates and lower animals jump away from impending dangers is triggered by a "quick and dirty" pathway leading information *directly* from the sensory thalamus into the amygdala. This pathway is not very sophisticated (e.g. it may take a piece of wood for a snake), but it is very economic because it does not need the sensory information to go "up" the evolutionarily newer cortical areas to trigger a coping response. The tag "secondary", "complex" or "cognitive" is applied to more sophisticated emotions, e.g. the fear to fail in one's job. For LeDoux, secondary fear is triggered by a brain pathway that passes through the cortex before leading into the amygdala. This means that bodily states typical of fear are mediated by thoughts; in even simpler words, that *beliefs* trigger emotions.

Like in the case of the somatic marker hypothesis, drawing a distinction between primary and secondary emotions implies that behaviour depends on the interaction of two rather separate parts of the brain: a primitive and instinctual one, depending on the activity of the amygdala (something like the "sensitive" part of the soul Aristotle mentions in the *De anima*) and a more sophisticated one, depending on the cortex (Aristotle's "rational" part). In this framework, Elliot's behaviour is explained as an impairment of that interaction. By lacking primary responses (he lacks galvanic skin responses), he lacks the capacity to integrate them with cortical activity, i.e. he lacks the capacity to have secondary fear towards risky situations. His thoughts (recall that he *knows* what he should do) do not reach the amygdala, hence no fear arises.

Partitions similar to the Aristotelian one can be found in the attempts to model emotional agents through virtual architectures. For example, Aron Sloman (e.g. 2001 for a recent version) claims that the characteristic structure of emotional agents is the superimposition of two three-layered architectures. The first one is divided into perceptual, reasoning and action processors, as in traditional AI. The second one is divided into reactive, deliberative and meta-management mechanisms and should account for the traditional distinction between primary and secondary emotions, plus a third set of *tertiary* emotions. The reactive layer produces automatic actions when certain conditions are satisfied (roughly, it plays the role of the amygdala). The deliberative layer is responsible for reasoning, planning, predicting and explaining; it considers, compares and selects various possible actions, thanks to its memory capabilities (it plays the role of the cortex). The meta-management layer is in some sense self-reflective, being able to monitor and even act on the internal processes going on at the lower levels (it implements e.g. the loss of control over the deliberative layer, which for Sloman is a tertiary emotion).

Supporters of this and similar architectural approaches have referred to neurological accounts like Damasio's and LeDoux' to say that their models are *biologically valid*. Evolution, they claim, has created such levels and we have to understand how they interact in order to explain the complexity of emotion. I suggest that this way of looking at the mind is rather *driven* by compositional accounts that are much older than evolutionary theory. The distinctions between body and mind, senses and intellect and, similarly, primary and secondary emotions are deeply rooted in the history of Western thought. More than a product of evolution, they are the heritage of Aristotle's partition (e.g. Descartes in *The Passions of the Soul* distinguished basic and complex

passions, Spinoza in his Ethica tried to derive all emotions from pain, pleasure and desire, McDougall in *The Sociology of Psychology* individuated seven basic instincts, etc.). They are all somehow related to the dichotomy of passion vs. reason. They are the lenses through which neurologists look at the brain, they represent the dominant framework in AI, and they have been used to interpret the stages of evolution as a progressive tendency towards intellectualisation (it is a fact that evolution adds bits of mechanism to brains and organisms; however, what I reject is the idea that it adds separate layers, where the newer ones are "more intellectual" than the older ones. I rather see the process of intellectualisation as emergent upon the *integration* of the newer bits of mechanism with the older ones). They not only constitute the predominant epistemology in emotion theory; in most cases, they are seen as real, naturally developed distinctions. In short, I suggest that when artificial models fit neurological ones, it is not because they are biologically valid; rather, it is because all of them are products of the same conceptualisation.

Even if recent models of the mind admit a quite sophisticated and feedback-like interaction between cognition and emotion (in Sloman's agents, for example, all the levels of one architecture interact among themselves and with all the levels of the other architecture), I believe that their *modularity* and *hierarchical structure* still poses too many constraints that prevent conceiving of that interaction in the subtle way required to understand the behaviour of the gamblers. What goes on during the gambling test suggests that wise long-term decision making is a question of integrating and harmonising *over time* the activity of several micro-components, such as somatic responses, feelings, memories and drives. They reciprocally and continuously influence one another, constructing successful behaviour as these interactions unfold.

The performance required from the gamblers consists in a continuous adjustment of different *cognitive* processes: checking the effects of their choices, comparing them with previous ones, figuring out the next best move, looking for a confirmation of the efficiency of the strategy adopted, changing strategy or keeping it. As the game proceeds and as the subjects start to realise that they are using a good strategy, *feelings* and *somatic responses* also start to arise and contribute to the final choice. But, crucially, all this does not happen in distinct moments, nor does it imply a strictly sequential causal chaining between the mentioned components. Somehow, feelings and somatic responses whisper at the beginning, and insofar as the subject becomes confident, they speak louder till eventually clearly heard. Does the subject become confident that her strategy works because she starts feeling the somatic tips, or does she feel them because she has become confident? This is the old question that has tormented emotion theorists since Plato's *Euthyphro*: do we love anything because we consider it holy, or do we consider it holy because we love it?

This is, I believe, the wrong question to ask. In asking "what comes first?" we force our science and epistemology to look for a leading process, a first event that triggers all the others, and we then have to deal with the cited perennial problems. The question leads to overlooking the *complexity* of the reciprocal adjustments involved in the process. The distinction between primary and secondary (or even tertiary) emotions is an effect of such simplification. It assumes that every emotion has to be triggered, in the first place, by either a thought or a direct, reflex-like mechanism. This forced choice has been imposed in particular since the formulation of the James-Lange theory. As it is

known, James and Lange claimed that perception can directly trigger a coping reaction, without the intermediate step of emotional feeling; instead, the emotional feeling is the set of somatic and visceral responses that constitute the coping reaction (James, 1884). The cognition-emotion debate that followed has required us to decide whether or not these somatic and visceral responses need cognition in order to be triggered. The questions raised by Arnold reflect this way of inquiring.

4. Complexity as a new framework

We need a framework that avoids the wrong question. We should not ask whether the somatic response or the decision making comes first in the construction of the successful strategy, because none of them is a causally independent event. To decide which are the losing decks, the gambler has to engage in the game, pay attention to what is going on, try to remember what has happened in each trial, etc. Unless these cognitive activities are supported by somatic responses and feelings, they are not sufficient for a successful performance. The positive role played by somatic responses and feelings strictly depends, in turn, on the fact that the player is engaging in certain cognitive activities and attempting to determine consciously which are the bad decks. Somatic responses and feelings provide useful information only if continuously paired with such processes. Thus none of these elements is a causally independent process; the successful strategy emerges out of their continuous reciprocal influences.

The dynamical systems approach to the mind can help clarify this account, offering a powerful explanatory conceptual framework to settle the cognition-emotion debate. In particular, one of its most useful gifts is the positive explanatory role attributed to *circular causation*. The concept is nicely illustrated by Kelso (1995) through the so called Rayleigh-Bénard instability. Take a pan and pour some oil in it, then apply heat from below. As the temperature difference between the top and bottom of the oil layer increases, you will notice the emergence of rolling motions in the oil (also called "convection rolls"). This phenomenon is due to the *collective action* of the many molecules of which the oil in the pan is composed and, crucially, is not led by any privileged component. Applying heat to the bottom of the pan determines a pattern of organisation in the oil molecules different from the one characterising a small temperature difference between the bottom and top of the oil layer. The process is *self-organising* and its unfolding over time is continuously determined by the interactions between the emerging rolls and the molecules. The organisation of the molecules cause the rolling motions, and the rolls cause changes in the pattern of organisation of the molecules. In other words, the emergent phenomenon constrains the behaviour of its component parts.

In general, the dynamical systems approach to cognitive science has taken advantage of the notions of circular causation, collective activity and self-organisation to undermine traditional distinctions based on hierarchical approaches to the mind. For example, dichotomies like knowing vs. doing, perceiving vs. moving, internal structure vs. external expression of it (e.g. in

language), holding pre-established objectives vs. *making* decisions have all been challenged by the notion of self-organisation. The crucial intuition is *that there is no steering process in the head* that leads all the others, the so called lower ones. As Kelso (1995) briefly puts it, there is no self in the self-organising system. What is important is the continuous process of self-adjustment and adaptation to the environment. There is no (either top-down or bottom-up) message passing between distinct, independent and hierarchically organised levels, but rather *global phenomena* emergent upon the organised activity of their micro-components. In more biological and organic terms, *homeostasis* leads the behaviour of the system, not any ghost in the machine.

In this framework, emotion as well can be seen as a complex process constructed upon the interactions and local feedback of many components, none of which is the leader. Fear, for example, is not triggered by a first, neutral, non-valued, non-emotional event. What is there is a self-sustained organism whose interactivity with the environment is intrinsically *valenced* (because of its previous history). The objects that fall within its domain of interactions get emotion-laden accordingly. To value something as dangerous requires us to be already inclined to fear it; and to be inclined to fear something requires an intrinsic appreciation of its harmfulness. The cognitive (or evaluative) and the emotive moments are not separated, nor sequentially ordered. They are part and parcel of the same organic activity unfolding in time.

The most recent dynamical approaches have produced quantitative models (i.e. models that implement differential equations) simulating behaviour and psychological phenomena. This reinforces the hope that it is possible to propose a new framework powerful enough to go beyond qualitative characterisations and to inspire empirical research. To put some flesh on this, I am going to discuss Thelen et al.'s (2001) model. This will provide the occasion to compare the behaviour of the gamblers with a case of motor behaviour which has been dissected by the analytical scalpel of dynamics. Just as this dissection dissolves explanatory difficulties due to previous hierarchical and sequential conceptualisations of the explanandum, a dynamical account of the gamblers' behaviour would dissolve difficulties due to misleading conceptualisations of the cognition-emotion relationship.

5. Dynamical explanations at work

Thelen et al.'s is a dynamical systems theory of infant perseverative reaching proposed to explain what psychologists, since Piaget, call the "A-not-B error". Such an "error" occurs when infants (usually 7 to 12 months of age), after reaching for a toy hidden at a certain location A, keep reaching for that location even after the toy is moved and hidden to a second, close-by location B. The phenomenon is particularly puzzling because the infants' performance crucially depends on several factors, like e.g. the attractiveness of the object, the distance between the locations A and B, the presence of cues differentiating the locations, and the time delay between the hiding event and when the infant is admitted to search. *Any variation in these parameters can modify the children's behaviour and prevent the error.* Besides, the error has been shown to take place even when *no object* is involved and infants are asked to reach for

locations that are cued by hand waving. This has suggested Thelen and colleagues that, unlike what Piaget thought, the A-not-B error does not say anything in particular about *the formation of the concept of object*. They suggest that it is rather a *motor error*, a case of *perseverative reaching* that depends on the dynamics of the interactions between the infants' brain, body and environment. Furthermore, the error has been shown to take place in *older* children (around 2 years old) when the task is complicated by hiding the object in a sandbox. This supports the authors' idea that the error is part of a behaviour manifested at all ages, in different degrees of complexity.

In the traditional psychological literature, the effects of all those factors have been tested through many variations of the original Piaget's experiment and the results have been used to support different hypotheses about the mechanisms involved in the formation of the concept of object. For example, some have said that the A-not-B error occurs because 7-12-month-olds have a poor memory for the hiding place and poor control to inhibit previously acquired motor responses. Or, they still have to mature coordinating abilities for their movements in order to *do* what they *know* they should do. In fact, a further puzzling element during the canonical A-not-B error experiment is that children seem to look in the right direction, although children keep reaching for the wrong location, they seem to *look* at the right one.

Thelen and colleagues strongly reject this dichotomy between knowing and acting:

> We deeply disagree with the widely held assumption that knowing and acting are modular and dissociable. Indeed the cornerstone of our dynamic model is that "knowing" is perceiving, moving, and remembering as they evolve over time, and that the error can be understood simply and completely in terms of these coupled processes. (Thelen et al., 2001, p.4)

They take *seeing* as being just another form of activity and not the expression of knowledge. For them, to say that knowledge guides action implies "that there lives, in the baby's head, a creature that is smarter than the body it inhabits" (p.3). The target of their criticisms are all the explanations that posit an intellectual faculty leading all the others.

They also claim that the tendency to indulge in modular accounts is what makes all explanations of the A-not-B error (and its variations) *incomplete*; as a result, none has managed to account for the complexity of the factors involved. This, I think, is also what has happened with theories of emotion. They have focused on either cognitive, feeling or bodily states; or they have attempted to account for *all* these factors, but in a strictly sequential causal way.

Thelen et al.'s model is intended to account for *all* the variations shown by the A-not-B error. It is a dynamical theory of perseverative reaching that illustrates how the movements of the infants depend on the subtle cooperativity of various components. The differential equations used to describe the process contain variables for the behavioural alternatives of the participants (reaching, looking), for the delay between the hiding and the permission to reach, for the distance of the location, for the attraction-strength of the object and, crucially, for the motor variation induced by previous reaching trials (memory). The model proposed has been tested with simulations (illustrating the variations of the experiments) and has reproduced the performance of human subjects.

6. A dynamical revision of the notion of appraisal

What are the implications of this new framework for theories of emotion? In which sense can Thelen et al.'s model help understand emotions? What is the relation between the A-not-B and the gambling task?

I am going to try to answer these questions by referring to a popular concept in emotion theories, i.e. the concept of *appraisal*. I will show that the ways it has been used so far are unsatisfactory because they have forced it into the traditional distinctions reviewed before. In so doing, they have overlooked the most interesting part of the story, i.e. the *real processes* that the concept is supposed to stand for. The dynamical framework, by contrast, allows a finely grained analytical zoom into these processes and gives back a more comprehensive understanding. This results in greater explanatory power. It encompasses previous, partial and incomplete explanations, and it also allows looking at seemingly different processes as instances of the same one.

The notion of appraisal has been used by psychologists, neurologists and philosophers to explain how different emotional responses are triggered. Why do we run away from an approaching bear? Because we appraise (or "evaluate") it as dangerous. Why is Othello jealous? Because he appraises Desdemona as unfaithful, and so on. The debate is complicated by the fact that the notion is not only used to support the idea that beliefs or cognitions trigger emotions; some (e.g. Prinz, in press) argue that in the case of reflex-like emotions the appraisal is *somatic*, i.e. non-cognitive.

In what I think is the first dynamical theory of emotion, Dewey (1894) noticed that tags like "dangerous" or "unfaithful" are *already* emotionally valenced. To say that we run away from an approaching bear because we evaluate it as dangerous is an empty explanation unless it accounts for the concept of dangerous in the first place. Dewey's explanations appealed to evolutionary theory and he concluded that we constitute the approaching bear as dangerous already in the act of perceiving it. Decades later, Magda Arnold blamed him for using:

> … really radical expedients. In order to explain why we escape from the bear, Dewey is obliged to say not only that action comes before the emotion, but that it even comes before perception. It is the action which constitutes the object. But how can we move toward anything unless we first see it as an object, located somewhere in space? (Arnold, 1960, p.119; my emphasis).

Indeed, in her criticism Arnold put the dynamical approach in very clear terms. Moreover, she anticipated what some neurologists and cognitive scientists argue nowadays, i.e. that perception and action are not distinct processes and that they are not triggered by "neutral" representations of objects (Churchland et al., 1994).

Similarly, I suggest that emotional responses are not triggered by neutral objects appraised in a certain way at a later moment. This suggestion is supported by the dynamical approach to the brain explicitly called for by Freeman (1991). LeDoux (1996) provides many examples of brain activity that support it (despite his more traditional distinction reviewed earlier). Importantly, he explicitly refuses the idea that there is a *centre* of emotion; the

limbic system, he claims, is a theoretical construct introduced to fill in the no man's land between the innermost parts of the brain and the cortex. The emotional brain is rather the *whole brain*, and he shows this by illustrating several loops in which the amygdala is involved. Without getting into too many details, let me just hint at one of them.

The amygdala receives activation from cortical and subcortical areas and, importantly, also projects back to them, thus influencing the stimuli received from them. One could look at this action as a self-modulating process of the amygdala. Indeed, the projections of the amygdala to cortical areas are much more numerous than the ones from the cortex to the amygdala, and they are distributed in a more complex way. So, for example, let us consider what happens in the brain when a visual stimulus is supposed to reach the amygdala from the cortex. It first goes through the primary cortex, then to a secondary region, then to an area in the temporal lobe and finally to the amygdala. The amygdala projects to the temporal lobe, but also to the previous stages of the pathway. In other words, the stimulus from the cortex is continuously modulated by the action of the amygdala, in a way that would be too simplistic to describe as a feedback mechanism. The action of the amygdala over the areas that project to it is *multiple*; it influences several inter-communicating parts at once, which in turn *continuously change* the way in which the stimulus is passed from the higher to the lower levels. Hence, instead of going a straight pathway *down* across various layers, *the stimulus enters a looping process of continuous modulation*. We can say that the stimulus projecting to the amygdala is already amygdala-laden (see Hardcastle, 1999, for a defence of the idea that emotional percepts are already cortex-laden). Why does it work like that?

We can begin to see the reason if we consider that perceptual and motor activities are motivated by the amygdala itself (and by other parts to which it is highly connected). The amygdala stimulates the motor systems and, subsequently, prepares the sensory systems to receive forthcoming information. The sensory systems transmit the information back to the amygdala and the other connected parts, which re-initiates the process accordingly (Freeman, 1991). The fact that the amygdala receives amygdala-valenced sensory information thus suggests that the role of such information is to direct the attention of the amygdala to *relevant stimuli*. In other words, the amygdala mediates perception and action in a complex way. Sensory systems pick up information that is emotionally relevant (amygdala-valenced). Motor activities are both the antecedent and consequent of this activity of the sensory systems: they allow the senses to scan the environment in search of the relevant stimuli, and they are determined by what the senses pick up. The other way around, we can say that perceptual processes are both the antecedent and consequent of motor activities: they determine how the motor systems will act according to what is perceived as relevant and they pursue their information-seeking activity on the basis of where the motor systems direct them (see also Churchland et al., 1994). In sum, the amygdala motivates our actions and perceptions and filters perceptual and motor information.

As I said, this is still only a small part of the story. The amygdala also has many connections with long-term and working memory networks. In both cases, activation loops between the amygdala and other parts of the brain play the fundamental part in forming memories. Besides, it regulates the release of hormones in the bloodstream and its activity is itself regulated by the hormones

that reach the brain. It thus plays an important role in modulating brain and body activities (LeDoux, 1996).

This glimpse into the workings of the amygdala should have suggested that the organism does not perceive *neutral objects* upon which it subsequently acts, or that it subsequently appraises (why should it, if the objects were neutral?). Rather, the objects it encounters are already valenced by amygdala-induced seeking activities, selected through evolution. The organism is self-organising and the objects that enter its phenomenology contribute to this circular sustaining; besides, this happens through continuously emotionally valenced processes.

Going back to the A-not-B task and to the gambling experiment, I suggest that both are cases of motor planning driven by some valenced targets (reaching for the object, making money). Elliot's behaviour during the gambling test can, I think, also be seen as a case of perseverative reaching rather than as a dissociation between knowing and acting. While the perseverative reaching of the infants performing the A-not-B error is due to immaturity, Elliot's is due to a brain lesion. In both cases, however, the perseverative attitude depends on the lack of poise of perceptual and motor processes. If this is the case, then Elliot's impairment tells us that somatic responses play an important role in achieving and keeping such poise. It would be very interesting to measure the infants' physiological responses during the A-not-B error to track the dynamics of their arousal in the motor process. The hypothesis is that the poised integration of physical arousal within the dynamics of perception and action is the necessary ingredient to correct the perseverative attitude.

In the (healthy) players, such integration allows the correct evaluation of the situation and the related successful behaviour. Recall that subjects initially act without a plan, making a guess most of the times; as time goes by, they start appreciating the results of their choices and, thanks to the reciprocal tuning of somatic responses, feelings, memories and drives, they understand which are the good decks. The evaluation of the situation as favourable or not, as rewarding or failing, thus *co-steers* various processes from within, building up as they unfold. Similarly, the evaluative capacities of babies in the A-not-B test develop together with their motor and perceptual abilities, as they grow older. The passing of time allows the integration of various components into a successful reaching movement.

Both cases thus show the crucial role of *time* for the construction and realisation of plans for motor control. As Thelen et al. say, the A-not-B error is just a *phase* of a dynamics that belongs to all stages of life. This dynamics can be thought of as involving different stages of evaluation. Before 7 months, children do not reach for objects; their appraisal of the situation is at its minimum. Between 7 and 12, under certain conditions, they perform the A-not-B error. The appraisal of the situation seems to have reached the ability to guide a perseverative motor task, or a visual task, but not a counterfactual motor one. Crucially, these stages of appraisal depend in such a subtle way on the time scaled cooperativity of several factors, that sudden switches are pending. A small variation in the conditions of the experiment, and the appraisal unchains a different movement. And so on, similarly, for infants of more than 12 months, to 2 years, to adults. Given the temporal aspect of the gambling game, it could be possible to adapt Thelen et al.'s equations to model the gamblers' reaching and account for Elliot's different dynamics as a *transition* in his behavioural

state space.

In sum, the processes characterising the behaviour of the infants *and* the behaviour of the gamblers "are perceptual, *and* they are motor, *and* they are cognitive" (Thelen et al., 2001, p.8); to which we can now add that they are *emotional* and *evaluative*. In the same sense that cognition is the set of processes that hierarchical and modular accounts rather view as its effects, so appraisal is the set of responses that cognitive theories of emotion rather think of as its effects.

7. Conclusion: dissolving traditional problems

By way of conclusion, here is how we can dissolve the perennial problems mentioned by Magda Arnold and cited in the introduction.

The present view explains why emotion is dynamic, i.e. why, in Arnold's terminology, it is bound up with action. As we have seen, the key move is to deny that action is the output of a process in which, first, an object is perceived and, second, it is appraised. If it is required to follow a sequential order, the relation between appraisal and action will remain a contingent one and will never be able to account for the peculiar role of bodily responses in emotions. Indeed, philosophers like Anthony Kenny (1963) have exploited such causally sequential accounts to deny that bodily processes are related to emotion in any important way, and even to deny that experimental psychology can tell anything interesting about the relation between emotion and behaviour.

As far as the second point raised by Arnold is concerned (how is emotion aroused?), the notion of dynamical appraisal undermines the distinction between "perceived objects" and their "valence". Non-valenced human percepts do not exist; it is because percepts are valenced that they arouse emotions. This does not exclude that emotions are something like "personal reactions to situations"; rather, even if there might be pancultural and biologically old evaluative tendencies, the construction of appraisals depends on the history of the organism's interactions with its environment. The way in which Arnold formulates the problem is misleading because it seems to imply that physiological approaches make emotions objective and impersonal states, and that only the process of appraisal (in the traditional, modular version) makes emotion "personal".

The third point addresses the question of whether emotions are organising or disrupting. The traditional middle way proposes that emotions are sometimes disruptive and sometimes organising. Aristotle in the *Nicomachean Ethics* famously suggested we should have the right feelings at the right time towards the right people; anger, if moderated, is useful. Hume in his *Treatise* said that the calm emotions do and should lead our behaviour, while the violent ones disrupt it. Similarly, most middle-way supporters say that the disrupting aspect is directly proportional to the intensity of the emotion; the approach is thus sympathetic towards the organising role of emotions. However, it still seems to suggest that reasonable behaviour is a matter of controlling them by means of some other faculty. For the dynamical approach, to control emotions means *to nudge the whole emotional system from within*, tuning the organism-environment interactions on the basis of ongoing evaluations. There is no

independent faculty capable of disrupting or organising the organism or its behaviour; what is organised or disrupted is the dynamics of the whole being. Thus "the difference between an emotion that accompanies goal-directed striving and an emotion that interferes with it" is a difference in harmonisation. Successful goal-directed striving corresponds to the harmonised and adaptive coupling of feelings, drives and actions.

The final question (how are the physiological changes that go with emotion really produced?) remains a hard one; in dynamical terms, it challenges us to specify the details of local feedback processes and to track the action of multiple collective variables. This, of course, is where the real work remains to be done.

8. Acknowledgments

Many thanks to Andy Clark, Margaret Boden, Marco Giunti and Steve Potter for useful comments on the first draft of this paper. Thanks also to the members of the E-Intentionality Group at the COGS School of the University of Sussex for discussion of these topics.

References

Aristotle (1961). De anima. (Translated by David Ross). Oxford: Clarendon Press.
Aristotle (1980). The Nicomachean ethics. (Translated by David Ross). Oxford: Oxford University Press.

Arnold, M.B. (1960). Emotion and personality. New York: Columbia University Press.
Arnold, M.B. (1970). Perennial problems in the field of emotion. In M.B. Arnold (Ed.). Feelings and emotions. New York: Academic Press.

Churchland, P.S. (1996). Feeling reasons. In P.M. Churchland and P.S. Churchland (Eds.). On the contrary: critical essays. Cambridge MA: MIT Press.

Churchland, P.S., Ramachandran, V.S. and Sejunwski, T.J. (1994). A critique of pure vision. In C. Koch and J. Davis (Eds.). Large scale neuronal theories of the brain. Cambridge, MA: MIT Press.

Damasio, A. (1994). Descartes' error: emotion, reason and the human brain. New York: G.P. Putnam.

Damasio, A. and Damasio, H. (1994). Cortical systems for retrieval of concrete knowledge: the convergence zones framework. In C. Koch and J. Davis (Eds.). Large scale neuronal theories of the brain. Cambridge, MA: MIT Press.

Descartes, R. (1912). The passions of the soul. In R. Descartes (1912). The philosophical works. Vol. 2. (Translated by E.S. Hold and G.R.T. Ross). Oxford: Oxford University Press.

Dewey, J. (1894). The theory of emotion. In J. Dewey (1882-1898). The early works. Southern Illinois University Press.

Freeman, W. (1991). The physiology of perception. Scientific American, 264, 78-85.

Hardcastle, V. (1999). It's OK to be complicated: the case of emotion. Journal of Consciousness Studies, 6, 237-249.

Hume, D. (1978). A treatise on human nature. (Edited by L.A.Selby-Brigge, 2nd ed.). Oxford: Clarendon Press.

James, W. (1884). What is an emotion? Mind, 9, 188-205.

Kelso, J.A.S. (1995). Dynamic patterns: the self-organisation of brain and behavior. Cambridge MA: MIT Press.

Kenny, A. (1963). Action, emotion and will. London: Routledge and Kegan Paul.

LeDoux, J. E. (1996). The emotional brain. New York: Simon and Schuster.

McDougall, W. (1908). An introduction to social psychology. London: Methuen.

MacLean, P.D. (1990). The triune brain in evolution. New York: Plenum.

Plato (1914). Euthyphro. (Translated by H.N. Fowler). London: Heinemann.

Port, R. F. and van Gelder, T. (Eds.) (1995). Mind as motion. Cambridge MA: MIT Press.

Prinz, J. (in press). Emotional perception. Oxford: Oxford University Press.

Sloman, A. (2001). Varieties of affect and the CogAff architecture schema. Proceedings of the AISB'01 symposium on emotion, cognition and affective computing, York: University of York, pp.39-48.

Spinoza (2000). Ethica. (Edited and translated by G.H.R. Parkinson). Oxford: Oxford University Press.

Thelen, E. and Smith, L.B. (1994). A dynamic systems approach to the development of cognition and action. Cambridge MA: MIT Press.

Thelen, E., Schöner, G., Scheier, C. and Smith, L.B. (2001). The dynamics of embodiment: a field theory of infant perseverative reaching. Behavioral and Brain Sciences, 24 (1), pp.1-86.

Tranel, D., Bechara, A. and Damasio, A. (2000). Decision making and the somatic marker hypothesis. In M.S. Gazzaniga (ed.). The new cognitive neuroscience. Cambridge MA: MIT Press.

NOTES

(**1**) *Skin conductance* (or *skin resistance*) is traditionally used as a measure of a subject's level of emotional arousal. Changes in skin conductance are easily detected via electrodes placed on two palmar active sites (usually the subjects' fingers) and connected to a polygraph. Electrical current passes along the pathway between the electrodes; when the skin's sweat glands secrete fluid, such a passage is reduced. Measuring the skin resistance thus means measuring the change in the amount of current conducted by the skin between the electrodes. back

(**2**) The amygdala is a small almond-shaped organ close to the hippocampus, considered primarily responsible for aversive reactions (including skin conductance). See LeDoux (1996). back

(**3**) The notion of "limbic system" was introduced by MacLean (1990), who divided the brain into three parts: the reptilian brain, the emotional brain, and the mammalian brain. The limbic system refers to the emotional brain, i.e. to the area included between the thalamus and the cortex. LeDoux (1996) believes that this distinction is untenable and that the notion of limbic system is a just way of filling in the "no man's land" between the thalamus and the cortex. However, it is still a popular notion. back