

## ***Non-classical cognitive science and emotion***

---

*In former times a Raja sent for all the blind men in his capital and placed an elephant in their midst. One man felt the head of the elephant, another an ear, another a tusk, another the tuft of its tail. Asked to describe the elephant, one said that the elephant was a large pot, others that it was a winnowing fan, a ploughshare, or a besom. Thus each described the elephant as the part which he first touched, and the Raja was consumed with merriment.*

Buddhist parable<sup>1</sup>

In this final chapter I argue that the three non-classical approaches discussed in the previous chapters may be combined to produce a single integrated non-classical approach. In the second section, I illustrate this approach by reference to the interruption theory of emotion.

### ***6.1. Integrated non-classical cognitive science***

In chapter two, I described how, in the period between 1950 and 1980, cognitive scientists tended (i) to assume that the mind was a domain-general mechanism, (ii) to identify its boundary with a physical feature of the agent (either the boundary of the brain, or the boundary of the body), and (iii) to work with discrete models. In chapters three to five, I showed how, after 1980, these assumptions came to be challenged by various 'deviant' schools of thought within the cognitive science community. Evolutionary psychologists argued that the mind was composed of many domain-specific mechanisms. Proponents of the so-called 'situated' approach argued that the mind had flexible boundaries. And there were calls to model the mind in continuous terms by those enamoured of dynamical systems theory.

In this section, I try to lay bare the conceptual links between these three non-classical approaches. First, however, I want to examine the links between the assumptions of the classical approach.

#### *The disunity of classical cognitive science*

The three assumptions of domain-generality, internalism and discreteness are logically independent. It is perfectly possible to imagine a cognitive

---

<sup>1</sup> The parable of the elephant is attributed to the Buddha by the *Udana*, one of the scriptures of the Theravada or Hinayana school. This version of the story is taken from *Buddhism: An Introduction and Guide*, by Christmas Humphreys (Harmondsworth: Penguin, 1951), p.11.

scientist of a classical bent dropping any one of them and retaining the others, or relinquishing any two of them and keeping one. All permutations are possible from a strictly logical point of view. Nor do there seem to be any strong theoretical reasons of a non-logical kind for linking the three together. One is forced to the conclusion, then, that the conjunction of these three assumptions in the first decades of cognitive science was a mere historical accident.

The word 'accident', however, should not be taken to imply that there is no good explanation for the fact that the first generation of cognitive scientists subscribed to these three assumptions. It is just that the explanation should appeal to historical reasons rather than theoretical or logical ones. For example, Andrew Wells has suggested one possible historical explanation for the widespread acceptance of assumption (ii): internalism. He claims that Turing originally conceived of the finite state control and the infinite tape memory in his 'universal machine' as equivalent to the agent and the environment, respectively. According to Wells, it was only the increasing tendency of computer scientists to hardwire the control to a finite memory and package the resulting system into a single box that led Turing's original distinction to be increasingly blurred. The result was that cognitive scientists interpreted control and memory as internal components *within* the cognitive agent rather than as two qualitatively different sources of variance – the organism and the environment (Wells, 1998: 275). In other words, it was a contingent fact about the design of computing machinery that led cognitive scientists to adopt a strong commitment to internalism.

This point might be extended more generally to other aspects of classical cognitive science. The basic idea of CTM is so general that it gives no real guidance about how artificial minds should be designed, or by what criteria they should be evaluated. Yet, as a matter of fact, many classical cognitive scientists probably identified the term 'computer' far too closely with the actual machines of their day, which were built for very particular purposes (and certainly not always to provide models of the human mind, although they were later construed as such). The demands of mathematical rigour from those who wanted machines to ease the burden of doing sums meant that these machines had to be consistent, reliable, non-random, accurate, and predictable.<sup>2</sup> These aims are best met by discrete-state machines. Whether or not this explanation is the correct one remains to be seen, but this would clearly make an interesting project in the history of science.

The unity of classical cognitive science is a historical curiosity, then, that has no interesting theoretical explanation. One might be tempted to infer from this that, just as the three classical assumptions are not conceptually

---

<sup>2</sup> This point was brought home to me by Geoffrey Miller (personal communication).

linked, so also their contraries are quite independent of one another. This, however, would not be a valid inference. It might be true, but one cannot assume it on purely logical grounds. In fact, I think it is false. I think that there are good grounds for thinking of the three non-classical approaches as forming a single conceptual bundle. There are good grounds, in other words, for combining the evolutionary, situated and dynamical approaches into a single 'non-classical' approach that is far more theoretically coherent than the classical approach ever was.

### *The unity of non-classical cognitive science*

This is not to say that the non-classical approaches are equal partners. On the contrary, I think that the conceptual unity of the integrated non-classical approach I am proposing comes from taking the evolutionary approach to be primary. In other words, if you start with an evolutionary approach, you will probably also want to adopt a situated approach, and perhaps a dynamical approach too. But if you start with a situated approach you are much freer to retain the classical assumptions of domain-generality and discreteness, and if you start with a dynamical approach, you are not thereby given any guidance on the questions of domain-generality and internalism.

Here is how the evolutionary approach leads us to the other non-classical approaches. First, it leads us to think in situated terms, because natural selection will always favour those cognitive agents that can offload as much of their computational burden onto the environment as possible. Natural selection is an economiser, and computation is expensive. George Williams made a similar point in 1966 when he argued that, other things being equal, natural selection will always favour obligate adaptations over facultative ones (Williams, 1966). Obligate adaptations develop willy-nilly, while facultative ones develop one way in one set of circumstances and another way in other circumstances. Obligate adaptations are computationally cheap; the advantages of polymorphism have to be considerable before natural selection will give up the cheap alternative.

This point is reinforced by computer simulations of evolution. When artificial perceptual discrimination systems are allowed to evolve by techniques of artificial life, the end result is often a highly situated system that offloads as much computation onto the world as possible. In other words, such systems exploit very specific details of their local environment, not general features of all possible environments. For example, a system for picking out squares might simply evolve a straight-edge detector if the only shapes in its artificial environment were squares and circles. Only if it also regularly encountered other straight-edged shapes such as triangles would it need to build in extra knowledge.

*Evolution and dynamics*

The evolutionary approach also leads cognitive scientists to pay attention to the grain of our scales (are there enough states in the model?), to the details of timing, and to chaotic behaviour – all the things which, I argued, are the distinctive traits of the dynamical approach. The details of timing are particularly important to evolved cognitive agents who, unlike the systems that tend to result from traditional human design methods, must deal with a multiplicity of concerns in real time. Evolutionary psychologists are also interested, of course, in longer time-scales than those that apply to the second-by-second control of behaviour. Ontogeny and phylogeny are also inherently temporal processes, so there may be a distinctive advantage in adopting the dynamical approach when trying to understand them.

Ontogeny and phylogeny have traditionally been described in terms that are, at least implicitly, digital. During the past ten years, however, dynamical accounts of these processes have grown in popularity (Depew and Weber, 1995). There are now interesting, though still rather speculative, accounts of both ontogenetic and phylogenetic processes that attribute to them complex nonlinear phenomena such as bifurcation.

Many traits, for example, are classed as innate or acquired, or analysed into innate and acquired components. This binary opposition, however, is too crude for most explanatory purposes. Here, a dynamical approach can help to make it clear that innateness is a question of degree. The resources of dynamical systems theory can also be applied directly to the question of how innate a given trait is by providing a mathematically precise model of Waddington's epigenetic landscape (Waddington, 1940). The various factors influencing the shape of the landscape can each be represented as a different dimension in an  $n$ -dimensional state space, and the trajectories through this state space will then represent the possible paths taken by individual development or the evolution of a species. Nor is there any need, in such a model, to partition these factors on the basis of whether they are 'genetic' or 'environmental'. Such a distinction serves no useful purpose (other than assisting animal and plant breeders), and the dynamical model makes this clear by treating all factors as equivalent sources of variance.

If one adopts the evolutionary approach to begin with, then, there are good reasons why one should also take a situated and a dynamical approach. The evolutionary approach, therefore, can be the basis of an integrated non-classical approach that differs from the classical approach in taking a view of minds as:

- (i) Massively domain-specific
- (ii) Leaky
- (iii) Inherently temporal, and involving continuous as well as discrete systems

The integrated non-classical approach also differs from the classical approach in that these three tenets are much more closely linked by theoretical reasons than the assumptions of domain-generality, internalism and discreteness. The integrated non-classical approach is, in other words, truly *integrated* in a way that the classical approach is not.

#### *From possibility to actuality*

Another difference between the integrated non-classical approach and the classical approach is that, at moment of writing, the former only exists as a theoretical possibility. For several decades, almost all cognitive scientists subscribed to all three assumptions that define the classical approach. Many still do, though the number is gradually diminishing. However, there are very few cognitive scientists who reject all three classical assumptions. Many cognitive scientists today reject one, and some reject two, but hardly any challenge the classical approach on all three fronts. There are few, if any, proponents of the integrated non-classical approach I propose.

If anything, the proponents of one classical approach often seem to be more concerned to compete with proponents of the others rather than with building links between them. Exchanges between proponents of the various non-classical approaches to cognition can sometimes take on a very shrill tone, as if these approaches were mutually exclusive.

This is less true of relations between the situated and the dynamical approaches, where there have been some attempts to tie them together. Andy Clark, for example, has argued that the resources of dynamical systems theory are strongly preferable for understanding embodied, embedded agents. In particular, he claims that dynamical systems theory provides 'an explanatory framework that (1) is well suited to modelling both organismic and environmental parameters and (2) models them both in a uniform vocabulary and framework, thus facilitating an understanding of the complex interactions between the two' (Clark, 1997: 113). A framework that invokes digital components lacks these advantages.

However, this is far from being the common view among cognitive scientists working with a situated approach or those working under the

banner of dynamical systems theory. Most of the time, each of these two schools of thought operate without much reference to the other. The same is true of evolutionary psychology with respect to both the situated and dynamical approaches. One looks in vain at the bibliographies of the various books and papers published by leading evolutionary psychologists for any mention of any works by the leading proponents of situated cognition (such as Andy Clark, Francisco Varela, George Lakoff and Rodney Brooks), nor of the earlier sources on which these writers draw (such as Heidegger, Merleau-Ponty and J. J. Gibson).<sup>3</sup> Likewise, most evolutionary psychologists seem to be unaware of the burgeoning work in dynamical cognition. It seems that evolutionary psychologists have, at least up to now, adopted a very narrow conception of cognitive science, one which is exclusively classical in all respects except its commitment to the evolved nature of mind. When reading Cosmides and Tooby, or Steven Pinker, one gets the impression that cognitive science is still entirely dominated by the disembodied, digital approach that characterised the discipline in its early days.

This is particularly unfortunate for, as I have just argued, the evolutionary approach can provide the foundation for an integrated non-classical approach. Evolutionary cognitive scientists are not logically obliged to adopt a situated and a dynamical approach, but there are good theoretical reasons why they should do so. By failing to realise this, evolutionary psychologists have so far missed the opportunity to lead a revolution in cognitive science.

The explanation for this may for have more to do with social and rhetorical reasons than with any failure of imagination on the part of evolutionary psychologists. The main priority for evolutionary psychology in its early days, in the late 1980s, was to establish itself as a credible research program in its own right. At that time, evolutionary theory was perceived as rather tangential to the business of discovering mental structure by most cognitive scientists. In challenging this assumption, as well as arguing for the domain-specificity of many mental processes, evolutionary psychologists had enough on their hands. Tying the evolutionary approach too closely to the situated and dynamical approaches, which were also perceived as 'young Turks' by the older generation of cognitive scientists, would have made it even harder to establish credibility (and thus to get jobs and funding). At a time when departments of cognitive science were run by those of a classical bent, retaining the classical assumptions of internalism and discreteness was probably necessary for evolutionary psychology for purely tactical reasons.<sup>4</sup>

---

<sup>3</sup> Geoffrey Miller is somewhat of an exception. He is the only evolutionary psychologist to make abundant references to J. J. Gibson and ecological psychology.

<sup>4</sup> This point was suggested to me by Geoffrey Miller (personal communication).

Things are somewhat different now, though. The evolutionary approach is certainly not accepted by all – indeed, there are still sections of the cognitive science community that are vehemently opposed to it – but it has at least established itself as a major player. It has already acquired many of the status symbols of a thriving research program: scholarly journals that are peer-reviewed, learned societies, textbooks, academic positions, undergraduate courses, and annual conferences. Evolutionary psychology has also achieved something that other schools of thought in cognitive science never have, not even the classical approach: widespread popular appeal. This last feature, however, may actually have *hindered* the process of gaining academic acceptance (perhaps because classical cognitive scientists are jealous).

Now that evolutionary cognitive science is no longer an embryonic discipline but a rapidly maturing research program, the time would seem ripe for it to question its adherence to the classical assumptions of internalism and discreteness. If evolutionary psychologists have enough nerve and imagination, they could transform the integrated non-classical approach I have proposed here from a mere theoretical possibility into a historical reality.

#### *Pluralism and compatibility*

How would such an integrated non-classical cognitive science relate to its classical forbear? Confrontation would not necessarily be the order of the day. As I have argued in the previous chapters, the cognitive scientist is not necessarily forced to make a stark choice when faced with the issues of domain-generality and discreteness (internalism is a different matter, as I will argue shortly). It is quite possible to build hybrid models that combine domain-general mechanisms with domain-specific ones (as Fodor did in *The Modularity of Mind*), or that combine discrete-state machines with analogue systems (e.g. Tani, 1999: 152). Thus, instead of talking about a 'paradigm shift' or a 'scientific revolution', then, the relationship between the integrated non-classical approach and the classical one might therefore be described more appropriately as one of assimilation, akin to the way that the theory of general relativity is sometimes described as absorbing Newton's theory of gravity as a special case applying only to a limited domain. This way of talking might help to avoid perpetuating the spurious confrontations that have dogged cognitive science in recent years. Andy Clark has argued persuasively for an eclectic approach to the mind in which we need to combine a variety of explanatory styles, including both the componential explanations typical of the classical approach and the dynamical explanations of more recent years. He suggests that progress in cognitive science will consist of 'adding new tools' to the explanatory tool-kit, rather than abandoning

those we already have. After all, if *the mind* were so simple that a single approach could unlock all its secrets, we would be so simple that we couldn't understand the theory!<sup>5</sup>

Still, if calls for pluralism are not to descend into sloppy thinking, we must be precise about the nature of the compatibility between the classical approach and the various non-classical approaches. Compatibility comes in various kinds.<sup>6</sup> In particular, the claim that two approaches, A and B, are 'compatible' could be construed in at least three different ways:

- (i) A and B provide different kinds of explanation
- (ii) A and B are mutually inter-translatable
- (iii) A and B explain different phenomena

I have argued that the evolutionary approach is compatible with the classical approach in sense (i), that the situated approach is compatible with the classical approach in sense (ii), and that the dynamical approach is compatible with the classical approach in sense (iii). I will now spell these compatibility claims out in more detail.

The classical approach and the evolutionary approach provide different *kinds* of explanation. Specifically, the former is fundamentally concerned with providing *design* explanations (how minds work), while the latter is concerned with *functional-historical* explanations (why minds work the way they do). Now, there may be ways of deriving constraints on the former kind of explanation from the latter; this is precisely what Cosmides and Toobey suggest when they propose that cognitive models should always be evaluated according to the evolvability criterion (see section 3.1). However, this possibility has not yet been conclusively demonstrated, as I showed by reference to the debate about domain-specificity. However, even if this were eventually to be proven to be the case, it would not call into question the general point that providing synchronic hypotheses about mental structure and providing diachronic hypotheses about the origins of such structures are different kinds of explanatory project. The links between them, if any, are empirical, not conceptual; we can only derive constraints about mental design from functional explanations by enriching the latter with a whole set of empirical assumptions about the adaptive problems posed by particular environments for particular lineages of organism.

---

<sup>5</sup> Adapted from a phrase quoted by Andy Clark (Clark, 1997: 175). Clark states that he was unable to trace the originator of this remark.

<sup>6</sup> I owe this point to Michael Wheeler (personal communication).

The classical approach and the situated approach are mutually inter-translatable. As I argued in chapter four, whenever the proponent of the situated approach locates the input-output boundary of a particular computational system at some point outside the organism, the classical cognitive scientist can always re-describe this system in traditional internalist terms. Proponents of the situated approach can argue against this internalist move on the grounds that it is dogmatic and is of dubious explanatory value, but they cannot rule it out on purely logical grounds. The leaky mind hypothesis, when stated baldly, is the contradictory of the internalist view, but when it comes to applying these two approaches to real-world systems, there are no purely logical grounds for preferring one over the other. This logical compatibility, however, is not that exciting.

Finally, the classical approach and the dynamical approach are compatible in sense (iii): they explain different phenomena (and sometimes different aspects of the same phenomenon, which amounts to the same thing). I argued in chapter five that, insofar as continuous models can be taken as a separate class, they are better suited than discrete models to explaining 'low-level' processes such as pattern-matching and limb-coordination. Discrete models, on the other hand, are better suited to modelling 'high-level' processes such as forward planning. Hence the cognitive scientist is not forced to choose between an exclusively discrete approach and an exclusively continuous one. She can construct hybrid models that use both digital and analogue components. The same point applies to the debate about domain-generality. It is possible to conceive of minds that employ both domain-general mechanisms and domain-specific ones. This, indeed, was the main thrust of Fodor's proposal in *The Modularity of Mind* (Fodor, 1983).

Calls for a kind of *super-integrated* cognitive science, combining the insights of the integrated non-classical approach and the classical one, must be careful to distinguish between these different kinds of compatibility. Otherwise, they risk leading us to blur the questions raised by the non-classical approaches, and thus to obscure their importance. It is hard to argue with calls for broad-mindedness and pluralism, but if such calls are to amount to anything more than the politically-correct view that 'everyone must have prizes', they must tempt us into thinking that there are no real disagreements. Methodological pluralism is not an end in itself; it is simply a way of clearing aside the false oppositions so that we may concentrate on the genuine ones. The ultimate aim of cognitive science, after all, should not be to provide a cosy umbrella under which those of any persuasion can take shelter, but to answer the questions about how minds work and why they work the way they do.

## **6.2. An integrated non-classical approach to emotion**

In section 3.2, I put forward a theory of emotion based on a proposal by Herbert Simon. According to this theory, emotions are defined in functional terms as interruption mechanisms. An interruption mechanism is one that can interrupt ongoing activity and temporarily take control of behaviour in the service of survival or reproductive goals. In section 4.2, I developed the interruption theory by showing how some of the computational burden it required could be offloaded from the brain onto the body and the environment. In section 5.2, I asked whether further refinements to the theory could be made by paying attention to the continuous features of emotion as well as the discrete ones.

The interruption theory can serve as an example of the integrated non-classical approach that I proposed in the previous section. It has all the hallmarks that define such an approach: domain-specificity, leakage, and continuity. Furthermore, the evolutionary approach is primary. The view of emotions as interruption mechanisms is based on the functional accounts of emotion provided by evolutionary psychology. Thus I started with evolutionary considerations, and used the situated and dynamical approaches serve to round out the theory.

The interruption theory, as I have sketched it out in this thesis, needs much more conceptual refinement. However, the bare bones are at least clear. The theory is already capable of generating specific hypotheses about the design of emotional systems. Such hypotheses could be tested by implementing these designs in artificial agents. Indeed, without this vital step, the interruption theory would remain of limited use to cognitive science.

#### *Classical versus non-classical approaches to emotion*

The interruption theory is clearly non-classical in flavour, but this does not mean that it is incompatible with classical models of emotion. In the previous section, I outlined various ways in which theories can be compatible: providing different kinds of explanation, being mutually-intertranslatable, and explaining different things. Is the interruption theory compatible with classical appraisal theory in any (or all) of these different ways?

##### *(i) Kinds of explanation*

The interruption theory provides a very different *kind* of explanation to that provided by the classical account of emotion. Appraisal theory and the propositional attitude theory both explain *how* emotions can be representational. The interruption theory takes this as given, and draws on evolutionary considerations to explain what emotions are

representations of. Emotions, it claims, are representations of changes that are relevant to the achievement of some biological (i.e. survival-related or reproduction-related) goal. To be more precise, it is the function of emotions to represent such changes to the organism so that behaviour can be modified accordingly.

This general account needs to be fleshed out for each individual emotion. For each emotion, we need to say which biological goal the emotion is designed to serve. This provides a criterion for individuating emotional mechanisms: one goal, one mechanism. For example, fear is the emotion designed to serve the goal of avoiding potential physical injury (trauma), whether as a result of being attacked by another organism, or as a result of some process in the nonliving environment such as an avalanche. The function of fear is thus to represent this danger to the organism so that ongoing activity can be interrupted, when necessary, to avoid the danger.

The interruption theory complements the classical approach because the functional account it provides offers a new way of distinguishing between cognition and emotion. The classical approach had a rather conservative aim; to show how emotions could be seen as thoroughly representational, and thus count as true mental phenomena according to CTM. Appraisal theory and the propositional attitude theory achieved this aim by construing emotions as judgements. But this success was purchased at the price of destroying the only widely accepted account of how emotions differed from thoughts. Prior to appraisal theory and the propositional attitude theory, most psychologists had accepted Hume's view that passion was to be distinguished from reason by reference to the concept of representation; reason was representational, while passion was not. The classical approach rejects this way of making the distinction between cognition and emotion, but offers no other way to make it. The interruption theory does. If emotions are interruption mechanisms, then, if cognition is a distinct kind of process, it must be the kind of process that cannot interrupt any other.

(ii) *Inter-translatability*

When the interruption theory is supplemented with considerations drawn from the situated approach, emotions can be seen as processes occurring in a system whose boundaries are not co-terminous with any single feature of the organism. On this view, although it is sometimes useful to view emotions as entirely neural processes, at other times it is useful to see them as processes that leak out of the brain into the rest of the body and even into the external world. Humans, in particular, have found ways

of offloading the computational burden involved in deciding when to interrupt ongoing activity onto parts of their environment. Other animals succeed in doing this too, though to a lesser extent.

From a strictly logical point of view, there is nothing to prevent the classical cognitive scientist from re-describing the examples of embodied and embedded emotions in terms of a purely neural process that has complex feedback loops with bodily and environmental processes. The input-output boundary can always be moved back to the borders of the brain. Yet this type of compatibility claim is purely scholastic. There is more of an argument to be had here when one compares the two approaches with regard to explanatory fruitfulness rather than mere logical consistency. The design-based methodology of cognitive science means that the proper way to assess a research program is not via post-hoc logical assimilation to a given theory, but via practical research. In other words, the mere fact that two approaches are mutually intertranslatable should not lead us to overlook the possibility that one approach might lead consistently to much better working models.

(iii) *Explanatory domains*

If dynamical considerations can enrich the interruption theory by leading to a greater attention to the temporal features of emotion, this does not necessarily lead to any incompatibility with the classical approach. The classical models of emotion may not have been sufficiently sensitive to the details of timing, but nor did they deny the importance of such details. They simply ignored them. In this case, the interruption theory can be seen as supplementing the classical approach to emotion by providing explanations of different kinds of emotional phenomena. There is no real conflict here.

*Domain-general*

The fact that the interruption theory is compatible with the classical approach to emotion in all these ways should not tempt us into thinking that there are no areas of disagreement at all. There does seem to be at least one major point of disagreement – the issue of domain-general. In the classical models, such as the OCC model, all emotional stimuli are processed by a single emotion-generating mechanism. In the interruption theory, however, each emotion is implemented by a distinct mechanism that attends only to the kind of input that is relevant to that emotion.

However, it would be false to infer from this that the interruption theory is incompatible with the classical approach on this point. Although the actual models of emotion developed by appraisal theorists have all been domain-general, this need not have been the case. The domain-general

appraisal-type models is a historical accident, not a theoretical necessity. There is nothing *inherently* domain-general about appraisal theory. The basic idea that emotions are judgements arrived at by attending to particular features of external (and possibly internal) stimuli is compatible with both a domain-general and a domain-specific theory of emotional mechanisms. Thus it is not quite correct to say that the interruption theory is opposed to the classical approach on this point. It is more accurate to say that the interruption theory is opposed to most of the models designed by appraisal theorists so far. Besides, taking a domain-specific view of emotion does not imply that one is also committed to a domain-specific view of cognition. It is possible to conceive of a hybrid model of the mind in which emotional processes are domain-specific while cognitive processes are domain-general.

#### *An integrated theory of emotion*

The interruption theory can be seen then, not just as an integrated *non-classical* approach to emotion, but as an integrated approach *simpliciter*. To call it 'non-classical' implies a wholesale rejection of the ideas of domain-generality, internalism and discreteness, taken as a single package. The interruption theory, however, is much more sophisticated than this. As I have just argued, it is compatible (in various different ways) with many aspects of the classical approach to emotion, and many aspects of the classical approach to cognition. To describe the interruption theory as 'non-classical' would be to obscure this compatibility.

This is all well and good. But since understanding emotions, at least from a cognitive perspective, entails building emotional machines, the litmus test of the interruption theory will not rest on such compatibility claims but on whether or not it leads us to better machine models of emotion. If it does, then this would not only vindicate the interruption theory itself, but would also support my proposal for an integrated approach to cognitive science in general.