

Hume and the enactive approach to mind

Tom Froese

t.froese@gmail.com

Centre for Computational Neuroscience and Robotics (CCNR)

Centre for Research in Cognitive Science (COGS)

University of Sussex, Brighton, UK

Abstract. An important part of David Hume's work is his attempt to put the natural sciences on a firmer foundation by introducing the scientific method into the study of human nature. This investigation resulted in a novel understanding of the mind, which in turn informed Hume's critical evaluation of the scope and limits of the scientific method as such. However, while these latter reflections continue to influence today's philosophy of science, his theory of mind is nowadays mainly of interest in terms of philosophical scholarship. This paper aims to show that, even though Hume's recognition in the cognitive sciences has so far been limited, there is an opportunity to reevaluate his work in the context of more recent scientific developments. In particular, it is argued that we can gain a better understanding of his overall philosophy by tracing the ongoing establishment of the enactive approach. In return, this novel interpretation of Hume's 'science of man' is used as the basis for a consideration of the current and future status of the cognitive sciences.

Keywords. Hume; cognitive science; enactive; life; mind; human nature

Note: This is the penultimate draft (September 2008) of a paper which is forthcoming in the journal of Phenomenology and the Cognitive Sciences.

1. Introduction

David Hume (1711-76) is generally considered to be one of the most important British philosophers. In fact, some aspects of his work continue to be influential even outside of scholarly circles, for example in the philosophy of science (Rosenberg 1993). However, even though he attempted an “experimental” investigation of the mind inspired by the success of the natural sciences (*T.* Intro. 7), his work¹, like that of most philosophers, has so far been hardly acknowledged in the cognitive sciences. The general aim of this paper is therefore to provide some helpful pointers for the establishment of a mutually informing relationship between contemporary Hume studies and recent advances in the scientific study of mind and cognition.

In addition, the ongoing development of alternatives to this computationalist mainstream provides us with an opportunity to reevaluate the relevance of Hume’s philosophical work. It will be argued that the cognitive sciences have only recently been developing in a direction that enables them to formulate a more coherent and comprehensive interpretation of Hume’s theory of mind than has previously been possible. This is already the case with the emergence of embodied dynamicist cognitive science (e.g. Clark 1997; van Gelder 1998; Beer 2000; Wheeler 2005), but especially so with the ongoing development of the enactive paradigm (e.g. Varela, Thompson & Rosch 1991; Weber & Varela 2002; Thompson 2004; Di Paolo 2005; Thompson 2007; Di Paolo, Rohde & De Jaegher, in press).

The establishment of this mutual connection is therefore not only of interest for scholarly reasons, but also because it provides the background for a philosophical discussion of the various issues involved in the ongoing paradigm shift within the cognitive sciences. As such, the paper will only be marginally concerned with an evaluation of the vast literature of Hume studies. Instead, it offers an original interpretation of Hume’s work in the light of the most recent scientific developments, especially those of enactive cognitive science, and on this basis will raise some novel issues as well as provide some speculations about future directions of research.

1.1 Hume’s “science of man”

Hume is traditionally regarded as the last philosopher of the British empiricist tradition, and is generally acknowledged to be one of the most important philosophers ever to have written in the English language. His perceived significance has customarily been derived from the skeptical challenge he posed to the assumptions of Cartesian and rationalist philosophy. For example, Hume challenged the claim that cognition is essentially a form of abstract reasoning taking place in some disembodied realm of mental substance. Instead he considered cognition to be a natural phenomenon arising from a combination of perceptions, habits, passions, and certain associative principles of the mind. Thus, one of the ways in which Hume establishes a purely naturalistic conception of human nature is by a thorough subordination of

¹ This paper will draw mainly on Hume’s *A Treatise of Human Nature: Being an attempt to introduce the experimental method of reasoning into moral subjects* (1739-40), *An Abstract of ... A Treatise of Human Nature* (1739-40), *An Enquiry Concerning Human Understanding* (1748), and his posthumously published *Dialogues Concerning Natural Religion* (1779). Quotations taken from these works will be abbreviated by *T.*, *Abs.*, *EHU.*, and *DNR.*, respectively.

reason to feeling and instinct (Kemp Smith 1905). Moreover, the scope of his naturalization is not limited to reason but extends to human nature as a whole:

Hume's theory sees every aspect of human life as naturalistically explicable. It places man squarely within the scientifically intelligible world of nature, and thus conflicts with the traditional conception of a detached rational subject. (Stroud 1977, p. 13)

The methodology which Hume proposed in order to achieve this aim was to study the mind in the form of a "science of man", namely by applying the same kind of "experimental philosophy" which had already been immensely successful in advancing the science of nature (*T.* Intro. 7). Especially influential in this regard was the experimental method developed by Newton (Buckle 2001, pp. 83-90). Thus, although Hume's specific relationship with Newtonian science is still a matter of scholarly debate (cf. Schliesser 2007), he can generally be regarded as one of the forefathers of the modern cognitive sciences (e.g. Fodor 2003, p. 2; van Gelder 1998; Garrett 1997, p. 9; Biro 1993; Varela, Thompson & Rosch 1991, p. 129-130).

Hume's attempt to replace rationalist philosophy with an experimental study of the mind based on "experience and observation" (*T.* Intro. 7) was largely motivated by his concern with the "present imperfect condition of the sciences" (*T.* Intro. 2). This situation was especially problematic due to the fundamental explanatory gaps which Descartes had created. Thus, an essential aspect of Hume's project was to question the entrenched assumptions of the Cartesian tradition, namely the "boundary between the human and the animal, and between the mental and the physical" (Buckle 2001, p. 234).

One of the problems of assuming a fundamental distinction between human and animal intelligence is the lack of explanatory parsimony since separate theories are needed to account for each of them. In contrast, "any theory, by which we explain the operations of the understanding [...] will acquire additional authority, if we find, that the same theory is requisite to explain the same phenomena in all other animals" (*EHU.* 9.1). Accordingly, Hume often further verified his theories regarding human nature by determining whether they could also account for animal behavior (e.g. *T.* 1.3.16; *T.* 2.1.12; *T.* 2.2.12; *EHU.* 9). Moreover, he is not afraid to point out the patent absurdity of the Cartesian tradition: "no truth appears to me more evident, than that beasts are endow'd with thought and reason as well as men. The arguments are in this case so obvious, that they never escape the most stupid and ignorant" (*T.* 1.3.16.1). This appeal to animal thought and reason should not be misunderstood in an intellectualist manner. For Hume the primary cognitive operations are habitual tendencies based on experience rather than inferences made by some rational faculty. Thus, the way in which he dissolved the Cartesian boundary between man and animal was to argue that intelligent behavior is not generally derived from abstract reasoning and argumentation:

Were this doubtful with regard to men, it seems to admit of no question with regard to the brute creation; and the conclusion being once firmly established in the one, we have a strong presumption, from all the rules of analogy, that it ought to be universally admitted, without any exception or reserve. (*EHU.* 9.5)

In this manner a philosophical distinction is dissolved under the pressure of a more coherent science based on the continuity of life and mind. Indeed, it has been argued that Hume “treats mental functioning as importantly analogous to bodily functioning; and, by discerning similarities between animals and humans in this respect, likewise treats animal and human natures as analogous” (Buckle 2001, p. 234). In the context of a more unified science of cognition it is therefore to Hume’s great credit that he treats “men as much more like the animals than most earlier theories had done” (Stroud 1977, p. 2).

This integration of human and animal cognition under one comprehensive theory of mind provided Hume with an explanatory advantage, but the naturalization of human nature was also motivated by more fundamental concerns. This is because, while the majority of philosophers of mind continued to be lost in endless and irresolvable metaphysical speculations, the natural sciences, which are dependent on the human mind as a condition of possibility, were left without any solid foundation. The revolution of the natural sciences was indeed well underway thanks to the metaphysical protection provided by Cartesian substance dualism, but the relationship between the *res extensa* and the *res cogitans* was completely lost in the bifurcation (cf. Jonas 1966, p. 58). Thus, at the same time as natural science received its specific domain in the form of the material world, namely by banishing the subject to another realm, it also became blind to the constitutive conditions which make its existence possible. In order to resolve this deeply problematic situation it was necessary for those enabling conditions themselves to become an object of study: “consequently we ourselves are not only the beings, that reason, but also one of the objects, concerning which we reason” (*T.* Intro. 4). In other words, by naturalizing the *res cogitans* it becomes conceivable to close this explanatory gap without simultaneously undermining the natural sciences. And not only that; the science of human nature is in a position to provide a solid foundation for all of our domains of knowledge:

‘Tis evident, that all the sciences have a relation, greater or less, to human nature; [...]. Even *Mathematics*, *Natural Philosophy*, and *Natural Religion* are in some measure dependent on the science of MAN; since they lie under the cognizance of men, and are judg’d of by their powers and faculties. (*T.* Intro. 4)

While this foundational importance of studying the mind in a systematic manner is not often explicitly acknowledged in the cognitive sciences, Hume was very much aware that his science of human nature necessarily entailed some radical consequences:

There is no question of importance, whose decision is not compriz’d in the science of man; and there is none, which can be decided with any certainty, before we become acquainted with that science. In pretending therefore to explain the principles of human nature, we in effect propose a compleat system of the sciences, built on a foundation almost entirely new, and the only one upon which they can stand with any security. (*T.* Intro. 6)

We would do well to take a moment to reflect upon the significance of this passage. What is at stake here is not only the establishment of an appropriate foundation for the sciences; it is at the same time an opportunity to consider afresh all the big questions: What is life? What is mind? And, most importantly, what does it mean to be human?

To be sure, when it comes to such matters there is an unavoidable temptation to shape the answers to fit our preconceptions, or to assume certainty where none is to be had. Thus, from this perspective Hume's preferred manner of addressing these questions appears in a new light. When he says: "as the science of man is the only solid foundation for the other sciences, so the only solid foundation we can give to this science itself must be laid on experience and observation" (*T.* Intro. 7), he is demanding nothing less than an honest and open engagement with the situation of our own existence.

1.2 Outline of the paper

The aim of the next sections is to compare and contrast some essential aspects of Hume's philosophy of mind with the central claims of several relatively distinct approaches which currently co-exist in the cognitive sciences: computationalism (Section 2), embodied dynamicism (Section 3), and enactivism (Section 4). There are aspects to Hume's writings which lend themselves to be associated with each of these approaches, and as such his work cannot be said to fully belong to any one of them in a straightforward manner. Nevertheless, it will be concluded that Hume's philosophy is best approached from the perspective of more recent developments, in particular with respect to the emergence of enactivism. Finally, the extrapolation of this convergent trajectory is used as the basis for some musings about the future of the cognitive sciences.

2. Hume and computationalism

Since their inception as an institutionalized field in the 1970s, the cognitive sciences have been dominated by a research program known as *computationalism*. Proponents of this tradition are generally committed to a representational theory of perception as well as a computational theory of cognition. In this section Hume's theory of mind will be evaluated according to how compatible it is with the two most distinct variations of computationalism, namely cognitivism (Section 2.1) and connectionism (Section 2.2).

2.1 Hume and cognitivism

One of the most popular variations of computationalism, and the approach which lay the initial foundations for the cognitive sciences, is known as *cognitivism* (e.g. Fodor 1975; Pylyshyn 1984). This view has one of its most famous articulations in Newell and Simon's (1976) physical symbol system hypothesis, which states that "a physical-symbol system has the necessary and sufficient means for general intelligent action". From this traditional perspective cognition is usually viewed as centrally controlled, disembodied, and decontextualized planning as epitomized by abstract reasoning. Accordingly, the mind is conceptualized as a computer, and cognition is viewed as the computational manipulation of internal mental representations.

The philosophical foundations of cognitivism can be traced back to continental rationalism (Dreyfus & Dreyfus 1988), namely to the very tradition which Hume heavily criticized and wanted to replace with an experimental 'science of man'. In contrast to the claims of the rationalists, Hume uncovered severe problems for the

position that abstract reasoning can serve as the foundation of intelligent behavior (Kemp Smith 1905). We will review some important aspects of his alternative account of the mind in the following sections. What is important in the current context is that Hume explicitly rejects the conception of a separate representational faculty of *intellect*, and this fact accounts for a good deal what is most distinctively ‘empiricist’ about his philosophy (Garrett 1997, p. 20-25). In this manner he can be said to have attacked a foundational assumption of cognitivism, namely the existence of some unifying rational controller in the mind². Similarly, the *imagination*, for Hume the most important aspect of the understanding, is a non-reflective faculty that naturally moves from experience to belief (Biro 1993). In summary, Hume’s hypothesis is “*that all our reasonings concerning causes and effects are deriv’d from nothing but custom; and that belief is more properly an act of the sensitive, than of the cogitative part of our natures*” (T. 1.4.1.8). In other words, causal ‘reasoning’ is the result of habitual tendencies and belief is due to an affective sentiment rather than a logical inference. Since Hume explains the primary operations of the mind in explicit contrast to any intellectualist construal, it follows that his philosophy is not very amenable to a cognitivist interpretation.

Nevertheless, there is one aspect of Hume’s theory of mind which lends itself particularly well to be interpreted from such a cognitivist perspective, namely his commitment to the ‘Theory of Ideas’ (e.g. Fodor 2003). For Hume there exist two distinct kinds of perceptions: (i) *impressions*, by which he means all our sensations, feelings, passions, and emotions, and (ii) *ideas*, which he conceived of as images or copies of these impressions as used in thinking and reasoning (T. 1.1.1.1). This ‘Theory of Ideas’ and the ‘Copy Principle’, namely that “*all our simple ideas in their first appearance are deriv’d from simple impressions, which are correspondent to them, and which they exactly represent*” (T. 1.1.1.7), are essential to Hume’s philosophy because he uses them to reject claims regarding the possibility of innate ideas (e.g. Abs. 6). However, at the same time this theoretical framework also makes some parts of his philosophy of mind vulnerable to a traditional rationalist interpretation. Indeed, it has been argued that:

One thing that works against a consistent and comprehensive naturalism in Hume’s own thought is his unshakable attachment to the Theory of Ideas. That theory impedes the development of his program in several directions in which he might otherwise have pursued it. (Stroud 1977, p. 224)

For example, similar to the Cartesian (and computationalist) explanation of perception in terms of sense-represent-plan-move cycles (cf. Wheeler 2005, pp. 67-70), the body is often portrayed by Hume as an independent entity which passively furnishes the mind with simple impressions out of which more complex perceptions and ideas can then be assembled. As a case in point consider Hume’s account of the perception of solidity: “An object, that presses upon any of our members, meets with resistance; and that resistance, by the motion it gives to the nerves and animal spirits, conveys a

² Note that Hume’s critical opposition toward rationalism has arguably also received an experimental foundation in recent artificial intelligence research because in that field “the rationalist tradition had finally been put to an empirical test, and it had failed” (Dreyfus & Dreyfus 1988).

certain sensation to the mind” (*T.* 1.4.5.13)³. And since “impressions of touch are simple impressions” while the idea of “solidity necessarily supposes two bodies, along with contiguity and impulse” (*T.* 1.4.5.14), it follows that the idea of solidity must be assembled by the imagination out of a combination of several simple perceptions.

While it will be argued later on that there are also indications that Hume sometimes assigned a more constitutive role to embodiment (cf. Section 3.3.1), even in those cases he never fully recognizes the importance of active bodily movement for the constitution of perceptual objects. Thus, even though his proposed operative principles of the mind are largely based on the customary conjunction of different perceptions, he does not seem to have ever entertained the idea that these conjunctions could also be given structure by our embodied action. In this particular respect Hume’s philosophy is quite amenable to the rationalist tradition, which also conceives of perception as essentially a passive internal process that is fundamentally distinct from embodied actions. Still, despite this similarity it also needs to be emphasized that Hume explained this perceptual process in terms of pre-reflective mental operations and without recourse to the problematic metaphor of the Cartesian homunculus which performs cognitive operations on the incoming ‘data’.

Nevertheless, this contrasting style of explanation has not prevented Fodor (2003) from arguing at length that Hume’s theory of mind is more or less interchangeable with the ‘Representational Theory of Mind’. To be fair, Hume does follow his predecessors in regarding the imagination as a representational faculty of the mind. In this particular respect he differs from the rationalist tradition only in explicitly distinguishing the memory as another distinct representational faculty (Garrett 1997, p. 20). For Hume both the imagination and the memory have the function of producing ideas after the original impressions have subsided:

We find by experience, that when an impression has been present with the mind, it again makes its appearance there as an idea; and this it may do after two different ways: Either when in its new appearance it retains a considerable degree of its first vivacity, and is somewhat intermediate betwixt an impression and an idea; or when it entirely loses that vivacity, and is a perfect idea. The faculty, by which we repeat our impressions in the first manner, is call’d the MEMORY, and the other the IMAGINATION. (*T.* 1.1.3.1)

Note, however, that Hume’s distinction of representational faculties into memory and imagination – and no others – constitutes a rejection of the Cartesian ideal of a higher representational faculty of intellect (Garrett 1997, p. 39). In addition, even though it is possible to interpret Hume’s notion of ‘idea’ as a kind of internal mental representation, even Fodor (2003, p. 133) is forced to admit that “Hume’s semantic empiricism doesn’t allow mental representations of anything except what can be given in a specious present, namely, the *content* of an experience at a time”. Accordingly, the impact of Hume’s rationalistic heritage is not as extensive as it might first appear. On the contrary, Fodor actually has to reject many of Hume’s central philosophical claims in order to make his study of human nature fit with his preferred interpretation:

³ According to some anatomists of Hume’s time, ‘animal spirits’ were a kind of nerve fluid inside nerve-tubes, which was the material source of nervous transmissions in animals and humans (*EHU*, p. 235).

Take away Hume's empiricism, and his motivation for the copy theory goes too. Take away the empiricism *and* the copy theory, and what's left is a perfectly standard Representational Theory of the Mind, one that's compatible with as much (or as little) nativism as the facts turn out to require. (Fodor 2003, p. 42)

Thus, bearing in mind that Fodor (2003) is forced to claim that Hume's philosophy "needs to be purged of his empiricism" (p. 84), and that one of the aims of his own interpretation "is to abstract from the aspects of Hume's theory of mind that are dictated primarily by his epistemology" (p. 33), it should be evident that the two frameworks are not very compatible. This difficulty of interpreting Hume's work in a consistent and comprehensive manner from a cognitivist perspective should not come as a surprise, especially considering that such orthodox cognitive science is generally Cartesian in character (Wheeler 2005, p. 55).

2.2 Hume and connectionism

It is worth mentioning that there is a similar incompatibility between Hume's theory of mind and *connectionism*, an approach to cognitive science which began to challenge the cognitivist mainstream in the early 1980s (e.g. McClelland *et al.* 1986; Smolensky 1988). In contrast to cognitivism's insistence on characterizing cognition in the form of abstract symbol manipulation, connectionism views cognition and perception in terms of the emergence of global coherent states in a network of simple components. To be sure, there are some elements in Hume's philosophy which are compatible with such a sub-symbolic conception of the mind. For example, it has been demonstrated that we can make use of the connectionist approach to evaluate his theory of abstract ideas (Collier 2005) as well as his writings on the continued existence of unperceived objects (Collier 1999). Furthermore, Hume's 'bundle theory' of personal identity, namely the idea that "what we call a *mind*, is nothing but a heap or collection of different perceptions, united together by certain relations, and suppos'd, tho' falsely, to be endow'd with a perfect simplicity and identity" (*T.* 1.4.2.39), indicates some additional resemblance between connectionism and his proposed 'science of man'. However, despite the existence of these similarities, and even though Hume's *Treatise* provided one of the most influential discussions of associationism for philosophy, it turns out that "contemporary cognitive science seems to owe more to Locke and James than any of the other players in the associationist tradition" (Harnish 2002, p. 20).

Arguably one of the reasons for this incompatibility is that connectionism generally shares with cognitivism the same constitutive assumptions that underlie Cartesian psychology. Indeed, most of connectionism's disagreement with cognitivism is only over the particular nature of computationalist cognition and representation (symbolic for cognitivists, sub-symbolic for connectionsists), rather than over computationalism as such (cf. Thompson 2007, p. 10; Harnish 2002, p. 275; Wheeler 2005, p. 8).

2.3 The problem of meaning

In order to further bring out the depth of disagreement between Hume's philosophy and computationalism it will be useful to take the 'frame problem' as our starting

point. This problem was first discussed in terms of classical artificial intelligence (McCarthy & Hayes 1969), but has since also been further developed in the philosophical literature (e.g. Wheeler 2005, pp. 178-184; Dreyfus 2007). For Dennett (1978, p. 125) the frame problem is the puzzle of how “a cognitive creature, an entity with many beliefs about the world,” can update those beliefs appropriately according to the circumstances in which it finds itself. More generally, the question is how a cognitive agent faced with a practically infinite context can determine what is relevant for its current situation in such a way that it can engage in effective action (Wheeler 2005, p. 179). Interestingly, Dennett (1984) has argued that “Hume, like virtually all other philosophers and ‘mentalistic’ psychologists, was unable to see the frame problem because he operated at what I call a purely semantic level, or a *phenomenological* level”. The rejection of this assertion will naturally lead us to a better appreciation of Hume’s non-Cartesian framework.

As a first step let us consider the conditions faced by a purely rational and disembodied intellect. Since there can only be logical or demonstrative reasoning in such a purely abstract domain, there can be no frame problem. It is impossible for the imagination to conceive anything contrary to a logical proof or demonstration because all conclusions necessarily follow from their premises: “In that case, the person, who assents, not only conceives the ideas according to the proposition, but is necessarily determin’d to conceive them in that particular manner, either immediately or by the interposition of other ideas” (*T.* 1.3.7.3). However, a hypothetical mathematical intellect is useless as a model of the mind because it cannot engage in any real-world action. Accordingly, we need to embody and situate this intellect in some manner, because only embodied systems can behave. Similarly, Hume says: “Mathematics, indeed, are useful in all mechanical operations, and arithmetic in almost every art and profession: But ‘tis not of themselves they have any influence. Mechanics are the art of regulating the motions of bodies *to some design’d end or purpose*” (*T.* 2.3.3.2).

However, by instantiating the rational intellect in this manner, for example by building a good old-fashioned robot according to the computationalist approach, we have in fact already made it vulnerable to the frame problem. As soon as the system is embodied in the world it is no longer dealing with a purely logical domain where conclusions necessarily follow from their premises. On the contrary, “in reasonings from causation, and concerning matters of fact, this absolute necessity cannot take place, and the imagination is free to conceive both sides of the question” (*T.* 1.3.7.3). This problem becomes even more pressing when we consider that the mind can also add all kinds of hypothetical considerations into the mix:

The imagination has the command over all its ideas, and can join, and mix, and vary them in all the ways possible. It may conceive objects with all the circumstances of place and time. It may set them, in a manner, before our eyes in their true colours, just as they might have existed. (*T.* 1.3.7.7)

Given this practically infinite actual and imaginary context, how is the mind able to operate such that only relevant aspects of any situation affect its actions? Hume’s first reply would be that the mind is constrained by the constancy and structure imposed by a combination of associative relations and past experience. However, Dennett (1984) and Deleuze (1953, p. 123) are in agreement that this associative network of relations is not sufficient, and that the problem is how such relations can be put to appropriate

use. For example, how would these relations determine which associative direction takes priority if they are equally traversable in several directions? To be fair, only the relations of resemblance and contiguity are purely relational, whereas causation also implies a direction because it “already contains within its own synthesis of time a principle of irreversibility” (Deleuze 1953, p. 124). Nevertheless, Hume is aware that even this directionality of causal relations is not sufficient to remove the problem:

It can never in the least concern us to know, that such objects are causes, and such others effects, if both causes and effects be indifferent to us. Where the objects themselves do not affect us, their connexion can never give them any influence; and ‘tis plain, that as reason is nothing but the discovery of this connexion, it cannot be by its means that the objects are able to affect us. (*T.* 2.3.3.3)

In effect, we can interpret this passage as Hume’s rejection of any theory of mind which centers on an essentially rational intellect that is embodied in some mechanical device, i.e. the Cartesian/computationalist vision of the human condition, as untenable. Reason alone, while able to provide us with information about relations of ideas and matters of fact, cannot provide any explanation of why we ought to choose one course of action over another. Accordingly, the focus on abstract reasoning becomes a problem because it does not sufficiently determine practice: “the real question is to know which effects *interest me* and makes me seek out its cause” (Deleuze 1953, p. 124).

In this manner we are lead from a consideration of the frame problem to the very heart of Hume’s critique of the rationalist tradition. Since personal preference cannot be decided by means of a logical argument it follows that rationality, while important in its corrective capacity, cannot be the whole story. Hume’s answer to this problem is that the rationalist conception of the relationship between reason and emotion needs to be turned upside-down: “Reason is, and ought only to be the slave of the passions, and can never pretend to any other office than to serve and obey them” (*T.* 2.3.3.4). In the end his theory of mind avoids the frame problem because both belief and reason primarily derive their effect by means of feeling and passion. In contrast to the rationalist tradition, for Hume the essence of human nature is primarily affective: “A passion is an original existence, or, if you will, modification of existence, and contains not any representative quality, which renders it a copy of any other existence or modification” (*T.* 2.3.3.5). Hume was thus foreshadowing a new appreciation of the frame problem that is currently being developed by the enactive approach to robotics (e.g. Di Paolo 2003), namely that the problem of meaning is not a problem of *knowledge* but rather of *being*.

Let us finish this section by reflecting the discussion of the frame problem back onto the ongoing debates within the cognitive sciences. First, it is worth noting that the significant role which emotions play for all of our cognitive abilities has also begun to be (re)discovered (e.g. Damasio 1994). Indeed, if we take the foundational importance of the passions in Hume’s science of human nature as an indication, we can expect that the recognition of the role of emotions is likely to increase in the future: “when [man] is studied scientifically, according to Hume, it will be seen that feeling, not reason, is responsible for his thinking and acting as he does” (Stroud 1977, p. 11). Second, there is another important lesson for the cognitive sciences that we can draw

from this discussion, one which has not been the focus of much explicit debate so far. The question is: What does the existence of the frame problem entail for our understanding of what it means to be human? Hume addresses this issue in his characteristically provocative style:

Where a passion is neither founded on false suppositions, nor chooses means insufficient for the end, the understanding can neither justify nor condemn it. 'Tis not contrary to reason to prefer the destruction of the whole world to the scratching of my finger. (*T.* 2.3.3.6)

It turns out that any theory which conceives of human nature as essentially defined by a rational intellect is not only faced by a problem in terms of accounting for our *practice*, but is at the same time burdened with a serious vacuum in terms of *ethics*. Thus, while the constitutive assumptions of the Cartesian tradition were originally an attempt to protect the human from scientific disenchantment, they have inadvertently led to a theory that leaves no room for an understanding of moral behavior.

3. Hume and embodied dynamicism

Since the early 1990s the computationalist orthodoxy has begun to be challenged by the emergence of embodied-embedded cognitive science (e.g. Clark 1997; Wheeler 2005; Varela, Thompson & Rosch 1991). This approach claims that an agent's embodiment and situatedness is constitutive of its perceiving, knowing and doing. Furthermore, the computational hypothesis has been challenged by the dynamical hypothesis that cognitive agents are best understood as dynamical systems (van Gelder & Port 1995). These developments can be broadly grouped together under the heading of *embodied dynamicism* (cf. Thompson 2007, pp. 10-13). While this approach has retained the connectionist focus on self-organizing dynamic systems, it incorporates this emergentist perspective into a non-computationalist framework which holds that cognition is a situated activity which spans a systemic totality consisting of an agent's brain, body, and world (e.g. Beer 2000).

This section will argue that embodied dynamicism provides a more coherent and comprehensive interpretation of Hume's theory of mind and that, indeed, "the Humean dream of a dynamics of cognition can now be seriously pursued" (van Gelder 1998). In order to better compare and contrast Hume's approach with embodied dynamicism, we will focus on three essential aspects: (i) a dynamical approach to cognition, (ii) a situated approach to perception, and (iii) an embodied approach to mind.

3.1 Cognition is a dynamical phenomenon

Hume was one of the first philosophers to make an attempt at using the terminology of *dynamical systems theory* to characterize the properties of mental processes. Indeed, it has even been argued that "there is no doubt that he tends to think of the mind in these mechanistic Newtonian terms, and that his model of mental dynamics has a profound influence on many of his philosophical conclusions" (Stroud 1977, p. 9). While the exact relationship between Hume's philosophy and Newton's work is still a matter of debate in contemporary Hume studies (Schliesser 2007), there are

certainly some interesting analogies between the two frameworks of investigation (cf. Buckle 2001, pp. 83-90).

Thus, it is to Hume's great credit that he made an effort to establish a theory of mind in which intelligence and behavioral coherence can be explained without the necessity of having to postulate some kind of centralized controller or Cartesian 'homunculus' (Biro 1993). The first attempts to capture this kind of insight in terms of dynamical systems theory was undertaken during the cybernetics era. Ashby (1947) used the mathematics of random step-functions to demonstrate that the existence of intelligent behavior does not require an equally intelligent origin. In effect, his claim is that "if we allow breaks to occur, then all dynamic systems change their organizations until they arrive at a state of equilibrium" (Ashby 1947). Similarly, Hume made it conceivable that any behavior, which from the point of view of an external observer appears as rational, can actually be the result of a few simple interacting elements and habits (e.g. *EHU*. 9.5). Moreover, even though he did not provide any mathematical formalism, in the *Dialogues* he provides an account of the emergence of adaptive behavior which anticipates the essence of Ashby's theory. For example, when talking about living systems Hume observes:

A defect in any of these particulars destroys the form; and the matter, of which it is composed, is again set loose, and is thrown into irregular motions and fermentations, till it unite itself to some other regular form. [When that form is destroyed] a chaos ensues; till finite, though innumerable revolutions produce at last some forms, whose parts and organs are so adjusted as to support the forms amidst a continued succession of matter. (*DNR*. Part VIII)

Even though the correspondence between these two theoretical accounts is striking, it has not yet been acknowledged in any scholarly investigation. A more detailed study of this connection should therefore be of interest for future scholarly research.

We have already briefly indicated in relation to connectionism (Section 2.2) that Hume's 'bundle theory' of identity can be interpreted as a rudimentary description of a kind of self-organizing dynamical system. However, while connectionism is more interested in how such dynamical systems settle into stable attractors, the dynamical approach to cognition is more concerned with transient dynamics, especially because it generally studies systems which are embedded within sensorimotor loops (cf. Cliff 1991). This latter focus on transient dynamics and environmental situatedness is actually much more amenable to Hume's 'bundle theory' than the connectionist approach:

[We] are nothing but a bundle or collection of different perceptions, which succeed each other with an inconceivable rapidity, and are in a perpetual flux and movement. Our eyes cannot turn in their sockets without varying our perceptions. Our thought is still more variable than our sight; and all our other senses and faculties contribute to this change; [...]. (*T*. 1.4.6.4)

Perhaps this difference in emphasis between connectionism and the dynamical approach to cognition can partly explain why the former does not seem to explicitly consider Hume's 'bundle theory' as one of its precursors. However, even though his theory shares more similarities with the dynamical approach, we will find that Hume

conceived of the mind to be special type of self-organizing system, and that this makes the ‘bundle theory’ most closely resemble the enactive account of autonomous identity (cf. Section 4.1).

Let us now consider that aspect of Hume’s philosophy which has traditionally been strongly linked with Newtonian dynamics, namely the propensity by which one idea introduces another in the mind: “The qualities, from which this association arises, and by which the mind is after this manner convey’d from one idea to another, are three, *viz.* RESEMBLANCE, CONTIGUITY in time or place, and CAUSE and EFFECT” (*T.* 1.1.4.1). These relations and the principle of association form part of the core of Hume’s theory of mind. He even hints at the possibility that they might be more fundamental than Newton’s laws of nature because, “so far as regards the mind, these are the only links that bind the parts of the universe together” (*Abs.* 35). In other words, for Hume the laws of association are a condition of possibility for Newton’s achievements, “as these are the only ties of our thoughts, they are really *to us* the cement of the universe, and all the operations of the mind must, in a great measure, depend on them” (*Abs.* 35). He even used the terminology of Newtonian dynamics to describe his discovery:

Here is a kind of ATTRACTION, which in the mental world will be found to have as extraordinary effects as in the natural, and to show itself in as many and as various forms. Its effects are every where conspicuous; but as to its causes, they are mostly unknown, and must be resolv’d into *original* qualities of human nature, which I pretend not to explain. (*T.* 1.1.4.6)

From this passage it is clear that Hume conceived of the association of ideas as a law of nature similar to the gravitational force described by Newton in his *Philosophiae naturalis principia mathematica* (1687). In this respect it is worth noting that Newton captured his mechanical and gravitational laws in the mathematics of differential and integral calculus, which he invented for this purpose. On the contrary, substantive science and mathematical formulations are notoriously absent from Hume’s work. To be fair, this should be expected considering that an experimental science of human nature was non-existent before Hume’s time, and that the required tools for a mathematical treatment of cognition were not yet available. Under these circumstances Hume did the best he could to uphold the aim that “in judging the actions of men we must proceed upon the same maxims, as when we reason concerning external objects” (*T.* 2.3.1.12). His approach can thus still be seen as broadly Newtonian (cf. Buckle 2001, pp. 70-90).

To be sure, there are some significant shortcomings in Hume’s work, and it can hardly be said that he successfully established a scientific tradition based on his envisioned ‘science of man’. Nevertheless, his proposal to understand the mind in terms of dynamics has recently begun to be realized. Thus, even though Hume’s theory of mind met with little acceptance during his own lifetime and has been largely neglected by orthodox cognitive science, at least “in recent years, the Humean alternative [to the computational theory of mind] has been gaining momentum” (van Gelder 1998). The dynamical approach to cognition acknowledges Hume for being the first to champion the radical notion that “the dynamics of matter would be paralleled by a dynamics of mind” (van Gelder 1998).

Still, while there is evidently some affinity between Hume's theory of mind and the dynamical approach to cognition, one crucial difference should also be noted. Whereas the latter is largely concerned with understanding the dynamics of behavior emerging from the brain-body-world systemic whole (e.g. Beer 2000; van Gelder & Port 1995), Hume is mainly focused on determining the dynamical structure of the mind by means of analyzing his own lived experience. This commitment to a first-person approach to the study of mind aligns Hume's proposed 'science of man' more closely with the enactive approach to cognitive science, especially the neuro-phenomenological research program which attempts to use the mathematics of dynamical systems to describe the structure of our lived experience (e.g. Varela 1996; 1999).

3.2 The mind is situated

The embodied dynamicist paradigm in the cognitive sciences puts strong emphasis on the importance of *situatedness* for natural cognition and perception (e.g. Beer 2000). The notion of situatedness generally denotes that an agent interacts with its environment directly rather than through some mediating internal element such as an internal mental representation or world model (Brooks 1991). It will be useful to distinguish between (i) situated *perception* and (ii) situated *cognition*. The notion of situated perception entails that the objects which we perceive are not mere representations of external objects in our minds. Instead, we experience objects directly and without any form of mediation. In the field of artificial intelligence and robotics this alternative to the representational approach to perception has been famously captured by Brooks' (1991) slogan that 'the world is its own best model'. Still, this understanding of basic perception does not preclude the possibility that representation can play a role in more detached forms of cognition (e.g. Clark 1997; Wheeler 2005). In contrast, the notion of situated cognition goes further than this by including the situated approach to perception and additionally claiming that the concept of an 'internal mental representation' is unintelligible in general (e.g. Harvey 2008). Thus, embodied dynamicists make use of the dynamical approach to cognition as a more viable alternative to the computationalist framework (van Gelder 1998).

Does Hume conceive of the mind as situated in this manner? In what follows we will first discuss textual evidence which makes it possible to argue that Hume is strictly committed to an understanding of (i) situated perception, and then provide some arguments that he is also promoting (ii) situated cognition at least to some extent.

3.2.1 The case of situated perception

It is important to be clear from the start that the question of how the perceptual process operates is not about how we normally experience this process (*T.* 1.4.2.38). Everyone in this debate agrees that when we perceive an external object we do not experience that object as an internal representation of an external object. On the contrary, when we perceive an external object we experience that object directly:

the generality of mankind, who as they perceive only one being, can never assent to the opinion of a double existence and representation. Those very sensations, which enter by the eye or ear, are with them the true objects, nor can they readily conceive that this pen or paper, which is immediately

perceiv'd, represents another, which is different from, but resembling it. (T. 1.4.2.31)

However, even though the representational theory of perception can agree with this *description* of our experience, it can still hold that we need to conceive of perceptual objects as internal representations in order to *explain* this phenomenon. Hume rejects this representational explanation because he wants to establish a 'science of man' with an experimental method that is based on experience and observation: "The only existences, of which we are certain, are perceptions, which being immediately present to us by consciousness, command our strongest assent, and are the first foundation of all our conclusions" (T. 1.4.2.47). In contrast, the representational theory of perception is based on a hypothetical distinction between the perceived object (which we directly experience) and some assumed 'external object in-itself' (which our experience supposedly represents to us). And since the existence of the latter is in principle unverifiable by experience, as is the supposed relationship it has with the former, this distinction can only be based on a speculative metaphysical doctrine. Accordingly, Hume's strict methodological stance undermines the representational theory, "for as to the notion of external existence, when taken for something specifically different from our perceptions, we have already shown its absurdity" (T. 1.4.2.2). Indeed, Hume spends a considerable amount of effort in order to demonstrate this absurdity of the representational theory of perception, and since that philosophical position is still the working assumption of most cognitive scientists today, it is worthwhile to briefly expand on this topic.

First, we can note the fact that when we reflect on our perceptual experiences we can observe that they always depend on our body to some extent and are at all times different and often interrupted, for example when we blink with our eyes while looking at this page. In this case, despite the variability of what is perceptually given in experience, we still believe in the presence of an external object, namely the page, rather than the annihilation of one page, darkness, and then the instantaneous creation of another page that looks almost identical to the previous one. How can we account for this experience of a stable world on the basis of our variable perceptions?

Hume argues that in order to reconcile these considerations some philosophers propose a system of 'double existence' by which they distinguish between perceptions and objects: the latter are supposed to constitute the continued existence and identity which the former are thought to be lacking. However, as Hume remarks "were we not first persuaded, that our perceptions are our only objects, and continue to exist even when they no longer make their appearance to the senses, we shou'd never be led to think, that our perceptions and objects are different, and that our objects alone preserve a continu'd existence" (T. 1.4.2.46). Accordingly, this philosophical system inherits all the difficulties associated with reflection about our normal perceptual situation, and is additionally "loaded with this absurdity, that it at once denies and establishes the vulgar supposition" (T. 1.4.2.56). In other words, the representational theory of perception is nothing but a fanciful fiction, "a monstrous offspring" resulting from a clash between the operations of the imagination and deliberate reflection (T. 1.4.2.52). It is beyond the scope of this paper to discuss Hume's resolution to this problem in any detail, so a few remarks will have to suffice. In general it can be said that for Hume the process of sense perception is more like active habitual constitution rather than passive representation or information processing. On

Hume's view, the relative stability of our perceptual world is due to an instinctual subjective accomplishment which, while not rationally justifiable, happens to be a viable way of living. In essence, he views our belief in the continued and independent existence of an external world as a kind of regulatory principle of the imagination which, with the help of memory's retention of past impressions, bestows a greater coherence and constancy on our variable flow of perceptual experiences (*T.* 1.4.2)⁴.

3.2.2 The case of situated cognition

From the preceding discussion it should be clear that Hume rejects the notion of 'representation' as meaningless in the context of perception. However, his stance on the role of internal mental representations for cognition is more ambivalent due to his commitment to the 'Theory of Ideas' (cf. Section 2.1). To what extent then can his theory of mind be interpreted in terms of situated cognition? We have already seen that there is a certain affinity between his 'science of man' and the dynamical approach to cognition (cf. Section 3.1), a framework which often rejects representational terminology. Similar to the dynamic approach to cognition, Hume argued that all animals (including humans) can learn to act appropriately from experience through habit without being "directed by any such relations or comparisons of ideas, as are the proper objects of our intellectual faculties" (*EHU.* 9.6). He thus viewed the cognitive capacity for abstract mental idea manipulation as secondary, while claiming that animal "instinct or mechanical power" plays a more fundamental role in generating intelligent behavior (*EHU.* 9.6). If the bulk of intelligent behavior can be accounted for without recourse to the notion of 'idea', then what role does this representational element play in his theory of mind?

There are two important reasons why Hume needs 'ideas' in addition to 'impressions' in order for his theory to successfully account for the operations of the mind: (i) the possibility of retaining past experiences in the representational faculty of the memory, and (ii) the possibility of deliberate reflection in the representational faculty of the imagination. While Hume supposes the former to be a necessary requirement for the establishment of custom or habit, which is the primary operation of the mind, the latter is required for more 'higher-level' cognition. Abstract reflection requires 'ideas' because Hume ascribes the two elementary types of perceptions very different motivational influences. For example, when they are about pain and pleasure: "Impressions always actuate the soul, and that in the highest degree; but 'tis not every idea which has the same effect" (*T.* 1.3.10.2). Thus, if we were solely driven by impressions we would come close to being a *purely reactive system*, always being driven by fleeting momentary needs and concerns. On the other hand, if every idea also influenced our behavior then our conduct would suffer from the unsteadiness and idle wanderings of our thoughts, continuously compelling us to undertake undesirable actions (*T.* 1.3.10.2). To avoid these two extremes, "Nature has, therefore, chosen a medium" such that the influence of an idea on the will is usually negligible but can

⁴ Hume often does not make a proper distinction between the physical or causal impact on the senses (sub-personal level) and the sense impressions present to the mind (personal level). Moreover, it is also questionable whether such atomic sense impressions can even be given in our experience (cf. Sellars 1956). Nevertheless, at least in the case of our experience of an external world, he explicitly offers a personal level explanation that takes as its starting point "only a single existence, which I shall call indifferently *object* or *perception*, [...], understanding by both of them what any common man means by a hat, or shoe, or stone, or any other impression, convey'd to him by his senses" (*T.* 1.4.2.31).

increase depending on the quality of belief associated with it: “The effect, then, of belief is to raise up a simple idea to an equality with our impressions, and bestow on it a like influence on the passions” (*T.* 1.3.10.3).

In summary, the main role of the notion of ‘idea’ in Hume’s theory of mind is that their existence allows for the possibility of two important mental abilities, namely being guided by past experience (habit) as well as deliberate action (reflection). Without such ‘ideas’ it might thus at first appear that Hume would be committed to saying that our behavior is completely determined by momentary, situation-dependent impression-action reflex cycles. And indeed, we can commend Hume for having recognized the importance of having some type of mediation between mind and the world in order to enable a new kind of autonomy (cf. Jonas 1966). The role of ‘ideas’ in Hume’s theory of mind thus fulfils two mediating purposes: (i) the agent is no longer purely reactive, since it is also influenced by the history of past experiences, and (ii) the agent has the ability to pause, deliberate and reflect (including on past experiences).

However, we have already seen that the ‘Theory of Ideas’ is one of the most problematic aspects of Hume’s theory of mind (cf. Section 2.1). On the one hand, it impedes his overall project of establishing a ‘science of man’, because it works against a consistent and comprehensive naturalism (Stroud 1977, p. 224-226). On the other hand, it makes his philosophy vulnerable to a rationalist re-interpretation that, for example, makes it possible for Fodor (2003, p. 8) to argue that “Hume’s cognitive science is a footnote to Descartes’s”. It therefore seems reasonable to assume that Hume would have embraced a fully non-representational theory of mind, if that theory had also been able to account for all the required cognitive capacities. The challenge for embodied dynamicism is thus to explain these mental capacities without in the end having recourse to the notion of ‘internal mental representation’. Is it possible for us to update Hume’s account in terms of recent advances in the cognitive sciences?

In order to account for (i) we can first note that it is possible even for a purely reactive system, i.e. a system whose outputs are at each moment only determined by its current inputs, to engage in non-reactive behavior due to the history of interaction resulting from its situatedness (Izquierdo & Di Paolo 2005). In the case of a dynamical system the circumstances are even more straightforward because such a system has an internal state that changes over time, and which thereby enables past events to influence the system’s current activity (Beer 2003). In fact, it has been shown that this internal state is sufficient for a situated model agent to successfully engage in behavior that requires associative learning and re-learning without any explicit learning rules (Izquierdo & Harvey 2007). Finally, there exists an ‘organismically-inspired’ approach to robotics which attempts to understand the constitutive dynamics of habit formation (Di Paolo 2003). It therefore appears that at least in these respects that Hume’s representational explanation in terms of ‘idea’ can be replaced with a dynamical systems approach to cognition in a relatively straightforward manner.

The autonomy afforded by (i) exemplifies varying degrees of independence from direct outside control, though these typically do not have the cognitive/motivational complexity that is additionally required for the freedom of (ii). However, in order to account for (ii) the situation is much less straightforward because these ‘higher-level’ cognitive abilities are often associated with intelligent behavior for which a

representational explanation seems more appropriate. Boden (2008), for example, argues that our freedom depends to a large extent on a variety of psychological resources such as the capacity for reasoning, means-end planning and deliberate self-monitoring, and that these cognitive capacities are best understood in terms of a complex representational architecture. Similarly, Hume wanted to account for this kind of human freedom as a philosophical foundation for his later writings on morals and politics, but the only way he could conceive of doing so was on the basis of ‘ideas’.

The question whether a pure dynamical systems approach can also handle such “representation-hungry problems” (Clark & Toribio 1994) is still very much the subject of current debate in the cognitive sciences. Accordingly, while the proponents of embodied dynamicism generally agree that the primary basis of cognition is non-representational, they are divided on whether this is also the case for ‘higher-level’ cognition. Three positions can be roughly identified: (1) some researchers argue that there are some cognitive tasks, especially those requiring environmentally detached reflection, which necessitate a set of representational processes to become operative (e.g. Wheeler 2005; Clark 1997), (2) some want to push the dynamical approach as far as possible, but are nevertheless open to the possibility that the notion of representation might still end up playing a role in our understanding of cognition (e.g. Beer 2003; van Gelder 1998), and (3) some find the concept of internal mental representation unintelligible, except when it used to explain the operation of mental processes in anything but a clearly metaphorical sense (e.g. Harvey 2008). It is evident from Hume’s strict commitment to the Theory of Ideas that his own position is most similar to (1):

[T]he understanding or the imagination can draw inferences from past experience, without reflecting on it; much more without forming any principle concerning it, or reasoning upon that principle. [...]. In general we may observe, that in all the most establish’d and uniform conjunctions of causes and effects, [...], the mind never carries its view expressly to consider any past experience: Tho’ in other associations of objects, which are more rare and unusual, it may assist the custom and transition of ideas by this reflection. (*T.* 1.3.8.13-14)

Is it possible to cash out this ability for detached reflection in dynamical terms without recourse to the notion of internal mental representation? At the moment a detailed alternative theory does not exist, and some embodied dynamicists argue that this is an in principle limitation of a purely dynamical approach (e.g. Wheeler 2005; Clark 1997). However, there are also some attempts in the cognitive sciences to develop such non-representational theory of ‘higher-level’ cognition. One of the most promising in this regard is enactive cognitive science because it recognizes the importance of mediacy between mind and world as the basis for living and human autonomy, and at the same time has the explicit aim of removing the language of ‘inner mental representation’ from all cognitive science discourse (e.g. Di Paolo, Rohde & De Jaegher, in press). Moreover, while the enactive community is aware that its approach faces a considerable challenge in accounting for such ‘representation-hungry’ or ‘higher-level’ cognition, this problem is beginning to be addressed in a systematic manner from the bottom up (e.g. Stewart, Di Paolo & Gapenne, in press).

It is likely that Hume, from the point of view of his ‘science of man’, would have welcomed these recent developments in the cognitive sciences.

3.3 The mind is embodied

The notion of embodiment has been one of the major conceptual driving forces of the recent developments in the cognitive sciences (e.g. Varela, Thompson & Rosch 1991; Clark 1997; Gallagher 2005; Thompson 2007). While there is no one definition of embodiment that is generally accepted in the cognitive sciences, there is nevertheless at least some broad agreement. Embodiment is often taken to mean that an agent is physically instantiated in the world (Chrisley & Ziemke 2002), but it can also refer more specifically to the fact that having a particular body makes a difference in terms of cognition (Di Paolo, Rohde & De Jaegher, in press). In this latter sense the notion of embodiment is closely related to situatedness. It can be interpreted to mean that cognition is shaped by the kind of experiences that come from having a body with various sensorimotor capacities, and that sensory and motor processes are fundamentally inseparable since they continuously modulate each other during action in the world (Varela, Thompson and Rosch 1991, p. 173).

3.3.1 How the body shapes the mind

To begin with we can distinguish two different kinds of statements on the role of the body that broadly correspond to the two different interpretations of embodiment offered above, namely (i) the body as a physical source of perceptions, and (ii) the body as shaping the mind. While the former role of the body is also advocated by the rationalist tradition (including computationalism), the latter is more specific to embodied dynamicist approaches (including enactive cognitive science).

Hume’s most common statements regarding the role of the body is that it is a special kind of object which passively furnishes the imagination with a certain type of perceptions, namely impressions of sensation. These “original impressions or impressions of sensation are such as without any antecedent perception arise in the soul, from the constitution of the body, from the animal spirits, or from the application of objects to the external organs” (*T.* 2.1.1.1). More specifically, those perceptions arising from the constitution of the body include “all bodily pains and pleasures”, while the external organs provide “impressions of the senses”. An example of the latter was already discussed in terms of the perception of solidity (cf. Section 2.1). In order to demonstrate this bodily dependence of our various perceptions Hume invites the reader to conduct some simple “experiments, which convince us, that our perceptions are not possest of any independent existence” (*T.* 1.4.2.45):

When we press one eye with a finger, we immediately perceive all the objects to become double, and one half of them to be remov’d from their common and natural position. But as we do not attribute a continu’d existence to both these perceptions, and as they are both of the same nature, we clearly perceive, that all our perceptions are dependent on our organs, and the disposition of our nerves and animal spirits. This opinion is confirm’d by the seeming encrease and diminution of objects, according to their distance; by the apparent alterations in their figure; by the changes in their colour and other qualities from our sickness and distempers; and by an infinite number of other

experiments of the same kind; from all which we learn, that our sensible perceptions are not possest of any distinct or independent existence. (*T.* 1.4.2.45)

Does Hume also make similar claims regarding the possibility that our manner of embodiment shapes the way our mind operates? As a starting point we can note that in Hume's theory all mental operations are at least indirectly shaped by embodiment, because all sense impressions (i.e. the basic constitutive elements of his theory of mind) and some other perceptions are causally dependent on the body (e.g. *T.* 1.4.2.45; *T.* 1.4.5.30; *T.* 2.1.1.1). Furthermore, there are strong indications that Hume also conceived of the possibility of a more direct influence of the living body on mental functioning. Indeed, remarks to this effect can be found throughout his work. In the *Treatise*, for example, after explaining causal reasoning in terms of custom or habit he refers to this reasoning ability as an "instinct" that is a part of animal nature (*T.* 1.3.16.9). In contrast to rationalist appeals to logical necessity, he claims that "nature may certainly produce whatever can arise from habit: Nay, habit is nothing but one of the principles of nature, and derives all its force from that origin" (*T.* 1.3.16.9). Similar passages can be found in the *Enquiry* where he argues that intelligent behavior is more like an animal "instinct or mechanical power" (*EHU.* 9.6) than a form of abstract reasoning directed by a comparisons of ideas. And toward the end of his life he writes that "reason, in innumerable instances, is observed to arise from the principle of generation, and never to arise from any other principle" (*DNR.* Part VII), thereby implying that reason is conditioned by the animal organization (Kemp Smith 1935, p. 141).

Accordingly, it might be expected that Hume elaborates this supposed grounding of intelligence and reason in bodily properties (e.g. the 'mechanical power' and the 'principle of generation') by appealing to physiology. However, he only very occasionally makes reference to the workings of the body or nervous system (e.g. 'animal spirits') throughout his work, and never in any operational detail. Still, there is one noteworthy exception. In relation to the constitutive origins of the principle of association (i.e. the relations of resemblance, contiguity and causation), Hume proposes a hypothetical neuro-physiological explanation "in order to account for the mistakes that arise from these relations":

I shall therefore observe, that as the mind is endow'd with a power of exciting any idea it pleases; whenever it dispatches the spirits into that region of the brain, in which the idea is plac'd; these spirits always excite the idea, when they run precisely into the proper traces, and rummage that cell, which belongs to the idea. But as their motion is seldom direct, and naturally turns a little to the one side or the other; for this reason the animal spirits, falling into the contiguous traces, present other related ideas in lieu of that, which the mind desir'd at first to survey. (*T.* 1.2.5.20)

In other words, Hume proposes that the physical structure of the nervous system is constitutive of the associative variations in our thought. This passage is as close as Hume gets to providing a more detailed description of how particular ways of embodiment can shape the operations of the mind. Thus, although he claims that it would "have been easy to have made an imaginary dissection of the brain, and have shewn, why upon our conception of any idea, the animal spirits run into all the

contiguous traces, and rouze up the other ideas, that are related to it” (*T.* 1.2.5.20), anyone hoping for more detailed elaborations on this topic will be disappointed. Nevertheless, at least in general terms Hume recognized that our manner of embodiment is constitutive of cognition:

[Y]ou reason too hastily, when from the mere consideration of the ideas, you conclude that ‘tis impossible motion can ever produce thought, or a different position of parts give rise to a different passion or reflection. Nay ‘tis not only possible we may have such an experience, but ‘tis certain we have it; since every one may perceive, that the different dispositions of his body change his thoughts and sentiments. (*T.* 1.4.5.30)

Accordingly, Hume had all the ingredients for working out a more elaborate explanation of how *embodied action* can be constitutive of our perception and cognition (e.g. Noë 2004; O’Regan & Noë 2001). However, in this respect he was unable to fully free himself from the traditional philosophical bias toward the mind aspect of the mind-body relational whole. Thus, he begins the first book of the *Treatise* with the clarification that “the examination of our sensations belongs more to anatomists and natural philosophers than to moral; and therefore shall not at present be enter’d upon” (*T.* 1.1.2.1). He then repeats this warning at the start of the second book which, since it deals with the passions and emotions, might otherwise mislead the reader to expect a physiological account of bodily feelings (*T.* 2.1.1.2). This deference to the traditional boundaries of academic disciplines unfortunately prevents Hume from developing a more detailed understanding of how our embodiment, as well as the possibilities for embodied action this body affords, shape our mind in terms of both cognitive capacities and lived experiences (e.g. Gallagher 2005; Varela, Thompson & Rosch 1991).

3.3.2 Bridging the mind-body gap

These considerations of the role of the body for the mind appear to contradict Hume’s epistemological claims regarding the impossibility of knowing the ultimate origins of our sense impressions (e.g. *T.* 1.3.5.2). Thus, there are some scholars who question “whether Hume is entitled, on his premises, to draw in this way upon physics and physiology” (Ayer 1980, p. 37). However, Hume is not being inconsistent because there is no need for him to appeal to some supposed materialist or dualist metaphysics in order to make sense of the mind-body relationship. On the contrary, he argues that if it is claimed that a causal relationship between our mind and body “depends on the union of soul and body; I wou’d answer, that we must separate the question concerning the substance of the mind from that concerning the cause of its thought” (*T.* 1.4.5.30). And since we find by experience that changes in our body are united with changes in our mind, which provides the constant conjunction of experiences that according to Hume is necessary for us to form the idea of cause and effect, “we may certainly conclude, that motion may be, and actually is, the cause of thought and perception” (*T.* 1.4.5.30). In other words, he makes no metaphysical distinction between mind and body since both are encountered only from *within* the experiential domain. And “as the constant conjunction of objects constitutes the very essence of cause and effect, matter and motion may often be regarded as the causes of thought, as far as we have any notion of that relation” (*T.* 1.4.5.33).

To be sure, Hume clearly rejects the *absolute* question concerning the ultimate origin of experience as unintelligible and absurd (e.g. *T.* 1.3.5.2), but this does not prevent him from considering the *relational* question concerning the mind-body interdependency as given to us within experience. Hume's focus on the latter method is reminiscent of the dynamical approach, which bridges the mind-body gap by treating both cognitive and bodily activity as processes happening in time (cf. van Gelder & Port 1995). However, we will see later on that Hume's manner of addressing the 'hard problem' of consciousness (Chalmers 1996), namely by (i) rejecting the dogmatic reification of mind and world as independent substances, and (ii) analyzing the mind-body relationship within the domain of experience, is even more similar to enactive cognitive science's neuro-phenomenological method (cf. Section 4.4).

4. Hume and enactive cognitive science

The paradigm of *enactive* cognitive science originally emerged as a part of the embodied dynamicist approach in the early 1990s with the publication of the influential book *The Embodied Mind* by Varela, Thompson and Rosch (1991)⁵. However, while the enactive approach also emphasises the importance of embodiment, situatedness and dynamics for our understanding of mind and cognition, it has stood out from the beginning by promoting the cultivation of a principled phenomenological investigation of lived experience as a necessary complement to a standard scientific inquiry of the mind (e.g. Varela, Thompson & Rosch 1991; Varela 1996; 1999). Moreover, it has recently set itself apart even further by placing a systemic biological account of autonomous agency at the heart of its theoretical framework (e.g. Weber & Varela 2002; Thompson 2004; Di Paolo, Rohde & De Jaegher, in press). This complementary focus on biological (living) and phenomenological (lived) subjectivity clearly distinguishes the enactive approach from the rest of the competing paradigms in the cognitive sciences (cf. Thompson 2007).

In the previous section we have found that embodied dynamicism can offer a more consistent and comprehensive interpretation of Hume's work than computationalism. It follows that if there is a strong affinity between his work and enactive cognitive science, which includes the methods and insights of embodied dynamicism but is not exhausted by them, then it is best interpreted from that point of view. Accordingly, this section will focus on those aspects of enactive cognitive science which are relatively specific to this paradigm. Thus, we need to take into account its emphasis on living subjectivity in terms of biological autonomy and sense-making, as well as its focus on lived subjectivity in terms of phenomenology. And, most importantly, we must consider the defining feature of this enactive approach, namely the methodological integration of the biological and phenomenological domains in terms of the neuro-phenomenological research program.

⁵ Subsequently the label 'enactive' has also been associated with the sensorimotor account of perception of O'Regan and Noë (2001), especially due to Noë (2004). Here we will be specifically concerned with the broader enactive paradigm that began with Varela, Thompson and Rosch (1991). See Torrance (2005) and Thompson (2007, pp. 243-266) for a discussion of the relationship between these two approaches.

4.1 The insubstantial self

The biological foundation of enactivism builds on a long tradition of systemic thinking which can be traced to work done by the Chilean neurobiologists Maturana and Varela in the early 1970s (e.g. Maturana & Varela 1980). They coined the notion of *autopoiesis* (literally: self-production) in order to characterize the minimal organization of the living. In contrast to the popular functionalist or list-based approaches to the problem of defining life, Maturana and Varela proposed an operational characterization which is essential to all living beings. A recent definition of autopoiesis that has been used throughout the literature (e.g. Weber & Varela 2002; Di Paolo 2005) reads as follows:

An autopoietic system is organized (defined as unity) as a network of processes of production (synthesis and destruction) of components such that these components: (i) continuously regenerate the network that is producing them, and (ii) constitute the system as a distinguishable unity in the domain in which they exist.
(Varela 1997)

The paradigmatic example of an autopoietic system is the living cell. Autopoietic systems are autonomous systems due to their *organizational closure*, a notion which basically refers to the fact that the components of the system are both the cause and effect of their form (Thompson 2007, p. 65). Moreover, it is this reciprocal causality inherent in the living organization which grounds the autonomous identity and intentional agency of living beings (Weber & Varela 2002). In this sense the autopoietic organization forms the core of the enactive account of the constitution of living subjectivity (Thompson 2004).

How does Hume's theory of the self compare with this biological account of autonomous identity? We have already seen that Hume subscribes to what has been called a 'bundle theory' of self: "what we call a *mind*, is nothing but a heap or collection of different perceptions, united together by certain relations, and suppos'd, tho' falsely, to be endow'd with a perfect simplicity and identity" (*T.* 1.4.2.39). In this manner he rejects the existence of a substantial self as an adequate explanation for why we tend to attribute an identity to ourselves (Garrett 1981). Just like Hume did not ground our experience of a stable world in the existence of a supposed metaphysically independent world, neither is he looking to account for personal identity by appealing to a supposed substantial and independently existing self. Instead, he observes that:

The mind is a kind of theatre, where several perceptions successively make their appearance; pass, re-pass, glide away, and mingle in an infinite variety of postures and situations. There is properly no *simplicity* in it at one time, nor *identity* in different; [...]. The comparison of the theatre must not mislead us. They are the successive perceptions only, that constitute the mind; [...]. (*T.* 1.4.6.4)

The enactive approach has recognized Hume's discovery of the insubstantial nature of our identity as a valuable contribution (e.g. Varela, Thompson & Rosch 1991, p. 59-61), but it has not yet undertaken a more detailed evaluation of his theory. Is it

possible to relate the ‘bundle theory’ of self to the notion of autopoiesis? As a first step we can note that Hume takes inspiration from his account of organismic identity. He admits that providing an account of personal identity will not be easy since “to explain it perfectly we must take the matter pretty deep, and account for that identity, which we attribute to plants and animals; there being a great analogy betwixt it, and the identity of a self or person” (*T.* 1.4.6.5). Interestingly, there seems to be no generally accepted scholarly position as to why Hume singles out the living body as a model for the mind. Is there anything more to it than being the example which most easily led the imagination to attribute identity out of all the examples which he discussed in leading up to his account of personal identity (e.g. *T.* 1.4.6.8-12)? From the enactive perspective we can see that Hume was also interested in the way in which living bodies are organized:

But this [attribution of identity] still more remarkable, when we add a *sympathy* of parts to their *common end*, and suppose that they bear to each other, the reciprocal relation of cause and effect in all their actions and operations. This is the case with all animals and vegetables; where not only the several parts have a reference to some general purpose, but also a mutual dependence on, and connexion with each other. The effect of so strong a relation is, that tho’ every one must allow, in a very few years both vegetables and animals endure a *total* change, yet we still attribute identity to them, while their form, size, and substance are entirely alter’d. (*T.* 1.4.6.12)

Hume’s description of the identity of the living is remarkably similar to the definition of autopoiesis given above. He mentions both the reciprocal causality of the components as well as the constitution of an identity: “the true idea of the human mind, is to consider it as a system of different perceptions or different existences, which are link’d together by the relation of cause and effect, and mutually produce, destroy, influence, and modify each other” (*T.* 1.4.6.19). But how is such an identity, constituted by reciprocal causality, able to replace the Cartesian homunculus in terms of organizing the operations of the mind? Hume rejects any attempt at reification: “The identity, which we ascribe to the mind of man, is only a fictitious one, and of a like kind with that which we ascribe to vegetables and animal bodies” (*T.* 1.4.6.15). We can understand Hume’s recourse to the notion of a ‘fictitious identity’ as an attempt to express the insight that the identity of an organism is “nonsubstantially localized, and yet perfectly able to generate interactions” (Varela 1997). Hume expresses this point more clearly as follows:

In this respect, I cannot compare the soul more properly to any thing than to a republic or commonwealth, in which the several members are united by the reciprocal ties of government and subordination, and give rise to other persons, who propagate the same republic in the incessant changes of its parts. And as the same individual republic may not only change its members, but also its laws and constitutions; in a like manner the same person may vary his character and disposition, as well as his impressions and ideas, without losing his identity. Whatever changes he endures, his several parts are still connected by the relation of causation. (*T.* 1.4.6.19)

Thus, the analogy with the republic allows Hume to make more explicit what the mind’s organization (and autopoiesis) entails, namely *autonomy* in the sense of self-

governance (*auto* [self] *nomos* [law]). The concept of autonomy is not explicitly mentioned by Hume but it is implicit in that the republic may change its “laws and constitutions”. Similarly, the enactive approach characterizes autonomous systems as being those that “follow laws set up by their own activity. [...]. In order for a system to generate its own laws it must be able to build itself *at some level of identity*” (Di Paolo, Rohde & De Jaegher, in press). Thus, it is the autonomous system itself which, in relation to contextual constraints, continually gives rise to its own organization or ‘design’. And only because of this self-production can we say that the laws which the system follows are intrinsic to that system. Moreover, the self-constitution of such an autonomous identity necessarily entails the emergence of a correlative domain of interactions on the world proper to that identity (Thompson 2004). This corresponding constitution of a relational domain introduces an asymmetry into the relationship between the autonomous identity and its environment, and thus makes it possible for us to talk of the former as being a center of activity with a perspective on the latter (Di Paolo 2005).

4.2 Cognition is sense-making

In the previous section we have seen that Hume and the enactive approach share a similar interest in accounting for the constitution of identity in terms of a system that is defined by an organization of reciprocal causality. For the enactive approach this self-constituted identity forms the basis for a biologically grounded theory of cognition (Varela 1997), a theory that has recently been elaborated into the concept of *sense-making* (e.g. Weber & Varela 2002; Di Paolo 2005; Colombetti, in press). Indeed, this notion of sense-making, which can be understood as the creation and appreciation of meaning on the basis of exchanges with the world, has been put forward as the definitional property of a cognitive system (Di Paolo, Rohde & De Jaegher, in press). In this manner the enactive approach follows the rejection of a representational theory of mind to its natural consequence: through its self-constituted activity a cognitive system *enacts* a world of significance.

Does Hume’s theory of cognition also account for the constitution of a meaningful world in similar terms? For Hume the effective impact of an idea on the mind depends on its felt ‘vivacity’, which is normally negligible, but which can be significantly increased by what he calls ‘belief’. This notion of belief should not be understood in any reified manner as some kind of mental object, but rather in terms of a mental operation or *action*. Thus, Hume describes it as “that act of the mind, which renders realities more present to us than fictions, causes them to weigh more in the thought, and gives them a superior influence on the passions and imagination” (*T.* 1.3.7.7). Moreover, this act of belief is operationally effective not because it is logically convincing, but rather because it is fundamentally *affective*: “belief consists not in the nature and order of our ideas, but in the manner of their conception, and in their feeling to the mind” (*T.* 1.3.7.7). For Hume this affectivity entails that belief plays an essential role in determining our actions:

it is something *felt* by the mind, which distinguishes the ideas of the judgment from the fictions of the imagination. It gives them more force and influence; makes them appear of greater importance; infixes them in the mind; and renders them the governing principles of all our actions. (*T.* 1.3.7.7)

This conception of belief is similar to the enactive notion of sense-making, which “can also be seen as emotional, in that it provides values and induces actions accordingly (motivation)” (Colombetti, in press). How does belief become extended beyond the here and now such that it can be said to enact or constitute a whole world of significance? For Hume this transcendence of the current context is achieved by habitual projections on the basis of the retention of previous experiences:

Now as we call every thing CUSTOM, which proceeds from a past repetition, without any new reasoning or conclusion, we may establish it as a certain truth, that all the belief, which follows upon any present impression, is deriv'd solely from that origin. (*T.* 1.3.8.10)

As an illustration of this theory Hume provides the following example: “Suppose I see the legs and thighs of a person in motion, while some interpos'd object conceals the rest of his body. Here 'tis certain, the imagination spreads out the whole figure” (*T.* App. 4). How would the enactive approach account for the perception of the whole figure? Similar to Hume's approach, the theory proposed by Varela, Thompson and Rosch (1991) avoids the appeal to the metaphysics of representationalism and focuses instead on subjective accomplishment. Their goal is “not to determine how some perceiver-independent world is to be recovered” (p. 173). Instead, the focus of the theory remains within the capacities of the perceiver so as “to determine the common principles or lawful linkages between sensory and motor systems that explain how action can be perceptually guided in a perceiver-dependent world” (p. 173). Thus, while the enactive approach emphasizes the role of embodied action for perception, Hume emphasizes an act of the imagination. Even though he admits that “all actions and sensations of the mind are known to us by consciousness” (*T.* 1.4.2.7), he never appeals to the potential for embodied actions to structure the flow of experience. In this example, our experience of the whole person in motion is constituted by the immediate appearance of the relevant missing ideas through “their customary connexion with the present impression” (*T.* App. 4).

There are advantages to Hume's wider conception of the notion of *action*. His theory does not exclude the structuring potential of our experience of sensorimotor regularities, which have so far been the focus of the enactive approach, and it can also more easily account for perception in cases where it is difficult to appeal to embodied action:

For instance; I hear at present a person's voice, whom I am acquainted with; and this sound comes from the next room. This impression or my senses immediately conveys my thoughts to the person, along with all the surrounding objects. I paint them out to myself as existent at present, with the same qualities and relations, that I formerly knew them possess'd of. (*T.* App. 4)

We should not be misled by Hume's use of poetic language in this passage; the perception of the person in the next room is not achieved by some rational intellect. On the contrary, Hume insists that “when any object is presented, the idea of its usual attendant immediately strikes us, as something real and solid. 'Tis *felt*, rather than conceiv'd” (*T.* App. 9). Two important aspects may be noted about Hume's account of the constitution of an experiential world by the imagination: (i) it is a pre-reflective

process, because “custom operates before we have time for reflection” (*T.* 1.3.8.13), and (ii) the movement from current impressions (of sensation or memory) to belief in a world is a kind of perception because “all probable reasoning is nothing but a species of sensation” for which we must “follow our taste and sentiment” (*T.* 1.3.8.12). This instinctual process is indeed similar to the enactive notion of sense-making: it is a pre-reflective bringing forth of a context of significance for the perceiver that is dependent on that perceiver’s current circumstances and history of experiential regularities⁶.

4.3 The phenomenological connection

The enactive approach has emphasized the importance of taking first-person experience seriously since its inception (Varela, Thompson & Rosch 1991). This focus on the lived dimension of subjectivity is explored both in terms of philosophy (e.g. Thompson 2007; Bitbol 2002) and neuroscience (e.g. Varela 1996; 1999). As a consequence of this incorporation of first-person experience into the cognitive sciences it is becoming increasingly evident that we need to develop more rigorous methods for gathering and describing the phenomena in which we are interested (e.g. Roy *et al.* 1999; Varela & Shear 1999). Thus, in order to understand how to cultivate the necessary act of becoming aware in a more rigorous manner, the enactive approach is drawing inspiration from the Husserlian method of phenomenological reduction as well as the various practices of contemplative traditions (e.g. Depraz, Varela & Vermersch 2003).

How does Hume’s study of human nature relate to this focus on the investigation of lived subjectivity? As noted by Biro (1993), this is actually one of the areas in which Hume most strikingly differs from the mainstream cognitive sciences, since the latter are almost exclusively committed to measuring third-person (physiological or behavioral) data. Biro was referring to the methodological bias of computationalism, but his remarks are just as accurate with regard to most research in embodied dynamicism. However, when we interpret Hume’s choice of method from the perspective of enactive cognitive science, we can gain a sympathetic appreciation for his challenging attempt to ground the proposed ‘science of man’ on an experimental investigation of first-person experience.

It is worth noting from the start that Husserl, following his teacher Brentano, had a special admiration for Hume (Moran 2000, p. 69). Indeed, he considered him to be the first practitioner of a genuine phenomenology (Husserl 1929, §100). Husserl even credits Hume with the discovery of the problem of *constitution*, a phenomenological notion which refers to the immanent genesis of objectivities in subjectivity (Moran 2000, pp. 139-140). This is because Hume was also concerned with explaining our experience of objective features of the world, such as causality and the identity of

⁶ There is one notable difference in that Hume does not attempt to explain the origin of valence: “Nature has implanted in the human mind a perception of good or evil, or in other words, of pain and pleasure, as the chief spring and moving principle of all its actions” (*T.* 1.3.10.2). The enactive approach, on the other hand, often attempts to ground valence in the precarious situation of the self-constituting identity (e.g. Weber & Varela 2002). This additional move should not be misunderstood as a scientific explanation or logical argument; it is more of a structural match based on existential characteristics (cf. Jonas 1966).

objects, animals and persons, in terms of how these arise in consciousness. Thus, for him our natural belief in the continued and independent existence of ‘external’ reality is founded in the subjective accomplishment of the imagination (cf. Waxman 1994, pp. 238-244). This is in contrast to those philosophers or scientists who base their work on the natural attitude, which takes reality for granted without inquiring into the role of consciousness. Hume, Husserl, and the enactivists, on the other hand, are also interested in how this reality is disclosed, that is, how it comes to be given in our experience through the active engagement of the mind. Nevertheless, Husserl also criticized Hume because of his unwavering commitment to the ‘Theory of Ideas’, an underdeveloped understanding of intentionality (especially in terms of ideal objectivities), and the lack of a proper method for performing the phenomenological reduction (Husserl 1929, §100).

It is beyond the scope of this paper to compare and contrast Hume’s philosophy with Husserl’s phenomenology in any more detail (cf. Davie 1977). Nevertheless, we will briefly qualify Husserl’s general criticism regarding Hume’s apparent methodological inadequacy, especially because the role of first-person methods is also central to enactive cognitive science. It is evident that his work would have benefited from a more principled and less prejudiced engagement with his situated experience. At least he was in principle open to improve on these shortcomings, as indicated by his promotion of the ‘science of man’ as a collective effort. Indeed, much of the *Treatise* is written in the form of a debate between Hume and other (hypothetical) investigators. For example, an essential aspect of his method is an often repeated invitation for the reader to confirm his claims within their own experience. This appeal to intersubjective verification is, for instance, what defines his empiricist approach of validating the intelligibility of our theoretical ideas in terms of our experienced impressions. Thus, Hume asks the hypothetical opponent: “Does [the idea] arise from an impression of sensation or of reflection? Point it out distinctly to us, that we may know its nature and qualities. But if you cannot point out *any such impression*, you may be certain you are mistaken, when you imagine you have *any such idea*” (T. 1.2.5.28). To be sure, the constitutive import of intersubjectivity will remain largely neglected in this philosophical tradition until the later developments of Husserl’s phenomenology (cf. Zahavi 1996), but here it already finds its tentative beginnings.

Still, already Hume’s Newtonian experimental method has some similarities with the phenomenological method (Buckle 2001, pp. 82-83), though it is never developed in any explicit or principled manner. More significant in this regard is one of Hume’s most important philosophical essays, namely ‘Of the Standard of Taste’ (1757), which provides the fundamental critical theory that may have informed the projected fifth part of the *Treatise* that was to be called ‘On Criticism’ (Copley & Edgar 1998). This essay is mainly concerned with the philosophy of aesthetics, but it also outlines a number of practical skills required by any serious critic. While these criteria are discussed for the special case of artistic taste, they can easily be generalized in the form of a skilful method that enables one to discern the essential aspects of experiential phenomena. As will become evident from the criteria listed below (and similar to the phenomenological method), it takes considerable effort to master this skillful method as it is not naturally perfected in the “generality of men”.

First, the critic needs to acquire a ‘delicacy of taste’ which enables “every sense or faculty, to perceive with exactness its most minute objects, and allow nothing to escape its notice and observation”. Second, Hume observes that “nothing tends further to increase and improve this talent, than *practice*”. Moreover, this practice includes both repetition and variation: “so advantageous is practice [...], that, before we can give judgment [...], it will even be requisite that that [phenomenon] be more than once perused by us, and be surveyed in different lights with attention and deliberation”. Already on this point we can begin to see an affinity with enactive cognitive science’s concern with the establishment of a phenomenology as a practice (e.g. Varela 1996). Third, this practice needs to be complemented by frequent comparisons between various types and degrees of the phenomenon under study. Fourth, Hume states that “to enable a critic the more fully to execute this undertaking, he must preserve his mind free from all *prejudice*, and allow nothing to enter into his consideration, but the very object which is submitted to his examination”. This criterion prefigures to some extent the aims of Husserl’s methodological bracketing of our natural attitude or *epoché* (cf. Moran 2000, pp. 148-152). As a critic, says Hume, I must “forget, if possible, my individual being, and my peculiar circumstances. A person influenced by prejudice complies not with this condition, but obstinately maintains his natural position”. It takes effort to remove this prejudice. If the critic does not impose “a proper violence on his imagination” then “his taste evidently departs from the true standard, and of consequence loses all credit and authority”. Fifth, it is not enough just to once bracket our natural position. In order to avoid any bias creeping back in due to prejudice we need to employ “*good sense* to check its influence”, which makes sharp reasoning an essential skill for the critic. This last criterion is indicative of the necessity for deliberate redirection of attention during the process of becoming aware (e.g. Depraz, Varela & Vermersch 2003, p. 31).

Hume argues that by training the mind in this manner we will progress from a novice, who emphasizes the subjective aspect of qualitative phenomena, to an expert critic, who is able to discern objective qualities in the experience. A good case study for exploring the philosophical import of acquiring such a skill is wine tasting (e.g. Smith 2007). However, Hume himself never seemed to have entertained the possibility of using this method in order to develop the phenomenological foundation of his ‘science of man’ in a more principled manner. Still, even though this methodological aspect of his work is severely underdeveloped in contrast to enactive cognitive science (though the enactive approach is also just beginning to explore these issues with more seriousness), he is more explicit about the effects such training can have on our way of living. In an essay called ‘Of the Delicacy of Taste and Passion’ (1741), for example, he develops the existential consequences of becoming such a skilful critic.

He begins with the observation that “good or ill fortune is very little at our disposal” and since “great pleasures are much less frequent than great pains” this is a problem for people who “are subject to a certain *delicacy of passion*, which makes them extremely sensible to all accidents of life” and, as a consequence, often makes them act inappropriately. This is the case for the generality of mankind who base their happiness on external circumstances. However, Hume is persuaded that this problematic condition can be improved by cultivating a certain “delicacy of taste”, similar to the method outlined above. Note, however, that this is an existential transformation which is not exclusively positive: “It enlarges the sphere both of our

happiness and misery, and makes us sensible to pains as well as pleasures which escape the rest of mankind". Still, Hume argues that nothing is more "improving to the temper" and "favourable to love and friendship" than acquiring this skill.

The point of this brief detour was to show that there is a complex ethical dimension to enactive cognitive science's call for the establishment of phenomenology as a practice. For example, what are the effects of undergoing appropriate training on the personal life of the investigator? Is it acceptable to ask someone to change their way of being (not just their knowledge!) in the specified manner before they can engage in the scientific study of mind? Even though it might turn out that the overall benefits outweigh the potential drawbacks, and judging from the example of Buddhist meditation it looks like they will, these existential issues must still be properly acknowledged and discussed.

4.4 A methodology of mutual constraints

So far it has been argued that Hume's project shares commonalities with the various inspirations for enactive cognitive science, including its two founding traditions of autopoietic biology and Husserlian phenomenology. However, since the development of these traditions can also be pursued in relative independence of each other, it still needs to be demonstrated that his work also prefigures crucial aspects that are specifically *enactive*. For this purpose we will interpret his study of human nature from the perspective of what is arguably the defining feature of enactive cognitive science, namely the research program of *neuro-phenomenology* (e.g. Varela 1996; 1999; for a recent overview, see Thompson 2007, pp. 329-369).

Following Bitbol (2002), we can highlight several crucial elements of this research program, which was first formulated by Varela (1996) as methodological dissolution of the 'hard problem' of consciousness (Chalmers 1996). First, it explicitly acknowledges that any objective description arises as an invariant focus for a community of situated subjects with conscious experience. Second, it takes this situatedness as a natural starting point for defining an appropriate strategy of research, namely the establishment of a system of mutual constraints within the realm of experience from which the subject-object dichotomy arises. Third, it dissolves this dichotomy by proposing a radical redefinition of science, of nature, and of naturalization. For if we accept that nature is made of situated experiences as well as their manifold variants (both scientific and phenomenological), and that the naturalization of consciousness thus means including its disciplined contents within a strongly interconnected network of objects and experiences, then the 'hard problem' has disappeared. More precisely, the problem has been *dissolved* by a systematic reintegration of human (situated) life into the framework of the debate. In sum, neuro-phenomenology unifies the two founding traditions of enactive cognitive science through its circulation between the phenomenological (first-person) and biological (third-person) domains. In this manner the application of mutual constraints establishes the formerly separated branches at a novel, inclusive methodological level.

What can we reveal about Hume's 'science of man' when we interpret it from the perspective of neuro-phenomenology? To begin with it forces us to take a closer look at the starting point for Hume's investigations, because neuro-phenomenology derives from a conception of science as a situated activity that rejects any form of reification.

Hume similarly replaces any speculation about the metaphysical status of self and world with an analysis of their constitution in our experience. Indeed, his explicit intention was to develop a disciplined ‘experimental’ approach that takes experience as both the point of departure and the place to which we must always return:

And tho’ we must endeavour to render all our principles as universal as possible, by tracing up our experiments to the utmost, and explaining all effects from the simplest and fewest causes, ‘tis still certain we cannot go beyond experience; and any hypothesis, that pretends to discover the ultimate original qualities of human nature, ought at first to be rejected as presumptuous and chimerical. (*T. Intro. 8*)

In effect, this is just a restatement of the general aim and scope of the natural sciences, as applied to the ‘science of man’. But there is also an important difference between the two, in that the former is effectively detached and disengaged while the latter is essentially participatory⁷. Does Hume also acknowledge that scientists studying the mind are situated in this engaged manner? In fact, he was concerned that this participatory element of studying the mind has a tendency to give rise to interference when we want to study a particular situation, “as must render it impossible to form any just conclusion from the phænomenon” (*T. Intro. 10*). Nevertheless, this does not mean that he thought that it is in principle impossible to obtain the kind of structural invariants that define human nature. On the contrary, the basic scientific method remains the same in both cases, though it becomes more evident in the study of the mind that its descriptions arise from a community of situated subjects:

We must therefore glean up our experiments in this science from a cautious observation of human life, and take them as they appear in the common course of the world [...]. Where experiments of this kind are judiciously collected and compar’d, we may hope to establish on them a science, which will not be inferior in certainty, and will be much superior in utility to any other of human comprehension. (*T. Intro. 10*)

In addition, Hume contrasts his work in the *Treatise* with philosophy that he deems to be defect because of its alienation from common life. He claims that he has endeavored as much as possible to supply “those other measures of evidence on which life and action entirely depend, and which are our guides even in most of our philosophical speculations” (*Abs. 4*). This intention to take our lived situation seriously is unfortunately sometimes neglected by Hume in favor of theoretical arguments, though he does return to it occasionally, and especially when his skepticism becomes too much: “Here then I find myself absolutely and necessarily determin’d to live, and talk, and act like other people in the common affairs of life” (*T. 1.4.7.10*)⁸. The skeptical attitude, which has been induced through an

⁷ Note that this distinction between natural and mind science does not hold with respect to quantum mechanics, which is essentially participatory and structurally similar to the study of mind (cf. Bitbol 2002).

⁸ Hume reports that the limitations of human reason have made him ready to reject all belief and reasoning: “Where am I, or what? [...] What beings surround me? [...] I am confounded with all these questions, and begin to fancy myself [...] environ’d with the deepest darkness, and utterly depriv’d of the use of every member and faculty” (*T. 1.4.7.8*). It is worth noting the specific manner in which he finds back into the natural attitude: “I dine, I play a game of back-gammon, I converse, and am merry with my friends” (*T. 1.4.7.9*). Here we have a first rudimentary phenomenological description of how

understanding of the limitations of reason, thus becomes mitigated by a natural and necessary return to concrete situatedness. In this manner “we might hope to establish a system or set of opinions, which if not true (for that, perhaps, is too much to be hop’d for) might at least be satisfactory to the human mind, and might stand the test of the most critical examination” (*T.* 1.4.7.14).

From this perspective we can understand Hume’s mitigated skepticism as a method of mutual constraints which, similar to neuro-phenomenology, removes prejudice and dogmatism, and thus makes it possible to discover steady principles “which will suit with common practice and experience” (*T.* 1.4.7.14). Moreover, Hume rejects any attempt at forcing common people to live up to the dogmatic expectations derived from abstract reflection, but rather wishes the influence to run the other way: “They do well to keep themselves in their present situation; and instead of refining them into philosophers, I wish we cou’d communicate to our founders of systems, a share of this gross earthy mixture, as an ingredient, which they commonly stand much in need of” (*T.* 1.4.7.14). In effect, Hume turns the rationalist tradition upside down by promoting the pre-reflective understanding that is shared by all of mankind as the final measure. Similar to neuro-phenomenology’s rejection of any reification of scientific concepts, the starting point within the experience of our lived situation is the basis for Hume’s strict dismissal of the metaphysical implications of our ideas. For example, when rejecting the notion that we have any absolute idea of ‘external existence’ he remarks:

Let us fix our attention out of ourselves as much as possible: Let us chace our imagination to the heavens, or to the utmost limits of the universe; we never really advance a step beyond ourselves, nor can conceive any kind of existence, but those perceptions, which have appear’d in that narrow compass. This is the universe of the imagination, nor have we any idea but what is there produc’d. (*T.* 1.2.6.8)

It is this starting point within lived experience which provides Hume with the motivation to reject the representational theory of perception (cf. Section 3.2.1) and the development of an alternative account focusing on the subjective accomplishment of the imagination (cf. Section 4.2 and 4.3). Moreover, his account of the conditions which make us attribute a continued and independent existence to the objects of our experience is explicitly derived from the perspective of our common situation (*T.* 1.4.2.31). However, this appeal to the constitutive dimension of our experience has also often led to Hume being summarily dismissed as an idealist or radical skeptic. Since such criticisms could also be leveled against enactive cognitive science, it is important to clear up this confusion. As Thompson (2007, p. 21) makes clear, “the point here is not that the world would not exist if not for consciousness. Rather, it is that we have no grip on what reality means apart from what is disclosed to us as real, and such disclosure necessarily involves the intentional activity of consciousness”. In other words, the enactivists are interested in how the mind operates such that we experience a real and meaningful world. Similarly, Hume is not committed to the sceptical metaphysical claim that it is impossible for the world that we experience to *exist* independently of the constitutive actions of a conscious observer. Rather, the claim is an epistemological one, in that no matter whether it exists in this manner or

the category of the real is grounded in intersubjectivity, a theme more fully developed by the later Husserl (cf. Zahavi 1996).

not, the mind cannot *know* what the fact of the matter is. Accordingly, the supposed metaphysical status of the world is of no consequence to how the mind actually constitutes the *experience* of an ‘independently existing external world’. This experience must therefore be approached without recourse to the *metaphysical* notion of an independently existing external world.

All of this amounts to the claim that we cannot rationally prove the validity of the natural attitude, but we can ascertain its constitutive conditions within experience. In other words, while the ultimate causes of our experiences are perfectly inexplicable by human reason, we do not need to appeal to such metaphysical speculations in order to answer specific scientific questions. The viability of our explanations is always determined in relation to other experiences: “We may draw inferences from the coherence of our perceptions, whether they be true or false; whether they represent nature justly, or be mere illusions of the senses” (*T.* 1.3.5.2).

This rejection of metaphysical speculation, especially the reification of self, body and world, and the corresponding return to our situatedness within lived experience sets the stage for addressing the ‘hard problem’ of consciousness: How are the objective description of human nature and the actuality of our experience related? This question is at the core of the neuro-phenomenological method, which responds by proposing the establishment of “mutual constraints between (i) the actuality which is necessarily *presupposed* by the description, and (ii) certain elements which *belong* to this description” (Bitbol 2002). The focus of enactive cognitive science has been to apply this method in terms of the mutual constraints between stabilized contents of experience and certain neurophysiological processes (e.g. Varela 1999). Is it also possible to find such an enforced mutual constraint between the subjective and the objective domain in Hume’s ‘science of man’? To be sure, we have already seen that his experiential reports can hardly be said to be sufficiently stabilized, and proper biological considerations are seriously lacking. Nevertheless, a plausible case can be made that an appropriate kind of mutual constraints is consistently employed as a part of his ‘science of man’, namely between the subjectivity of our starting point within lived experience and the objectivity of human nature, which Hume often describes as a kind of “anatomy” (e.g. *T.* 1.4.6.23).

A prominent example of this kind of circulation between subjective and objective phenomena, while always remaining within the experiential domain, is Hume’s study of constitution. For instance, in the case of the constitution of external objects, he remarks that an objective description of human nature reveals that we cannot directly perceive, nor rationally justify our belief in, the continued and independent existence of the world. However, he also observes that, due to our concrete lived situation, we must still believe it to exist in this manner because “Nature has not left this to his choice” (*T.* 1.4.2.1). Hume then proceeds to dissolve this apparent tension by rejecting any metaphysical speculation and instead addressing the gap between the objective description of human nature and the actuality of our lived situation from within the experiential domain:

We may well ask, *What causes induce us to believe in the existence of body?* but ’tis in vain to ask, *Whether there be body or not?* That is a point, which we must take for granted in all our reasonings. The subject, then, of our present

enquiry is concerning the *causes* which induce us to belief in the existence of body. (*T.* 1.4.2.1)

Thus, in Hume's 'science of man' there exists a circulation of mutual constraints between the subjective conditions that are presupposed by the objective description of human nature, and certain elements that belong to this objective description. Of course, the objectivity of human nature should not be understood in any absolute sense, as this has already been rejected on the basis of Hume's starting point in situated experience (cf. *T.* Intro. 8). Nevertheless, the constraints imposed by our lived situation are concrete since, any radical skeptics can only entertain their doubt "till the necessities of nature, unsatisfied, put an end to their miserable existence" (*EHU.* 12.23). Thus, any objective investigation of the constitutive dimension of subjectivity is in the end limited by an absolute *viability constraint*, the extent of which is determined by the requirements of our concrete actuality. Thus, neuro-phenomenology's concern with the reintegration of human life into the cognitive sciences is matched by Hume's deference to nature: "Indulge your passion for science, says she, but let your science be human, and such as may have direct reference to action and society. [...] Be a philosopher; but, amidst all your philosophy, be still a man" (*EHU.* 1.6).

5. The future of the cognitive sciences

It thus seems that the ongoing internal development of the cognitive sciences roughly coincides with a theoretical movement from a Cartesian rationalist to a Humean empiricist philosophy. As always, history repeats itself, but this time the dialectic has shifted to the empirical domain of the natural sciences. Throughout this paper we have witnessed this ongoing development in terms of a broadening scientific context as we have traced the history of the cognitive sciences.

Whereas the cognitivist understanding of Hume was limited to the 'Theory of Ideas' and the abstract reasoning it affords, connectionism was also able to integrate his distributed approach to the mind. Embodied dynamicism further expanded this interpretative shift by also including Hume's emphasis on dynamic, distributed, pre-reflective and non-representational processes as the primary basis of perception and cognition. Finally, with the advent of enactive cognitive science we were able to see that it is crucial for Hume's theory of mind, traditionally often only derived from the first book of the *Treatise* ("Of the Understanding"), to be complemented by and grounded in the topics on affectivity in the second book ("Of the Passions"). Also, enactivism's interest in the practice of the phenomenological method provided the basis for a new appreciation of Hume's proposed experimental method. Finally, this interpretative expansion culminated in an evaluation of Hume's overall 'science of man' in terms of enactive cognitive science's neuro-phenomenological research program.

If Hume's 'science of man' is indeed interpreted more consistently and comprehensively by more recent developments in the cognitive sciences, culminating in the enactive approach, is it perhaps possible to extend this trajectory and anticipate future directions of research? In other words, can we speculate about how the cognitive sciences would have to further develop in order to make an even more

comprehensive interpretation of Hume's work possible? For example, considering the importance of feelings, sentiments and passions for Hume's philosophy of mind, we might expect the cognitive sciences to focus more on the role of affectivity. We have already seen that the enactive approach is well positioned to make valuable contributions to this area, especially in term of the notion of sense-making (e.g. Weber & Varela 2002; Di Paolo 2005; Colombetti, in press). There are also ongoing efforts to use the framework of enactive cognitive science as a basis for gaining a better understanding of affect (e.g. Varela & Depraz 2005), values (e.g. Di Paolo, Rohde & De Jaegher, in press) and emotions (e.g. Thompson 2007, pp. 360-381; Colombetti, in press).

Even though we have exclusively concentrated on Hume's theory of mind in this paper, it is important to note that he was also deeply concerned with ethics, social theory, history and political philosophy (cf. Copley & Edgar 1998). In fact, it is possible to argue that "Hume is a moralist and a sociologist, before being a psychologist" (Deleuze 1953, p. 21), and that he placed his theory of mind in the service of these concerns. If this is the wider context which we can speculate that the cognitive sciences will have to begin to address, would the enactive approach be able to accommodate and even promote these changes? While it is beyond the scope of this paper to address this final context of Hume's work, it is worth pointing out that the enactive approach has already started to explore some of these further directions. For example, there is an opportunity to evaluate Hume's notion of 'sympathy', which plays a fundamental role in his study of human nature, from the perspective of the enactive concern with the phenomenological concepts of empathy and open intersubjectivity (e.g. Thompson 2007, p. 382-393). Indeed, the enactive approach is beginning to investigate the social dimension of cognition, a topic largely ignored in mainstream cognitive science, in terms of the autonomy of social interactions (e.g. De Jaegher & Di Paolo 2007), and the role of play for 'higher-level' cognition (e.g. Di Paolo, Rohde & De Jaegher, in press). It is also starting to apply its theoretical framework to the wider range of social and cultural activities in which human beings continually engage (e.g. McGann 2007). Finally, this ongoing shift toward an understanding of the mind, which treats the social dimension as an essential constitutive element, also opens up the possibility for a scientifically informed theory of moral and ethical behavior. Some initial steps in this direction have already been made, for example by Varela (1992) and Thompson (2007, p. 401-402).

In summary, following Thompson (2007, p. 411), we can say that the enactive approach is beginning to promote the general claim that "the knowing and feeling subject is not the brain in the head, or even the brain plus the body, but the socially and culturally situated person, the enculturated human being". Hume could not have agreed more. The challenge for enactive cognitive science is to live up to this revolutionary ambition.

Acknowledgements. Many thanks to Mike Beaton, Ezequiel Di Paolo, Joel Parthemore, Tom Ziemke, Charles Wolfe and two anonymous reviewers for their helpful comments on an earlier draft of this paper. Much of this paper was written during a stay in 2007 at the University of Skövde funded by the euCognition network (www.eucognition.org), as part of the European Cognitive Systems initiative.

References

- Ashby, W. R. (1947), "The Nervous System as Physical Machine: With Special Reference to the Origin of Adaptive Behavior", *Mind*, **56**(221), pp. 44-59
- Ayer, A. J. (1980), *Hume*, New York, NY: Oxford University Press, 2000
- Beer, R. D. (2000), "Dynamical approaches to cognitive science", *Trends in Cognitive Sciences*, **4**(3), pp. 91-99
- Beer, R. D. (2003), "The dynamics of active categorical perception in an evolved model agent", *Adaptive Behavior*, **11**(4), pp. 209-243
- Biro, J. (1993), "Hume's new science of the mind", in: D.F. Norton (ed.), *The Cambridge Companion to Hume*, Cambridge, UK: Cambridge University Press, pp. 33-63
- Bitbol, M. (2002), "Science as if situation mattered", *Phenomenology and the Cognitive Sciences*, **1**, pp. 181-224
- Boden, M. A. (2008), "Autonomy: What is it?", *BioSystems*, **91**(2), pp. 305-308
- Brooks, R. A. (1991), "Intelligence without representation", *Artificial Intelligence*, **47**(1-3), pp. 139-160
- Buckle, S. (2001), *Hume's Enlightenment Tract: The Unity and Purpose of 'An Enquiry Concerning Human Understanding'*, New York, NY: Oxford University Press
- Chalmers, D. J. (1996), *The Conscious Mind: In Search of a Fundamental Theory*, New York, NY: Oxford University Press
- Chrisley, R. & Ziemke, T. (2002), "Embodiment", in: L. Nadel *et al.* (eds.), *Encyclopaedia of Cognitive Science*, London, UK: Macmillan Publishers, pp. 1102-1108
- Clark, A. (1997), *Being There: Putting brain, body, and world together again*, Cambridge, MA: The MIT Press
- Clark, A. & Toribio, J. (1994), "Doing without representing?", *Synthese*, **101**(3), pp. 401-431
- Cliff, D. (1991), "Computational Neuroethology: A Provisional Manifesto", in: J.-A. Meyer & S. W. Wilson (eds.), *From Animals to Animats: Proc. of the 1st Int. Conf. on Simulation of Adaptive Behavior*, Cambridge, MA: The MIT Press, pp. 29-39
- Collier, M. (1999), "Filling the Gaps: Hume and Connectionism on the Continued Existence of Unperceived Objects", *Hume Studies*, **25**(1-2), pp. 155-170
- Collier, M. (2005), "Hume and Cognitive Science: The Current Status of the Controversy over Abstract Ideas", *Phenomenology and the Cognitive Sciences*, **4**(2), pp. 197-207
- Colombetti, G. (in press), "Enaction, sense-making and emotion", in: J. Stewart, O. Gapenne & E. A. Di Paolo (eds.), *Enaction: Towards a New Paradigm for Cognitive Science*, Cambridge, MA: The MIT Press, forthcoming
- Copley, S. & Edgar, A. (1998), "Introduction", in: S. Copley & A. Edgar (eds.), *David Hume: Selected Essays*, New York, NY: Oxford University Press, pp. vii-xxii
- Damasio, A. R. (1994), *Descartes' Error: Emotion, Reason, and the Human Brain*, New York, NY: G.P. Putnam's Sons
- Davie, G. (1977), "Edmund Husserl and 'the as yet, in its most important respect, unrecognized greatness of Hume'", in: G. P. Morice (ed.), *David Hume: Bicentenary Papers*, Edinburgh, UK: Edinburgh University Press, pp. 69-76
- De Jaegher, H. & Di Paolo, E. A. (2007), "Participatory sense-making: An enactive approach to social cognition", *Phenomenology and the Cognitive Sciences*, **6**(4), pp. 485-507
- Deleuze, G. (1953), *Empirisme et Subjectivité: Essai sur la Nature Humaine selon Hume*, trans. by: C. V. Boundas, *Empiricism and Subjectivity: An Essay on Hume's Theory of Human Nature*, New York, NY: Columbia University Press, 1991

- Dennett, D. C. (1978), *Brainstorms: Philosophical Essays on Mind and Psychology*, Montgomery, VT: Bradford Books
- Dennett, D. C. (1984), "Cognitive Wheels: The Frame Problem of AI", in: C. Hookway (ed.), *Minds, Machines, and Evolution: Philosophical Studies*, Cambridge, UK: Cambridge University Press, pp. 129-151
- Depraz, N., Varela, F. J. & Vermersch, P. (2003), *On Becoming Aware: A pragmatics of experiencing*, The Netherlands, Amsterdam: John Benjamins Publishing
- Di Paolo, E.A. (2003), "Organismically-inspired robotics: homeostatic adaptation and teleology beyond the closed sensorimotor loop", in: K. Murase & T. Asakura (eds.), *Dynamical Systems Approach to Embodiment and Sociality*, Adelaide, Australia: Advanced Knowledge International, pp. 19-42
- Di Paolo, E.A. (2005), "Autopoiesis, adaptivity, teleology, agency", *Phenomenology and the Cognitive Sciences*, **4**(4), pp. 429-452
- Di Paolo, E. A., Rohde, M. & De Jaegher, H. (in press), "Horizons for the Enactive Mind: Values, Social Interaction, and Play", in: J. Stewart, O. Gapenne & E. A. Di Paolo (eds.), *Enaction: Towards a New Paradigm for Cognitive Science*, Cambridge, MA: The MIT Press, forthcoming
- Dreyfus, H. L. (2007), "Why Heideggerian AI failed and how fixing it would require making it more Heideggerian", *Philosophical Psychology*, **20**(2), pp. 247-268
- Dreyfus, H. L. & Dreyfus, S. E. (1988), "Making a mind versus modelling the brain: artificial intelligence back at a branch-point", *Daedalus*, **117**(1), p. 15-44
- Fodor, J. A. (1975), *The Language of Thought*, Cambridge, MA: Harvard University Press
- Fodor, J. A. (2003), *Hume Variations*, New York, NY: Oxford University Press
- Gallagher, S. (2005), *How the Body Shapes the Mind*, New York, NY: Oxford University Press
- Garrett, D. (1981), "Hume's Self-Doubts about Personal Identity", *The Philosophical Review*, **90**(3), pp. 337-358
- Garrett, D. (1997), *Cognition and Commitment in Hume's Philosophy*, New York, NY: Oxford University Press
- Harnish, R. M. (2002), *Minds, Brains, Computers: An Historical Introduction to the Foundations of Cognitive Science*, Oxford, UK: Blackwell Publishers Ltd.
- Harvey, I. (2008), "Misrepresentations", in: S. Bullock, J. Noble, R. Watson & M. Bedau (eds.), *Artificial Life XI: Proc. of the 11th Int. Conf. on the Simulation and Synthesis of Living Systems*, Cambridge, MA: The MIT Press, pp. 227-233
- Hume, D. (1739-1740), *A Treatise of Human Nature*, D. F. Norton & M. J. Norton (eds.), New York, NY: Oxford University Press, 2000
- Hume, D. (1739-1740), "An Abstract of ... *A Treatise of Human Nature*", in: D. F. Norton & M. J. Norton (eds.), *David Hume: A Treatise of Human Nature*, New York, NY: Oxford University Press, 2000
- Hume, D. (1741), "Of the Delicacy of Taste and Passion", in: S. Copley & A. Edgar (eds.), *David Hume: Selected Essays*, New York, NY: Oxford University Press, 1998, pp. 10-13
- Hume, D. (1748), *An Enquiry Concerning Human Understanding*, T. L. Beauchamp (ed.), New York, NY: Oxford University Press, 1999
- Hume, D. (1757), "Of the Standard of Taste", in: S. Copley & A. Edgar (eds.), *David Hume: Selected Essays*, New York, NY: Oxford University Press, 1998, pp. 133-154
- Hume, D. (1779), "Dialogues Concerning Natural Religion", in: J. C. A. Gaskin (ed.), *David Hume: Principal Writings on Religion including Dialogues Concerning Natural Religion and The Natural History of Religion*, Oxford, UK: Oxford University Press, 1998
- Husserl, E. (1929), *Formale und transzendente Logik. Versuch einer Kritik der logischen Vernunft*, in: P. Janssen (ed.), *Husserliana XVII*, The Hague, Netherlands: Martinus Nijhoff, 1974

- Izquierdo-Torres, E. & Di Paolo, E. A. (2005), "Is an Embodied System Ever Purely Reactive?", in: M. Capcarrere *et al.* (eds.), *ECAL 2005: Proc. of the 8th Euro. Conf. on Artificial Life*, Germany, Berlin: Springer-Verlag, pp. 252-261
- Izquierdo, E. & Harvey, I. (2007), "The Dynamics of Associative Learning in an Evolved Situated Agent", in: F. Almeida e Costa *et al.* (eds.), *ECAL 2007: Proc. of the 9th Euro. Conf. on Artificial Life*, Berlin, Germany: Springer-Verlag, pp. 365-374
- Jonas, H. (1966), *The Phenomenon of Life: Toward a Philosophical Biology*, Evanston, Illinois: Northwestern University Press, 2001
- Kemp Smith, N. (1905), "The Naturalism of Hume (I.)", *Mind*, **14**(54), pp. 149-173
- Kemp Smith, N. (1935), "Introduction", in: N.K. Smith (ed.), *Hume's Dialogues Concerning Natural Religion*, London, UK: Oxford University Press, pp. 1-153
- Maturana, H. R. & Varela, F. J. (1980), *Autopoiesis and Cognition: The Realization of the Living*, Dordrecht, Holland: Kluwer Academic Publishers
- McCarthy, J. & Hayes, P. J. (1969), "Some philosophical problems from the standpoint of artificial intelligence", in: B. Meltzer & D. Michie (eds.), *Machine Intelligence 4*, Edinburgh, UK: Edinburgh University Press, pp. 463-502
- McClelland, J. L., Rumelhart, D. E. & the PDP Research Group (1986), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Vol. 2: Psychological and Biological Models*, Cambridge, MA: The MIT Press
- McGann, M. (2007), "Enactive theorists do it on purpose: Toward an enactive account of goals and goal-directedness", *Phenomenology and the Cognitive Sciences*, **6**(4), pp. 463-483
- Moran, D. (2000), *Introduction to Phenomenology*, London, UK: Routledge
- Newell, A. & Simon, H. A. (1976), "Computer Science as Empirical Inquiry: Symbols and Search", *Communications of the Association for Computing Machinery*, **19**(3), pp. 113-126
- Noë, A. (2004), *Action in Perception*, Cambridge, MA: The MIT Press
- O'Regan, J. K. & Noë, A. (2001), "A sensorimotor account of vision and visual consciousness", *Behavioral and Brain Sciences*, **24**(5), pp. 939-1031
- Pylyshyn, Z. (1984), *Computation and Cognition: Toward a Foundation for Cognitive Science*, Cambridge, MA: The MIT Press
- Rosenberg, A. (1993), "Hume and the philosophy of science", in: D. F. Norton (ed.), *The Cambridge Companion to Hume*, Cambridge, UK: Cambridge University Press, pp. 64-89
- Roy, J.-M., Petitot, J., Pachoud, B. & Varela, F. J. (1999), "Beyond the Gap: An Introduction to Naturalizing Phenomenology", in: J. Petitot, F. J. Varela, B. Pachoud & J.-M. Roy (eds.), *Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science*, Stanford, CA: Stanford University Press, pp. 1-80
- Schliesser, E. (2007), "Hume's Newtonianism and Anti-Newtonianism", in: E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Summer 2007 Edition, URL = <http://plato.stanford.edu/archives/sum2007/entries/hume-newton/>
- Sellars, W. (1956), "Empiricism and the Philosophy of Mind", in: H. Feigl & M. Scriven (eds.), *The Foundations of Science and the Concepts of Psychology and Psychoanalysis*, Minneapolis, MN: University of Minnesota Press, pp. 253-329
- Smith, B. C. (2007), "The Objectivity of Tastes and Tasting", in: B. C. Smith (ed.), *Questions of Taste: The Philosophy of Wine*, New York, NY: Oxford University Press, pp. 41-77
- Smolensky, P. (1988), "On the proper treatment of connectionism", *Behavioral and Brain Sciences*, **11**(1), pp. 1-23
- Stewart, J., Di Paolo, E. A. & Gapenne, O. (in press), "Introduction", in: J. Stewart, O. Gapenne & E. A. Di Paolo (eds.), *Enaction: Towards a New Paradigm for Cognitive Science*, Cambridge, MA: The MIT Press, forthcoming
- Stroud, B. (1977), *Hume*, New York, NY: Routledge

- Thompson, E. (2004), "Life and mind: From autopoiesis to neurophenomenology. A tribute to Francisco Varela", *Phenomenology and the Cognitive Sciences*, **3**(4), pp. 381-398
- Thompson, E. (2007), *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*, Cambridge, MA: The MIT Press
- Torrance, S. (2005), "In search of the enactive: Introduction to special issue on enactive experience", *Phenomenology and the Cognitive Sciences*, **4**(4), pp. 357-368
- Van Gelder, T. (1998), "The dynamical hypothesis in cognitive science", *Behavioral and Brain Sciences*, **21**(5), pp. 615-665
- Van Gelder, T. & Port, R. F. (1995), "It's About Time: An Overview of the Dynamical Approach to Cognition", in: R. F. Port & T. van Gelder (eds.), *Mind as Motion: Explorations in the Dynamics of Cognition*, Cambridge, MA: The MIT Press, pp. 1-43
- Varela, F. J. (1992), *Un Know-How per l'Etica*, trans. as: *Ethical Know-How: Action, Wisdom, and Cognition*, Stanford, CA: Stanford University Press, 1999
- Varela, F. J. (1996), "Neurophenomenology: A Methodological Remedy for the Hard Problem", *Journal of Consciousness Studies*, **3**(4), pp. 330-349
- Varela, F.J. (1997), "Patterns of Life: Intertwining Identity and Cognition", *Brain and Cognition*, **34**(1), pp. 72-87
- Varela, F. J. (1999), "The specious present: a neurophenomenology of time consciousness", in: J. Petitot, F. J. Varela, B. Pachoud & J.-M. Roy (eds.), *Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science*, Stanford, CA: Stanford University Press, pp. 266-314
- Varela, F. J. & Depraz, N. (2005), "At the Source of Time: Valence and the constitutional dynamics of affect", *Journal of Consciousness Studies*, **12**(8-10), pp. 61-81
- Varela, F. J. & Shear, J. (1999), 'First-person Methodologies: What, Why, How?', *Journal of Consciousness Studies*, **6**(2-3), pp. 1-14
- Varela, F. J, Thompson, E. & Rosch, E. (1991), *The Embodied Mind: Cognitive science and human experience*, Cambridge, MA: The MIT Press
- Waxman, W. (1994), *Hume's theory of consciousness*, Cambridge, MA: Cambridge University Press
- Weber, A. & Varela, F. J. (2002), "Life after Kant: Natural purposes and the autopoietic foundations of biological individuality", *Phenomenology and the Cognitive Sciences*, **1**, pp. 97-125
- Wheeler, M. (2005), *Reconstructing the Cognitive World: The Next Step*, Cambridge, MA: The MIT Press
- Zahavi, D. (1996), *Husserl und die transzendente Intersubjektivität: Eine Antwort auf die sprachpragmatische Kritik*, trans. by: E. A. Behnke, *Husserl and Transcendental Intersubjectivity: A response to the Linguistic-Pragmatic Critique*, Athens, OH: Ohio University Press, 2001