What Buddhism Taught Cognitive Science about Self, Mind and Brain

Asaf Federman
Institute of Advance Study & Department of Psychology
Warwick University
a.federman@warwick.ac.uk

Abstract

In the past twenty years, new optimism about the relevance of Buddhism to cognitive science has been expressed by a number of established researchers. In this article I ask what are the conceptual roots of this optimism, and which forms of development it inspired, with particular focus on selfhood, embodiment and meditation. The latter contains three distinct points of contact that are also reviewed: the introduction of first person methods, neuroscientific research of meditation, and using meditation in psychotherapy. I argue that the dialogue between Buddhism and cognitive science is part of a bigger concern that accompanies late modernity since the 19th century regarding the gap between first and third person accounts of reality. In particular it taps on a growing discontent with the Cartesian outlook on the self and its place in the world. However, while Buddhism and cognitive science both reject a similar notion of substantial selfhood, what they offer in return is different. It is often overlooked that in Buddhism fact is interwoven with value, while in science they are still further apart. This makes the claims about the compatibility of the two systems somewhat naive, and explains why recently the «dialogue» takes the form of neuroscientific research of meditation: work that hardly changes or challenges the foundations of science.

Keywords: Cognitive sciences; Buddhism; Neurosciences; Neurophenomenology; Embodiment.

Resumen. Lo que el budismo enseñó a las ciencias cognitivas sobre el yo, la mente y el cerebro

En los últimos veinte años, un número de investigadores reconocidos han expresado un nuevo optimismo acerca de la relevancia del budismo en relación con la ciencia cognitiva. En este artículo, me pregunto cuáles son las raíces conceptuales de este optimismo, y en qué formas de desarrollo se inspira, con especial atención a la individualidad, la encarnación (embodiment) y la meditación. Este último contiene tres puntos distintos de contacto que sirven también de revisión: la introducción de los métodos de primera persona, la investigación neurocientífica de la meditación y la meditación en la psicoterapia. Sostengo que el diálogo entre el budismo y la ciencia cognitiva es parte de una preocupación mayor que acompaña a la modernidad tardía desde el siglo XIX, con respecto a la diferencia entre las descripciones de la realidad en primera y tercera persona. En particular se fundamenta en un creciente descontento con la perspectiva cartesiana sobre el yo y su lugar en el mundo. Sin embargo, mientras que la ciencia cognitiva y el budismo coinciden en recha-
If there is any religion that would cope with modern scientific needs it would be Buddhism.

(Not) Albert Einstein

On the Rhetoric of Buddhism and Cognitive Science

The words in the motto above are taken from a longer quote that is attributed to Albert Einstein but are most probably not his. They appear in a number of popular publications, most of which are on the Internet (e.g., Niimi, 2002), in a way that begs the question: why authors virally reproduce it? A simple answer would be that it rubberstamps the scientific potential of Buddhism by attributing it to the most celebrated scientific genius of the 20th century. But in most cases, the statement is not reiterated by Buddhists who seek to convince a Western skeptic, but by authors who work within a Western cultural context. Whether said by Einstein or not, this statement actually captures the gist of a fundamental historical truth about the arrival of Buddhism to the West: Buddhism has been perceived as an ally of science from as early as the 19th century, and continues to be so today.

In 1893, the Buddhist scholar and reformer Anagarika Dharmapala spoke to an American audience about Buddhism which is “free from theology, priestcraft, rituals, ceremonies, dogmas, heavens, hells” and is “a scientific religion... which is in harmony with geology, astrology, radioactivity and reality”. (in McMahan, 2004: 906). The historian David McMahan analyzes in detail why this rhetoric—which was obviously inaccurate as to the actual practice of Bud-

---

1. For a more details about the origin of the quote see (Ryan, 2007b) and (Ryan, 2007a).
dhism in Asia— has been created in the first place. One of his suggestions is that Buddhism provided a promise for remedy to a crisis of faith in Victorian society: it was naively perceived as a system that provided both spiritual solace and scientific rationality, thus bridging the rapidly opening gap between science and Christianity. By the beginning of the 21st century, much of the early naïveté about the prospects of the unification of Buddhism and science has been abandoned, although the above mentioned motto reminds us that some yearning is perhaps still maintained in the public’s imagination.

In the past twenty years, new optimism about the relevance of Buddhism to cognitive science has been seriously considered by a growing number of established researchers. This article asks what are the conceptual roots of this optimism, and which forms of developments it inspires. In particular, it looks at how «the self» arose as a philosophical focal point for Buddhism and cognitive science; at how (and whether) «embodiment» arose as a Buddhist-inspired theoretical alternative to classical cognitivism; and at the place of meditation in psychological and neuroscientific research. All this is preceded by a review of the existing typologies of the evolving relations between Buddhism and science.

**Mapping Exercises**

Alongside the evolving interaction between Buddhism and cognitive science there have been a number of attempts to map the process itself. These maps address both the actual and possible relationships between the two systems and are helpful for evaluating the direction of development. It is almost impossible to approach such a mapping exercise without acknowledging the typology of relations between religion and science that was suggested by the historian of religion Ian Barbour (1990). Barbour identified four categories of relationship: conflict, independence, dialogue and integration, which represent different types of perceived relations between science and religion.

Conflict is the view that science contradicts religion, and that they are mutually exclusive. This view has recently regained some attention with the publication of *The God Delusion* (Dawkins, 2006), but had been vigorously stated before (Russell, 1930; Russell, 1935). Independence is the view that science and religion exclusively cover different domains that do not overlap. Stephen Jay Gould is probably the most known advocate of this view:

> The magisterium of science covers the empirical realm: what the Universe is made of (fact) and why does it work in this way (theory). The magisterium of religion extends over questions of ultimate meaning and moral value. These two magisteria do not overlap, nor do they encompass all inquiry. (Gould, 1999: 6)

Independence is the view that science and religion exclusively cover different domains that do not overlap. Stephen Jay Gould is probably the most known advocate of this view:

Dialogue is the view that although religion and science are different, it is possible to compare them and to demonstrate that they are analogous in one way or another (Barbour, 2002: 350). Integration takes dialogue further to claim that analogies actually mean similarities. This may result in establishing one system that covers both religious beliefs and scientific theories, or it may result in integrating both to create a single entity (ibid.).

Buddhist scholars pointed out that most of the interaction between Buddhism and science happens within Barbour’s dialogue category, and suggested more nuanced distinctions between sub-categories. Richard Payne identified three apologetic attitudes to dialogue (Payne, 2002b). The first identifies one aspect of Buddhism and shows how it is conducive to science; the second claims that Buddhism agrees with scientific methods; the third claims science now discovers truths that had been always known in Buddhism. According to Payne the rhetoric that accompanies these attitudes covers up a common wish to promote Buddhism as a potential correction to science, assuming the superiority of the former. It is interesting to note that most authors who write about how Buddhism is conducive to cognitive science, including some of the researchers mentioned in footnote 1, run the risk of being interpreted as apologetic, as if trying to justify the superiority of Buddhism. But according to Payne, there can be other more constructive ways of dialoging. In particular he mentions naturalizing Buddhist phenomenology, critically examining it by the methods of cognitive science, and introducing the Buddhist model of cognition as a testable hypothesis. Such a model, he suggests, may provide an alternative to a Cartesian model in cognitive science (Payne, 2002a).

Drawing on Barbour’s work, José Cabezón (2003) suggests another typology of relations between Buddhism and science. His typology distinguishes between three categories: Conflict/ambivalence; Identity/similarity; and Complementarity. The identity/similarity category overlaps with Payne’s category of apologetic dialogue. Like Payne, Cabezón identifies claims about similarities in content (Buddhism claims things which are also claimed by science) and similarities in method (Buddhist knowledge is achieved by rational or empirical investigation, just like scientific knowledge). In both cases there are good reasons to treat these claims with critical caution. Claiming that cognitive science has independently discovered what was known in Buddhism for a long time may be true but does not go beyond simply stating a remarkable coincidence. Unless careful research manages to explain why and how these coincidences occurred, and in what historical conditions, the mere observation means very little. In addition, in many cases what is stated to be identical (e.g., karma and evolution), turns out to be false under scrutiny (karma is an ethically-based metaphysics, evolution is a biological theory). Regarding similarities in method there is a risk of imposing modern categories on pre-modern religious system. Although Buddhist scholars point to Buddhism being empirical, rational and non-dogmatic (Jayatilleke, 1963; Jayatilleke, 1971; Hoffman,
1987) these do not amount to a fully blown scientific epistemology. As the Buddhist scholar Stephen Evans has recently argued, this interpretation of Buddhist epistemology is questionable (Evans, 2007; Evans, 2008; Evans, 2009).

The most promising area of interaction comes therefore under complementarity. While advocates of similarity mainly endeavor to eradicate the differences between Buddhism and science, those who work under the complementarity assumption acknowledge the differences and seek to bridge over them. Cabezón reviews the work of Thurman, Goleman, deCharms, Wallace and Wilber and highlights this rhetoric:

Science deals with matter, Buddhism with mind. Science is the hardware, Buddhism the software. Science is rationalist, Buddhism experiential. Science is quantitative, Buddhist qualitative. Science is conventional, Buddhism contemplative. Science advances us materially, Buddhism advances us spiritually. (Cabezón, 2003: 50)

Goleman, for example, suggests that the similarities in content are superficial, and that the most interesting interaction between the Buddhist understanding of the mind and scientific psychology should exists in relation to their methods, which are significantly different (Goleman & Thurman, 1991: 100). Different methods bring about different complementary outlooks on human development, and thus combining them would lead to richer and fuller understanding of the human condition.

Payne and Cabezón’s typologies, with their discourse analysis pointers, are helpful for guiding our attention to pitfalls in interfacing Buddhism and science. Without ruling out the potential for genuine mutual enrichment, the typologies remind us that many statements about «a dialogue» are in fact apologetic claims in disguise. In this light, the following three sections are dedicated to examining the grounds on which Buddhism and cognitive science meet. They focus on the self, the embodied approach to cognition, and meditation.

The Self

Perhaps the most celebrated link between Buddhism and cognitive science is their seemingly similar perceptions of self. The Buddhist not-self doctrine (Sanskrit: anātman / Pāli: anatta) is widely acknowledged as the cornerstone of Buddhist philosophy, and what makes it both unique and challenging. If by «self» one means an eternal entity, like a soul, which exists at the centre of being and that defines our true and ultimate identity, most cognitive scientists would be happy to adopt a philosophy that denies it. It is not surprising that many publications that deal with the philosophical interaction between Buddhist and cognitive science dedicate significant chapters to this topic (e.g., Hayward, 1987; Varela, et al., 1991; Watson, et al., 1992; Watson, et al., 1999; Wallace, 2003).
However, one remarkable thing is the effort through which authors need to go in order to demonstrate that the content under the headings «no self in cognitivism» and «no self in Buddhism» can fuel a meaningful discussion. In fact, when confronting the details, most of them cannot avoid admitting the discrepancies:

I envisage that under normal conditions the four stages [of creating the ‘I’ though perception]... are less sequentially distinct than the Abhidhamma system may suggest. (Lancaster, 1997: 134 slightly abridged)

Although it is recognized that Ordinary Man’s views [about the self] are inborn and useful, in some varieties of Buddhist discourse there still seem to be some moral opprobrium clinging to these ‘errors’... (Galin, 2003: 132)

For cognitive scientists mental processes, as opposed to contents, are not observable, while for Buddhists these may be the objects of meditation. (Watson, et al., 1992: 14)

These lines are cited here not to say that Buddhism and cognitive science are in complete disagreement, but that sweeping statements about similarities can be misleading. A statement like «the views of self in contemporary Western discourses in philosophy and science do no seem inimical to that of Buddhism» (Watson, 1999: 37) may be true but is logically weak. It means that on the face of it, scientists would not object to the image of illusory, disintegrated and non-substantial self that is found in most Buddhist traditions. It would be more difficult, however, to establish a real connection between the specifics of each side’s perspectives. To take two examples, the epistemological importance of holding no views about the existence of the self (Thanissaro-Bhikkhu, 1996) has no counterpart in cognitive science, and the detailed cognitive-philosophical account of self as a layered brain-dependent process that evolved for the sake of self-monitoring (Metzinger, 2003) cannot be found in Buddhism.

The literature on «not-self» in Buddhism and cognitive science is usually not dedicated to showing how the details of the Buddhist world-view directly contributes to science, or vise versa. In a more gentle way, it usually «bends» the two traditions toward each other to explain one’s philosophy with the language of the other. A typical chapter on the subject would describe one philosophy first, and then explain how the other may relate to it. As much as any cross-cultural study of concepts, the outcome is usually intellectually fascinating, but in most cases there are no immediate implications for the scientific study of cognition. On the other hand, such exercises in comparative philosophy are important for naturalizing Buddhist views, and weeding out those perceptions which clearly do not stand the test of empirical cognitive science.

A rather neglected article by the psychologist Brian Lancaster (1997) uses the Abhidhamma model of perception as a solution to the binding problem. As this is one of the finest examples of using Buddhist philosophy in cognitive science, and a rare one too, I will briefly summarize its main argument.
Lancaster takes the general idea of interfacing Buddhism and cognitive science one step further by calling upon a particular Buddhist theory—that of the Pāli Abhidhamma—and by applying it to a particular issue in cognitive science, namely, the function of memory within the «binding problem». The binding problem in cognitive science is defined in the following way: «How does the information initially processed by a multitude of independent modular systems become integrated into coherent representations for perception, memory, and action?» (Revonsuo & Newman, 1999). Put simply, Lancaster answers that there is a semblance of coherent representations, but they are not, in fact, representing any real existences of a coherent self. He suggests an operational model to explain how this may happen, especially in relations to how information is integrated into coherent representation in memory.

In this model perception is a process by which preliminary sensory analysis activates related memory models. The self, or the ‘I’, is one of those models, and is activated in all cases that lead to explicit memory which will be available for conscious recall. Lancaster suggests that after an item has been perceptually identified, a «self model» is attached to it. He calls this attachment «I»-tag (p. 131), because it tags, as it were, the encoded memory with relation to the self-model that «owns» it.

«I» tagging is therefore crucial for explicit memory, because it contextualizes the memory in relation to the most active model («I»), thus enabling a quick recall. On the other hand implicit memory (a type of memory in which the individual’s behaviour is influenced by previous experiences without being conscious of them) is explained in this model as memory without «I»-tagging (p.133). In these cases the perceptual process is incomplete, and though identification is achieved, there is no linking between the percept and the self-model. So far, this is one model in cognitive science that may or may not be verified by empirical research. But its inspiration came directly from the model of perception suggested in the Abhidhamma.

Lancaster suggests that «I»-tagging is similar to what is called in the Abhidhamma «slight» (paritta) object in the conscious process (vīthicitta).3 «I»—Tagging happens rather late in the perceptual process, only after the percept has been identified in relation with other stored memories. This means that with any process that reaches this stage, the self model is reinforced, and possibly amended. Although Lancaster does not mention this, this theory may explain why a short masked stimulus influences behaviour while being inaccessible to explicit recall— it was not «I» tagged and therefore was not incorporated into the self-model. Self, therefore, is understood in this theory as an ever arising, ever changing model that constantly amends itself in relation to new interpreted sensory input. It is the very model that links the other models, to create the impression of unified agency. There is no observer-self to

3. For more details about the Buddhist model see p. 159 and onward in Narada & Bodhi (2000).
which experiences are represented, but a model of ownership that is maintained from moment to moment. In Lancaster’s words:

«I» is not preexisting entity awaiting—as it were—the opportunity to know the percept; rather it arises as a contextually determined feature in the overall responses ensuing from stimulation of the sense organ. (Lancaster, 1997: 137)

This point, as much as the overall analysis of cognition, not only resembles the Abhidhamma analysis of the cognition, but also is consciously influenced by it. Lancaster dedicates a section of the article to a generally accurate description of the Buddhist position (Lancaster, 1997: 125-9), and another to discussing the merit of a mature dialogue with it (p. 133-8). To his credit, the Buddhist position is not simply taken at face value. In relation to the Buddhist momentariness doctrine, for example, it is even rejected (p. 134). The Buddhist description of the cognitive process identifies several consecutive and distinct stages, and even suggests that what seems parallel (like hearing and seeing at the same time) is in fact a result of serial processing. It is generally accepted in cognitive science that it is the other way around: cognition depends on many parallel functions rather than a line of sequential events. Even what seems like a linear process (logical reasoning would be an example) is a result of many parallel processes. To use an expression by the philosopher Daniel Dennett conscious human minds are «more-or-less serial virtual machines implemented—in inefficiently—on the parallel hardware that evolution has provided for us». (Dennett, 1993a: 218). Lancaster rejects this Abhidhammic outlook on the seriality of cognition, but sees a potential in the overall Buddhist perspective on the binding problem.

This attitude to Buddhist philosophy is a perfect example for Payne’s constructive dialogue category. The Buddhist view inspires a cognitive model, which in turn yields new hypotheses that can be empirically tested (for example the link between self perception and memory recall). Whatever does not fit within scientific knowledge (momentariness, serial processing) is left out. This of course sets a very clear hierarchical relation between the two participants in the dialogue. Buddhism can inspire, but the final word is that of science.

We may now turn to a more fundamental question about the understanding of the self in Buddhism and cognitive science. Despite some apparent similarities, the ethical and soteriological nature of the Buddhist project is often undermined in contemporary cognitive science. The Abhidhamma

4. Things are in fact more complicated. For example, in the Pali Abhidhamma, each mental event (citta) of the serial cognitive process is accompanied by many cognitive factors (cetasikā) that arise simultaneously with it. In any case, the model is significantly different from models in cognitive science.

5. In the case of parallel versus serial processing there is some room for charitable speculation. It may be that the Abhidhamma model does not describe the relations between underlying cognitive processes (serial) and their conscious manifestation (parallel). It may describe the relations between everyday phenomenology, which involves parallel experiences, and meditative phenomenology, which is more likely to bring about sequential experiences.
is part of a larger project of re-conditioning the mind to see that there is no coherent or singular self. But in sharp contrast to cognitive science, this is not the end product of an empirical project but a starting point of an ethical project. Nevertheless, the idea that a stable and continuing self is a semblance looks surprisingly similar in both sides of the dialogue. I suggest that what triggers the interest on both sides is a conceptual structure of rejecting a particular kind of self. These rejections, which are directed towards the Cartesian and the Brahmanical perceptions of self, create a gap that needs to be filled. Although the act of rejecting, the process of divorcing from a previous concept of substantial selfhood, is similar in both Buddhism and cognitive science, the actual content that fills the gap is different.

Acknowledging this —that Buddhism and cognitive science are situated on similar ends of paradigm shifts— can explain the difficulties that arise with bringing together the two systems’ views of selfhood. Both systems are united in rejecting the idea of an eternal immaterial soul that is separated from the body, and yet controls it from within. But this is more or less where the similarities end. Buddhism rejects a kind of self (ātman) which is eternal, blissful, and identical with the creative force of the universe (Brahman). It identifies the attachment to such a self as a source of misery, and thus provides logical considerations (philosophy) and practical exercises (meditation, morality) as antidotes. The prime motivation is therefore ethical and is part of a larger soteriological system. Cognitive science, on the other hand, lacks both the ethical motivation and the practical application. Its rejection of the Cartesian self is a result of scientific theorizing that emerges from experimental work. Descartes’ Error, as the neurologist Antonio Damasio puts it, was not ethical but ontological: reason and emotion are not separated faculties, and the mind is not incorporeal (Damasio, 1994).

There is, however, another level of complication here. In Buddhism the ethical dimension is always within a hand-reach distance from the factual assertions. Upholding a self-theory is both morally wrong (it brings about continuous existence in misery) and factually wrong (such self cannot be found). This connection between fact and value is of course common to other pre-modern systems and, if to believe Alasdair MacIntyre, was lost with modernism (MacIntyre, 1984). This may begin to explain why cognitive scientists usually look at the «factual» aspect of the Buddhist perspective (e.g., the «models» in the Abhidhamma), and do not hurry to adopt the otherwise inseparable soteriological perspective.

### Embodiment

The history of the embodied approach in cognitive science (sometimes misleadingly called Embodied Cognitive Science) is woven with Buddhist links and references. The embodied approach emerged from the criticism of the

6. I have demonstrated it in some length in my PhD thesis (Federman, 2008).
first ‘classical’ phase in cognitive science which focused on computation and disembodied Artificial Intelligence. Although the first phase of cognitive science re-introduced the mind as an object of research (following a period of mind-denying Radical Behaviourism), many researchers worked with the assumption that a computational-mechanical explanation would be sufficient for explaining and replicating intelligence. Gradually, however, critics have pointed out that there were empirical loopholes and theoretical problems in this approach. On the philosophical level it seemed to have preserved something of the Cartesian model: the mind was seen as a device of logical rationality and was perceived as being isolated and situated against a world that needs to be internally represented. A number of empirical results and theoretical considerations indicated that a sound theory of cognition and its replication in artificial intelligence could not be fully achieved within this classical cognitive framework. There were indications that perception, language, thought and consciousness are fundamentally embedded in living bodies and in their relations with their environments.

Although the conditions for moving into a second phase in cognitive science matured through debates within the scientific community, early articulations of both criticism and the alternatives were raised in connection with Buddhism. The idea that the mind is a living and embodied entity, not a disembodied reasoning mechanism, has been expressed in the Buddhist inspired *Shifting Worlds Changing Minds* (Hayward, 1987), and more so in *The Embodied Mind* (Varela, et al., 1991). Hayward, Varela, Thomson and Rosch all drew on existing discussions in cognitive science and endeavored to link them to aspects of Buddhist philosophy and practice. *The Embodied Mind* became a classic in the philosophy of cognitive science and is widely cited.

The question that concerns us here is the actual Buddhist contribution to the development of the embodied approach that is suggested in *The Embodied Mind*. The authors call for a revolution in cognitive science and bring forward a radical critique of the state of affairs. Their criticism is directed at the assumed division between an independently existing «external» world of objects and events, and their «internal» representation in the symbolic computational environment of the brain or mind. They suggest that both scientific findings and Buddhist thought challenge this idea and suggest an alternative.

The alternative consists of an «enactive» account in which mind is embodied, not detached, and the «world» is being created and is not pregiven. This

7. Main players included the Computation Theory of Mind (CTM) and Fodor’s Language of Thought (LOT) hypothesis. The success of this type of AI research is usually illustrated by the creation of chess-playing computers, and many other content-specific problem solving devices. Failures of this approach are usually in the area of language and perception. For more details see Horst, Clark, and Bechtel (Bechtel, et al., 1998; Clark, 2001; Horst, 2005).

8. Early critics included the psychologist James Gibson, the philosopher Hubert Dreyfus and to some extent John Searle. (see Dreyfus, 1979; Gibson, 1979; Searle, 1980; and for review: Bechtel, et al., 1998; Clark, 1999; Clark, 2001; Horst, 2005).
insight, they suggest, cannot be fully achieved without exercising «mindfulness», a practice directly absorbed from Buddhism, which will cause the practitioner «to experience what one’s mind is doing as it does it» (Varela, et al., 1991: 23). In other words, mindfulness meditation is the missing phenomenological link in cognitive science; it is a method for exploring and knowing what human experience is. This conception, that Buddhist mindfulness meditation is a valid scientific tool for investigating experience, appears also in Shifting Worlds Changing Minds (Hayward, 1987: 192-194), and is echoed in later discussions (Rosch, 1997), and more recently in the writing of Alan Wallace.⁹

Is it true that the «revolution» suggested in The Embodied Mind entailed a Buddhist input? Looking backwards on this publication, it is clear that the embodied approach in cognitive science was part of a bigger movement that was promoted by thinkers who had no affinity with Buddhist thought. Almost simultaneously with the Embodied Mind, Daniel Dennett published his Consciousness Explained that contains similar critique of the Cartesian residue in classical cognitive science and suggests an alternative (Dennett, 1993a). In fact, it contains a very embodied or «enactivist» account of consciousness.¹⁰ Varela et al indeed drive home the urgent need to include phenomenology, and the shortfalls of the phenomenological tradition in Western philosophy. But this played only a minor role in exposing the problems of classical cognitivism, and the re-framing of the philosophical landscape. In other words, the development of the embodied mind paradigm in cognitive science does not owe as much as one may wish to Buddhism. As far as I can tell, the «revolution» of The Embodied Mind did not lead to cognitive scientists radically changing their research methods, nor did it led to replacing materialism with mindfulness-based phenomenology. Perhaps it is better to see the changes as a reform of the scientific problem space (Dennett, 1993b): including the body and the environment when accounting for cognition. In the next sub-section I expand on a particular call for a paradigm shift in cognitive science through the method of neurophenomenology.

**Meditation research**

Meditation research includes several different projects that can be sorted under three headings. First, mindfulness meditation inspires the development of neurophenomenology. Second, neurological research investigates the physiological correlates of meditation. Third, cognitive research assesses the applicability and function of meditation in clinical settings.

---

⁹. For example, «What’s been hidden is meditation’s role as a precision tool for exploring consciousness and the universe scientifically—that is, using empirical methods similar to those intrinsic to the scientific method (Wallace, 2009: 1).

¹⁰. Having read The Embodied Mind Dennett contemplates the possibility of his being an enactivist without realizing it (Dennett, 1993b).
Neurophenomenology: Changing paradigms in cognitive research?

Classical cognitivism did not forsake the older «behaviouristic» aspiration to objective experimentation in psychology, but rejected the idea that behaviour can be fully explained by studying only environmental conditions. Instead of studying behaviour as if there were no cognitive states that related to it, researchers began to be interested in the states of the internal «machinery». It is important to note that the materialistic and mechanistic framework of behaviourism survived this transition. There was nothing in the new paradigm that suggested non-mechanical understanding of human behaviour. However, it did offer a new mechanical account of human behaviour that did not exclude the «internal» states of the «thinking machine».

Critics of classical cognitivism pointed out that the internal states accounted for are still far removed from the experience of mental states. Being materialistic, they say, cognitive science is too focused on third-person descriptions of cognition and the brain, and ignores the very fact that human beings seem to have private phenomenological worlds that include thoughts, emotions, wishes, sensations, perceptions and so on. The problem of explaining how a physical system, like the brain, gives rise to phenomenal experience is described by philosophers as the Hard Problem (Chalmers, 1995). One solution to this problem was brought forwards by Francisco Varela, one of the authors of The Embodied Mind, who was explicitly influenced by what he understood to be a Buddhist practice of mindfulness (Varela, 1996).

Varela’s suggestion, which he called neurophenomenology, was to take first-person experience as irreducible to anything else, and to develop a rigorous method for its exploration. Buddhist meditation is the prime inspiration of the idea of a rigorous method for exploring the phenomenal world (Varela, et al., 1991: 23-31). Neurophenomenology seeks to marry this specialized first-person investigation with research in cognitive science. At the heart of this suggestion lies the potential of utilizing trained attention to improve observation and report of first person experiences. Another aspect is creating an agreed formalized language for describing whatever is revealed by this type of attention (Varela, 1996).

What was the impact of this suggestion on cognitive science? It would be hard to say that neurophenomenology has created a revolution in the study of conscious experience, although it definitely inspired a number of interesting research studies (Lutz, 2002; Lutz & Thompson, 2003; Cosmelli & Thompson, 2007; Lutz, 2007; Christoffa, et al., 2009). It seems that neurophenomenology contains two aspects that are, to some extent, independent. First, it puts forward the realization that without accounting for subjective experiences cognitive science would be held back. Cognitive science, it is suggested, should address them directly, from the first-person perspective. The second aspect of neurophenomenology is an ontological claim about the irreducib-
ity of these states. Both aspects work fine within the framework of the Hard Problem, and indeed it received a rather positive response from Chalmers himself (Chalmers, 1997). Chalmers does remain skeptical, however, about the accuracy of mindfulness as an experimental tool because «the mere act of attention to one’s experience transforms that experience» (Chalmers, 1997: 44). Unless one argues that attention is somehow outside experience, one must accept that different kinds or levels of attention stand for different experiences, and at a final account there is no way of being mindful without inducing some changes to mood, clarity, concentration, thoughts or emotions. In other words, attention training itself changes the quality of experience.

In the context of classical cognitive science, neurophenomenology looks radical, because it shies away from the scientific tendency to focus on third-person data, primarily in the form of behavioural outputs and physiological measurements. But it is not exclusive in suggesting that subjective experience should not be excluded. Dennett, who insists that the Hard Problem is counterproductive and not real (Dennett, 1996), nonetheless suggests that science should take into account first-person data. He proposes to approach the phenomenal reports of subjects through a method that he calls «heterophenomenology» (Dennett, 2003). On the conceptual level this is the exact opposite of neurophenomenology, because it is about the experience of the other (hetero), not of one’s own experience. Dennett proposes to treat the introspective reports of others as if they were fiction, and suspend judgment (potentially forever) on whether they are really about conscious states. Introspective reports in this case are taken to be expressions of beliefs about subjective experience, not reports of anything that have special ontological status. Although this seems to directly contradict the working assumption of neurophenomenology, which takes subjective experience as fundamentally real, Tjeerd van de Laar has recently shown that in practice there is no different between the two approaches (van de Laar, 2008). The heterophenomenologist could always describe the data generated by participants’ reports in terms of their sincere beliefs, correlating with third-person data from, for example, EEG or fMRI measurements, with, or without, taking seriously the ontology of phenomenal facts suggested by these reports. The difference between hetero—and neurophenomenologies is more a matter of preferred ontology and is a question that cannot be settled on the basis of empirical research.

Neurophenomenology was indeed part of a shift towards a more phenomenologically-aware research in cognitive science. It’s Buddhist inspired ontology, however, did not emerge as an accepted alternative to the scientific materialism.

A final word about neurophenomenology should be said in regard to the training of attention. The assumption here is that it is possible, with training, to establish mindfulness that observes experience without changing it. I have mentioned above that Chalmers objects to the idea of bias-free mindfulness, but this does not mean that training participants should be ruled out altogether. The idea that some training is required to improve partici-
pants’ reports of their experiences is even accepted by the materialist Dennett. In neurophenomenology, however, the proposal is more radical than that. Here, experience is considered irreducible. Mindfulness is said to have privileged access to this realm and can therefore make real discoveries. Learning how to meditate is taken to be a process that clears biases and prevents distortions. Subjective reports in this case are accounts of something real, not just utterances of subjective beliefs.

It would be interesting to investigate to what extent this assertion applies in Buddhism. Is mindfulness meditation taken to be an empirical tool? The apologetic attitudes mentioned in the beginning of this article come to hand. The claims that Buddhist practice is scientific or empirical is a descendent of the «scientific Buddhism» discourse of Anagarika Dharmapala (McMahan, 2004) and is found elsewhere on the Buddhist side of the dialogue (Jayatilleke, 1971; Hoffman, 1987). As I already mentioned it has been recently challenged in a series of articles by the Buddhist scholar Stephen Evans (Evans, 2007; Evans, 2009).

In the case of «Mindfulness meditation», one should be reminded that in Buddhism it has always been part of a bigger project of transforming the mind. Even contemporary Buddhist Scholars, like the scientists-monk Matthieu Ricard, discuss meditation training as something that helps to «examine in depth the nature of the mind» in order to «liberate ourselves from the yoke of the ego and ignorance» (Ricard, 2003: 263). This is another reminder of how ethics and ontology are connected in Buddhism. Indeed re-uniting them within the scientific discourse would entail a revolution in our perception of truth, value and virtue. There is little doubt that this has not yet happened.

Neuroscientific research on meditation

Without aiming to challenge and change the entire scientific project, meditation research uses meditation as an aid for exploring the physiological correlates of certain types of cognition, and for achieving better understanding of what meditation is. This is by far the most proliferate field of investigation, especially since the maturation of two conditions: progress and availability of brain imaging technologies, and the presence of experienced meditators in the West. It is estimated that there have been more than 80 research publications on Buddhist meditation in neuroscience in the past 10 years, and more than 300 when including psychology and behavioural sciences in the search.

11. «For instance, it [training] might in some circumstances heighten the powers of subjects to articulate or otherwise manifest their subjectivity to investigators». (Dennett, 2003: 21)
12. Varela et al ask: «Can mindfulness/awareness meditation be considered a kind of experimentation that makes discoveries about the nature and behavior of mind...?» (Varela et al., 1991: 31). A positive reply follows.
13. Estimation is based on searching ISI Web of Knowledge database (www.isiknowledge.com accessed 17 Sept. 2009) with the term «meditation», limited by the appearance of one of the following words in the Topic category: Buddh*, Vipassana, Mindfulness, Shamatha,
The conceptual framework of this type of research is not complicated. It basically involves collecting and analyzing physiological and behavioural data from meditating participants before, during or after meditation sessions, and comparing them with control groups. In some cases phenomenological data is collected as well, in order to enrich results with meditators’ first-person perspectives. This does not mean, however, that all researchers work under the ontological assumptions of neurophenomenology. On the contrary, some influential studies are designed without measures of self-reporting, all the more so without references to an irreducibility of conscious experiences (Lazar et al., 2005; JhA et al., 2007). Other studies do refer to phenomenological data, although these references are used to justify further objective measurements (Lutz et al., 2004).

Studying meditation with neuroscientific methods takes part in a mainstream research project in cognitive science. Researchers use meditators as a valuable resource of otherwise hard-to-get data. Adept meditators devote many hours —sometimes a lifetime— to developing a narrow cognitive repertoire. This fact in itself creates an opportunity for investigating whether these practices induce long lasting neural changes, and may contribute to identifying neural correlates of certain self-induces cognitions. By no means does meditation have a privileged contribution to this type of neuroscientific interest. Similar studies on other kinds of training, in particular juggling and navigation, have famously provided similar insights on brain plasticity (Maguire et al., 2000; Draganski et al., 2004). Meditation research joins them in showing how self-induced mental states can change functional measurement (Lutz et al., 2004) and lead to long-term structural changes in the brain (Lazar et al., 2005; Davidson & Lutz, 2008; Luders et al., 2009). The unique angle of these findings is the fact that, unlike juggling balls or driving cabs in London, «self-induced mental states» do not involve observable behaviours.14

Because meditation involves first-person experiences with little or no behavioural indicators there is little doubt that verbal descriptions are necessary for evaluating what participants are actually doing. The need to describe meditative practices in cognitive-behavioural terms led researchers to collaborate with Buddhist scholars in developing common descriptive standards. The need for a common operational definition has been stated recently by the

---

14. This too may not be entirely unique. There are indications that actual and imagined movements possess similar neural substrates (Lotze et al., 1999).
neuroscientists Antoine Lutz and Richard Davidson and the Buddhist scholar John Dunne (Lutz et al., 2007). By reviewing three Buddhist practices that have received scientific attention, and by referring to traditional Tibetan taxonomies, they describe the cognitive differences between three types of meditative practices: focused attention, open presence, and non-referential compassion. Here, neuroscience joins forces with scholarship in the search for a better understanding of what meditation is. When this end is achieved, each type of meditation may be used to investigate its neural correlates thus providing a richer description of traditional taxonomies, and better understanding of neural structures.

Meditation research begins to naturalize traditional practices and describe them with physiological and functional terminology. Amishi Jha (2007), for example, draws on existing neuroscientific distinctions between dorsal and ventral subsystems of attention, and examine how they may be involved in different kinds of mindfulness practices. In her study she compared inexperienced participants who underwent an eight-week Mindfulness-Based Stress Reduction course (MBSR, see below) and experienced practitioners who underwent one-month intensive mindfulness retreat. She found that both groups improved attentional functions that are associated with the dorsal system. (This means that both groups improved top-down attention control: the ability to deliberately direct attention to objects of choice.) She also found that experienced participants who underwent one-month mindfulness retreat significantly improved their ability for attention readiness and attention to exogenous stimulus detection comparing with the first group. These results (if replicated) may indicate that intensive mindfulness retreats and weekly mindfulness course induce different attentional skills, regardless of their common titles as ‘mindfulness’ trainings. Further on, this may also feed back into debates within Buddhist communities about the different types of concentrative awareness, and into therapeutic debates about the efficacy of meditative trainings.

**Psychotherapeutic research on meditation**

In the past 20 years there has been an unprecedented increase in empirical research on meditation-as-therapy. This research typically seeks to establish whether mindfulness meditation is beneficial for treating or preventing certain conditions (e.g., stress, depression, eating disorders, cancer & addiction, for review and meta-analysis see Baer, 2003; Grossman et al., 2004). Most of it can be seen as stemming from the work in the Stress Reduction Clinic in the University of Massachusetts in which Jon Kabat-Zin and colleagues developed the Mindfulness-Based Stress Reduction programme (MBSR). This is an eight-week programme that includes 2.5 hours weekly sessions and daily home practice of about an hour, in which participants learn to develop «mindfulness». In MBSR «mindfulness» is described as «paying attention in a particular way: on purpose, in the present moment, and non-
judgmentally» (Kabat-Zinn, 1994: 4). While clinical results are considered to be promising, researchers are still investigating the underlying cognitive functioning of the training.15

Mindfulness-Based Cognitive Therapy (MBCT) is similar to MBSR but contains additional components of cognitive therapy for depression. It has been shown to be effective in preventing relapse in patients who had three or more episodes of major depression before the intervention (Teasdale et al., 2000; Segal et al., 2002). The efficacy of this preventive intervention is linked to an information-processing theory of depressive relapse that was developed independently of any Buddhist theory by John Teasdale and Philip Barnard (Barnard & Teasdale, 1991). A central aspect of this model is the recognition that certain automatic thought processes that are reinforced in a depression episode, maintain the episode, and reinforce negative cognitive models that increase the likelihood of a proceeding episode. Thus, the potential for relapse increases with each episode. The problem is that cognitive models that are developed in depression tend to be themselves depressogenic, thus causing further depression. In other words, over thinking about depression, or «feeling bad» about it, reinforces the depression itself.

In 1995 Teasdale and colleagues identified MBSR as potential intervention that may be able to break the vicious cognitive cycle of depression relapse. They indicate that focusing the attention on particular objects (like the breath) use the resources that are otherwise «necessary to support depression-enhancing or depression-maintaining processing cycles». (Teasdale et al., 1995). In this way the «central engine» of depression weakens. In addition, depression is associated with a cognitive tendency to entertain ineffective but uncontrollable problem-solving strategies, which are bound to fail and to reinforce negative self-perceptions. The principle of «acknowledgement and acceptance» of mental states in MBSR comes handy in dismantling the efficacy of these ineffective cognitions. Instead of trying to solve depression, patients learn to de-centre from it. Mindfulness practice goes beyond the cognitive therapeutic principle of «coping and controlling» and introduce «non-doing» as a preferred mode. Finally, the emphasis on relating to thoughts and emotions as momentary impersonal phenomena, rather than realities «there is pain», instead of «I am in pain» weakens the need for urgent action that is identified in Teasedale’s cognitive model as contributor for the depressive cycle.

A new preventive therapy package was developed from these considerations. Initially, Attentional Control Training (ACT) combined cognitive therapy and training in attention control, but was subsequently replaced by Mindfulness Based Cognitive Therapy (MBCT), which incorporated most of the original MBSR training programme with additional cognitive therapy

15. The most recent research project to ask how mindfulness works in preventing relapse of major depression is «Preventing depressive relapse/recurrence in NHS settings through mindfulness-based cognitive therapy (MBCT) – acronym PREVENT». See University of Exeter Mood Disorder Centre http://centres.exeter.ac.uk/mood/research.php (last accessed 18 Sept. 2009) and http://www.hta.ac.uk/project/1924.asp (last accessed 14 Jan 2011).
elements for depression (Segal et al., 2002). Here is an interesting instance of scientific work enhanced by Buddhist input. The clinical success of MBCT in relapse prevention (Teasdale et al., 2000; Segal et al., 2002) validates Teasedale’s earlier cognitive model (Teasdale et al., 1995) and helps to identify more clearly the cognitive risk factors in major depressions relapse.

There is, however, something problematic in this heroic conclusion. To what extent is MBCT Buddhist? It is difficult to deny that it was inspired by Buddhist practice, but in many respects it has already completed a long process of integration into a Western and modern discourse. Some Buddhists scholars criticize the definitions of mindfulness as paying attention on purpose, in the present moment, and nonjudgmentally (e.g., Wallace, 2006; Dunne, 2008). Even traditional sources seem to diverge from that. The descriptions of mindfulness in the Satipatthana-sutta, for example, prescribe a set of cognitive frameworks, or «judgments», through which the meditator learns to interpret reality. Although the meditator learns to disengage from the phenomenal world, he does endeavor to develop a particular perspective on it. This should come as no surprise when we are reminded that in Buddhism the distinction between fact and value is not as sharp as in modern science. For Buddhists mindfulness helps the meditator to see the world in a particular way: as devoid of selfhood, as transient, and as unsatisfactory. It is not an exercise in acceptance, but in self-induced change. Moreover, although it is concerned with «suffering», it is not directly concerned with what clinicians call depression, and is not satisfied with «normality» as its solution. But perhaps the fact that Buddhist meditation centers in Europe and the US offer retreats on MBSR and MBCT tell us that the living Buddhist traditions in the West are quite receptive to change.

Conclusion

The interaction between Buddhism and cognitive science has two major aspects. The first is a challenge to the Western scientific world-view, and more specifically to certain assumptions with classical cognitive science. The dialogue here has been part of a long history of «scientific Buddhism» although it certainly does not confine itself to claims of similarities between Buddhism and science. Nevertheless, something of the original apologetic nature remains intact when certain authors bring forward Buddhism as a remedy or justification for advocated change in science. The idea that classical cognitivism is an inadequate framework for explaining intelligent behaviour is now part of the widely accepted embodied paradigm. But to what extent Buddhism contributed to this paradigm shift? Two radical suggestions inspired by Buddhism

16. «...finally he [the meditator] learns to be aware of dharmas (plural). This has been rendered as ‘his thoughts’. But the dharmas that the text spells out are in fact the teachings of the Buddha, such as the four noble truths. The meditator moves from thinking about those teachings to thinking with them: he learns (to use an anachronistic metaphor) to see the world through Buddhist spectacles». (Gombrich, 1996: 35-6).
—that phenomenology is irreducible, and that mindfulness is an unbiased research tool— did not enter mainstream scientific thought. The milder suggestions that the computational-representational assumption is insufficient, and that a more ecological approach to cognition is required, developed largely independently of Buddhist input.

It is perhaps better to look at the published work on this subject from a different angle. The Buddhist inspired «revolution» was part of a more general concern about modernity and science, and in particular a growing discontent with the Cartesian world-view that pervaded many quarters of scientific thinking. This was not purely a matter of scientific accuracy (although there were precisely theoretical problems with empirical results as well) but a concern that materialistic science fails to explain the human condition in a meaningful way. To some extent Buddhism had the same function here as in turn-of-the-century Victorian imagination: a perceived bridge over the gap between human experience and the image portrayed by science. Although this has rarely been acknowledged, if at all, what enables this function is the fact-value fabric of Buddhism. From the scientific point of view, Buddhism brings extra value, and from the ethical point of view it brings (apparent) empirical credibility.

Interesting observations about the contribution of Buddhist thought to cognitive science always involve acknowledgment of the difference between the two systems. Exercises in comparative philosophy expose substantial differences between Buddhist and scientific metaphysics, motivations, and therefore their «problem spaces». They expose, for example, that the approach to selfhood in Buddhism is embedded within a wider ethical and soteriological concern, while cognitive science is motivated by curiosity and usually remains silent about ethics. As of today, the ethical dimension in science is confined to «ethical committees» which are certainly not interested in the liberation of the scientist’s mind. The idea of «liberation» is confined to political thought and some strands of psychotherapy. If there has ever been a project of re-marrying fact and value in Western thought it has definitely not matured.

The second aspect of the interaction between Buddhism and cognitive science is more limited in scope, but has already brought promising immediate results. Meditation research makes use of data from adept mediators to further our understanding of both meditation and the neurological or cognitive structures associated with it. Data from neuroimaging research on meditation supports the re-introduction of brain plasticity in neuroscience. There is nothing conceptually challenging in this research, it simply makes use of otherwise difficult to obtain data from unique subjects. Another strand of research, however, tries to naturalize meditation practices and explain in cognitive terms what meditators do. In many respects, this is an exercise in translation, and as such it forces more nuanced distinctions, more accurate terminologies, and in general, richer descriptions of ancient practices. Undoubtedly, new interpretations and insight into the different practices will join the long-standing taxonomical literature on meditation within the Buddhist tradition.
References


LOTZE, Martin; MONTOYA, Pedro; ERB, Michael; HÜLSMANN, Ernst; FLOR, Herta; KLOSE, Uwe; BIRBAUMER, Niels and GRODD, Wolfgang (1999). «Activation of Cortical and Cerebellar Motor Areas During Executed and Imagined Hand Movements: An FMRI Study». Journal of Cognitive Neuroscience 11 (5), 491-501.


