Foundations of Science, special issue on "The Impact of Radical Constructivism on Science", edited by A. Riegler, 2001, vol. 6, no. 1–3: 99–124.

Radical Constructivism in Biology and Cognitive Science

John Stewart
Theoretical Biology Group, Department of Advanced Studies,
Gulbenkian Institute of Science, Oeiras, Portugal
email: stewart@igc.gulbenkian.pt

Abstract.

This article addresses the issue of "objectivism vs constructivism" in two areas, biology and cognitive science, which are intermediate between the natural sciences such as physics (where objectivism is dominant) and the human and social sciences (where constructivism is widespread). The issues in biology and in cognitive science are intimately related; in each of these twin areas, the "objectivism vs constructivism" issue is interestingly and rather evenly balanced; as a result, this issue engenders two contrasting paradigms, each of which has substantial specific scientific content. The neo-Darwinian paradigm in biology is closely resonant with the classical cognitivist paradigm in cognitive science, and both of them are intrinsically objectivist. The organismic paradigm in biology, based on the concept of autopoiesis, is consonant with the paradigm of "enaction" in cognitive science the latter paradigms are both profoundly constructivist.

In cognitive science, the objectivism vs constructivism issue is internal to the scientific field itself and reflexivity is inescapable. At this level, strong ontological objectivism is self-contradictory and therefore untenable. Radical constructivism is self-coherent; but it *also* rehabilitates a weak form of objectivism as a pragmatically viable alternative. In conclusion, there is an even-handed reciprocity between "objectivist" and "constructivist" perspectives. Finally, the article examines the consequences of this conclusion for fields other than cognitive science: biology; physics and the natural sciences; and the human and social sciences.

Keywords.
Autopoiesis, cognitivism, neo-darwinism, objectivism, organism.

I. Introduction.

Radical constructivism is well established in many areas of the human and social sciences, but to date has made little impact in the natural sciences. A plausible reason for this is that human social reality is easily understood as being "constructed": it is clear that there was a time when human society did not exist; it is possible (in principle, and increasingly in practice) to study the pre-historical and historical processes through which this reality came into existence and evolved; and our growing ecological consciousness makes us only too aware that there may well come a time when human society will cease to exist on the planet Earth. In other words, the object of the human and social sciences clearly seems to be ontologically (and not just epistemologically) constructed. By contrast, classical physics - which in our culture is the very prototype of natural science - is based on essentially ahistorical objects and laws. Rosen (1991) has pointed out that the generic ideal of the natural sciences is to achieve a theory cast in the form of a state-determined dynamic system. This is clearly the case for Newtonian dynamics; Rosen argues that subsequent developments, notably relativity and quantum mechanics, do not weaken but if anything actually strengthen this ideal. The point here is that in principle, knowledge of the state of a state-determined dynamic system at any point in time implies (via recursive application of the dynamic law) knowledge of the state of the system at all future times and indeed for all past time as well. Thus, in a deep sense, nothing really changes; as Whitehead (1926) expressed so clearly, "there is nothing to evolve, because one set of external relations is as good as any other set of external relations". Thus, the scientific objects of the natural sciences are epistemologically constituted precisely in such a way as to appear that they are *not* ontologically constructed. It is thus not surprising that radical constructivism has little appeal.

This article is composed of three sections. The first section addresses the question of biology, which intuitively occupies an intermediate position between physics on the one hand and the human sciences on the other. I shall seek to show that there are two contrasting paradigms within biology, the currently orthodox neo-Darwinian paradigm, and an alternative paradigm based on the theory of autopoiesis. The neo-Darwinian paradigm takes as its central scientific object the gene; I shall argue that the gene is epistemologically constituted as timeless and ahistorical, so that this paradigm is understandably hostile to constructivism. By contrast, the autopoietical paradigm takes as its central object the ongoing process of self-construction by which living organisms maintain their own existence; this paradigm is essentially historical and intrinsically constructivist.

The second section of this article will address the question of cognitive science. Here, I propose to argue that the classical paradigm - the computational theory of mind - is formally closely analogous to the neo-Darwinian paradigm in biology; that it necessarily presupposes an obectivist ontology; and that it is therefore logically opposed to constructivism. By contrast, the founders of the autopoietical paradigm, Maturana and Varela (1980), explicitly consider that "Life" and "Cognition" are fundamentally the *same* scientific object[1], so that this theory constitutes an alternative paradigm in cognitive science as well as in biology; an alternative which is, of course, constructivist. The particularity of cognitive science, compared to physics and even to biology, is that it is intrinsically reflexive: cognitive science is *itself* a cognitive activity, and therefore inescapably forms part of its own object.

In the third section, I shall seek to show that the objectivist paradigm is incoherent when this reflexivity is taken into account. However, this does not spell unconditional victory for the constructivist paradigm. This latter paradigm is indeed self-coherent; but it *also* actually rehabilitates the opposing objectivist paradigm, on condition that the claims made by such a science are modestly pragmatic and not ontological.

II. Biology.

In the early part of this century, Whitehead (1926) wrote: "A thoroughgoing evolutionary philosophy is inconsistent with materialism. The aboriginal stuff, or material, from which a materialistic philosophy starts is incapable of evolution. This material is in itself the ultimate substance. Evolution, on the materialistic theory, is reduced to the rôle of being another word for the description of the changes of the external relations between portions of matter. There is nothing to evolve, because one set of external relations is as

---

[1] The general vague idea that life is a "knowledge-gaining process" does of course have a long history, notably in the work of Lorenz. However, Maturana and Varela have sharpened the point to constitute a fundamental *theory* , conjointly of life and cognition.

good as any other set of external relations. There can merely be change, purposeless and unprogressive. But the whole point of the modern doctrine is the evolution of the complex organisms from antecedent states of less complex organisms".

We can clearly recognize here the opposition, mentioned in the introduction, between the static, reified ontology of classical physics which is anti-constructivist, as contrasted with the relational ontology of process and organism proposed by Whitehead and which, in our terms, is clearly constructivist. For Whitehead, it is therefore evident that "the doctrine [of evolutionary biology] thus cries aloud for a conception of organism as fundamental for nature". It is therefore quite astonishing to find that the major current in contemporary biology, the neo-Darwinian paradigm, defines as its central scientific objects not organisms, but those static, formal entities that are called genes. This reversal, which gives rise to a non-constructivist biology, requires some explanation.

II.1. The neo-Darwinian paradigm.

Darwin's most original contribution was not so much the idea of evolution itself (that idea was shared by many of his contemporaries and predecessors, including his great rival Lamarck, and had even been put forward in a long philosophical poem, "Zoomania", by his own grandfather); it lay in proposing a *mechanism* whereby evolutionary change could come about (Lenay 1999). That mechanism, as is well known, is Natural Selection. Darwin amassed a large amount of empirical evidence to show that in natural populations, there is substantial heritable variation for most characters - hence the possibility of artificial selection by animal and plant breeders. The core of the theory is that - given this heritable variation - if certain forms are better adapted, they will leave more descendents which resemble them, and so over time the population will evolve. Darwin's theory was immensely controversial at the time, and provoked a storm of objections - the vast majority of which were quite irrational and easily disposed of by Darwin's henchman T.H. Huxley. However, some of the objections were rational; and among these Darwin himself recognized that one was particularly important, and actually threatened the viability of his theory. This was the objection put forward by an engineer, Fleeming Jenkin.

It was commonly accepted at that time that heredity consisted of a mixture of the influences of the two parents, and that offspring would therefore be intermediate between their parents. Darwin's own theory of inheritance, which involved "gemmules" secreted by the bodies of the parents, was in accordance with this sort of "blending inheritance". The consequence of this, however, is that any new variation will be "diluted out" by blending before Natural Selection has a chance to establish it. Even more seriously, 50% of the heritable variation in the population will be lost by blending at each generation; and so a very considerable source of new heritable variation is necessary at each generation. Darwin himself was vague and uncertain on this point, and tended to attribute this new variation to the inheritance of acquired characteristics. This, however, would actually undermine the theory of natural selection, as we shall now see.

The point is a logical one. If usage gives rise to new variations, and if these acquired characters are inherited, then this is already a mechanism for evolution and natural selection is no longer necessary. As a corollary, if acquired characters are adapted[1] (and experience shows that they often are - for example, the skin on the sole of the foot thickens and muscles grow stronger with use), then evolutionary change will ensure adaptation,

[1] Acquired characters are not *necessarily* adapted (for example, habits of drug use); the fact that they so often *are* adaptive itself requires explanation. Beyond a vague and rather empty petition of principle, neo-Darwinian biology is not illuminating on this point; a proper explanation requires an organism-centred account of ontogeny (see below).

again without any need to appeal to natural selection. In other words, if acquired characters are inherited, natural selection becomes superfluous and is at best a partial explanation of evolutionary change. Thus, Weismann's theory of a separation between "germ plasm" and "soma", and the denial *in principle* of the very possibility of the inheritance of acquired characters, is not a revisable option but is a crucial cornerstone of the neo-Darwinian paradigm. Weismann's theory led to the "rediscovery" of Mendel's "laws", which provide a solution to the problems caused by blending inheritance. The Mendelian theory requires a clear conceptual distinction between the "phenotype", the external appearance of an organism, and the "genotype" or the hereditary constitution of the individual. For a single phenotypic character, the genotype consists of two genetic "factors" (nowadays called "genes"), one being received from each parent and one of which (at random) will be passed on to each offspring. The key point is that these genetic factors are *particulate* and do not "mix" or blend when different types of factors coexist in a single heterozygous individual. "Segregation" is the technical term for the fact that when genes are passed on to the next succeeding generation, they "separate out" unchanged with respect to the form in which they were inherited from the preceding generation. In other words, "genes" are scientific objects which are epistemologically constituted as essentially unchanging, and as certainly unaffected by the physiological and developmental process of the living organisms which host them for a while before passing them on. Weismann himself raised the problematic question as to how the "germ plasm" can remain quite immune from all somatic processes, and yet at the same "direct" those processes; and came up with the neat solution that the germ plasm must make "working copies" which migrate to the cytoplasm and get involved in the dirty work while leaving the master copies intact. This is a brilliant characterisation of the relation between DNA and messenger-RNA, a century before their empirical discovery in molecular terms; it reinforces the concept that genes themselves do not actually *participate* in the ongoing processes of life.

Genes are not absolutely immutable, of course, otherwise evolutionary change would be impossible. However, genetic "mutations" are both rare and essentially random accidents. Thus, the synthesis between Darwinian natural selection and Mendelian genetics, which is the very basis of the neo-Darwinian paradigm, solves both the problems caused by blending inheritance of acquired characters. The genetic variation in a population is maintained intact from generation to generation, due to the particulate nature of the genetic factors, so that there is plenty of time for natural selection to act on new genetic variants; and new genetic variants arise by a rare random process, so that the maintenance of adaptation in a changing environment and systematic evolutionary change does absolutely require natural selection.

The neo-Darwinian paradigm is a gene-centred form of biology. I have dwelt at some length on the historical constitution of this paradigm, *before* the 1950's and the discovery of DNA, in order to show that the integration of Mendelian genetics is not a contingent or revisable feature; it is quite crucial to the epistemological viability of natural selection as "the" explanation of evolution. However, there is a price to this: the emphasis on genes has lead away from a consideration of the concept of "organism". This is not an accident; Mendelian genetics is a highly formal discipline and, however strange it may seem, it is not actually a theory of heredity. This apparently paradoxical point is so important that it merits some detailed, technical treatment.

In Mendelian genetics[1], the "genotype" as such cannot be observed directly[2]: it is a theoretical hypothesis or postulate[3]. *Differences* in genotype are inferred from observed *differences* in a phenotypic character; if there are no phenotypic differences, no genotypes can be inferred. It follows that there is not and cannot be any genetics of a character which is invariant in the species in question. Thus, genetics can and does explain the distribution of (for example) blue eyes vs brown eyes in the offspring of particular crosses; but if all the animals in the population have eyes, genetics does not and cannot explain why there are eyes in the first place that might be blue or brown. To take another example, genetics does not and cannot explain why the offspring of cats are kittens which grow up into cats that *resemble* their parents.

This is the sense in which genetics is not a science of heredity. The "blind spot" is even more complete if the character in question is common to all living organisms. Thus, the phenomenon of autopoiesis - the fact that all living organisms *produce themselves* by their ongoing dynamic processes, which we shall examine more closely below - is and can only be ignored by Mendelian genetics and, hence, by the neo-Darwinian paradigm. It is to be noted that this is a true blind-spot, in the sense that not only is there a key feature of living organisms that is not seen; from a neo-Darwinian perspective, it is intrinisically impossible to appreciate *that* there is a feature which is not seen. Nevertheless, this lacuna gives rise to a serious weakness in the neo-Darwinian account of evolution. In order to qualify as a fully-fledged scientific explanation, it should be possible to *predict* the course of evolution on the basis of the relative fitnesses of different alleles at each gene locus. To date, this has only been possible for small-scale micro-evolution, and even then only in a very few cases (sickle-cell anaemia, melanism in moths, sexual selection for tail-length in birds....)[4]. It is generally admitted that neo-Darwinian theory fails signally to account for macro-evolution. The unfortunate consequence is that the theory becomes largely non-falsifiable, and degenerates into purely retrospective speculations or "Just-So" stories which are quite uncontrollable (Saunders 1989).

What is the significance of these fundamental limitations of the neo-Darwinist paradigm for the question as to whether the scientific object of biology is epistemologically

---

[1] In what follows, I shall use the simple term "genetics" instead of referring pedantically each time to "Mendelian genetics". Note, however, that this carries a risk of confusion: Mendelian genetics is precisely *not* a science of the constructive processes of genesis, as in "phylogenesis", "ontogenesis" or Piaget's "genetic structuralism".

[2] Nowadays, of course, DNA can be observed; but the nucleotide sequence of a piece of DNA is *not* a genotype, and to confuse the two is a reification which generates grave conceptual confusion.

[3] This is not an objection: in physics, all the fundamental scientific objects, such as atoms, quarks, and electro-magnetic fields, cannot be directly observed as such. It should, however, as a brake on the slippery slope to objectivist reification.

[4] Even in the favourable case of micro-evolution, neo-Darwinian theory explains less than it might seem, because it *presupposes* both the variants and the complete phenotypes on which selection is based. In order to account for the phenotype itself, an organismic biology is indispensable. It is revealing that the "complete phenotype" harbours some surprises, even in the classical cases. Thus, in sickle-cell anaemia, differential reproduction results not just from the well-known resistance to malaria, but also from differences in fertility which are dependent on social class. Again, in the case of melanism in moths, differential fitness results not just from the well-known camouflage effect and differences in predation, but also from an increased hardiness and capacity to pass the winter of the melanic form. However, these very examples do illustrate that in the case of micro-evolution, neo-Darwinian theory can in principle make falsifiable predictions; so it is important not to throw out the micro-evolutionary baby with the bath-water.

constituted as an entity which is ontologically constructed through history? The answer is that from a neo-Darwinian perspective, the only scientific object that is properly constituted epistemologically is the gene; and as I have said, DNA genes remain essentially unchanged during the physiological processes of life and sexual reproduction. Consequently, the dynamic history of organisms and eco-systems is evacuated and "evolution" becomes an unintelligible description of changes in genes[1]. It is rather as though human history were reduced to a description of written inscriptions and "texts" which no-one is able to interpret. The result is that the dimension of "genesis" in the term "phylogeny" is completely lost. An analogous reduction has occurred in the case of ontogenesis, the process by which a fertilized egg-cell develops to become an adult organism. The biological discipline of embryology has practically disappeared, even as a description of the actual bodily morphology of the developing embryo, let alone as an explanation of the dynamics of the developmental system that underlies these morphogenetic processes. It has been replaced by accounts of the patterns of gene activation at different stages in development and in various differentiated tissues. As a matter of dogma, derived from the Weismannian concepts of germ plasma and soma, it is maintained that the genome constitutes a "genetic programme" which directs development; but this remains a pure hypostatic petition of principle, devoid of any genuine explanatory content (Oyama 1985).

II.2. Autopoiesis: an alternative paradigm in biology.

As I have noted, the absence of a theory of living organisms as dynamic processes of becoming is a "blind-spot", and this makes it difficult to perceive the lack from within the neo-Darwinian perspective itself. We require an alternative perspective to search for a response to Whitehead's appeal: "The doctrine [of evolution] thus cries aloud for a conception of organism as fundamental for nature. It also requires an underlying activity - a substantial activity - expressing itself in individual embodiments, and evolving in achievements or organism. The organism is a unit of emergent value, a real fusion of the characters of eternal objects, emerging for its own sake" (Whitehead 1926). In order to express this in more contemporary terms, we may find a basis if we note that living organisms are thermodynamically and materially open systems. As Prigogine and Stenghers (1979) have noted, such systems spontaneously gives rise to "dissipative structures". In order to emphasize the essentially dynamic nature of such "structures", the French philosopher Simondon (1964) has coined the term "individuation". In line with the insights of Prigogine, Simondon has noted that "dissipative structures" can arise in purely physical systems: for example, a whirlpool, a tornado, a star, or a growing crystal. However, such examples of physical individuation are characterized by the fact that they are essentially ephemeral: they last only as long as certain external conditions, over which they have no control, happen to be maintained. By contrast, biological entities are characterized by the fact that the processes of individuation which constitute them as recognizable entities *also* systematically prolong and maintain the boundary conditions which are necessary for these processes to endure indefinitely. All living organisms are mortal: if ontologically they could not die, it would be meaningless to assert that they are "alive". However, the death of a living organism is always essentially an accident; they are intrinsically potentially immortal. This is very well illustrated by the case of unicellular

[1] The majority of DNA is "junk DNA", and even in functional genes, the majority of codons have little or no functional consequences; so that if even "evolution" is reduced to a description of changes in DNA nucleotide sequence, there is not much *to* understand.

organisms[1]. To the extent that doubling in volume and dividing in two is not the same as dying, all the unicellular organisms alive today are as old as life on our planet, i.e. over 3000 million years - and they are still going strong. The global nature of this dynamic process has been well expressed by Lovelock (1988) under the name of the "Gaia hypothesis": the totality of living organisms on the planet Earth function in such a way as to dynamically and indefinitely maintain the conditions (the temperature, the presence of water, the composition of the atmosphere and so on) that are necessary for life itself.

Simondon's concept of processes of individuation that recursively sustain the conditions for their own continuation is strongly convergent with the concept of autopoeisis put forward by Maturana and Varela (1980), and also with the concept of dynamic systems that are "closed under efficient cause" put forward by Rosen (1991). In this alternative perspective, genes are important, and indeed ubiquitous in all forms of life as we know it today; but they are not in themselves the "secret of life". Much contemporary rhetoric notwithstanding (Lewontin 1993), genes do not "reproduce themselves". They have the distinctive feature that in the context of a living organisms, they can *be copied* accurately, and independently of their meaning[2]; but left to themselves they are remarkably inert[3] and do not do anything at all. They certainly do not "direct" the ontogeny of multicellular organisms[4]. The common misconception that the regularity of ontogeny requires "instruction" of the process by "information" from an external source, be it environmental or genetic, is a symptom of the deep-seated dichotomy in Western thought between "Matter" and "Form", allied to the assumption that "Matter" is in itself inert or at best chaotic, so that only "Form" (Platonic or otherwise) can be creative (Oyama 1985). However, if (with Whitehead and Prigogine) this concept is abandoned in favour

---

[1] The argument applies equally, in principle, to multicellular organisms; so the fact that the latter generally exhibit senescence and death through ageing itself requires explanation (and incidentally poses a severe problem for neo-Darwinism: how can senescence and death possibly be a selective *advantage* ? Weismann, notably, got into a terrible tangle over this question). I cannot adequately enter this fascinating subject here: I refer to Medawar (1956) for what seems to me essentially the correct answer.

[2] There is a fascinating analogy here with alphabetic writing. Ideographic characters, for example, or semitic scripts without vowels, cannot be properly copied without being interpreted. By contrast, alphabetic writing, which is based on a conventional coding of the phonetic forms, is "orthothetic", i.e. it can be controllably copied without being interpreted (indeed, as proof-readers know, it is often better controlled for accuracy if it is *not* interpreted). The social consequences of this have been immense.

[3] DNA is, indeed, the most inert of the molecular components of living organisms - which explains why it can be recovered intact from mammouths frozen in polar ice, Egyptian mummies, or Neanderthal hominids. At the same time, in the perspective of the alternative autopoietic paradigm, genes in the context of a living organism are not the essentially impervious, static entities that they are in the neo-Darwinian perspective: as McClintock showed over 60 years ago, they "jump" from site to site in the genome, and the genome as a whole is remarkably fluid (Ho 1997).

[4] Ironically, when it comes to *explaining* ontongeny, the neo-Darwinian paradigm which self-consciously aims at a mechanistic ontology actually attributes magical and vitalistic capacities to the genes (Oyama 1985). The infamous concept of "selfish genes" (Dawkins 1976) is a red herring here; the metaphor according to which genes are "homonculi" which knowingly calculate their selective advantage, although rather tendencious, can with care be reformulated in rigorous terms. My point, and Oyama's, is rather that what is "magical" is attributing genes with the power to "direct" ontogeny; it is like attributing a book such as the Coran, which is an inert object, with the power to "direct" an Islamic revolution. In the social realm, it is clear that the revolution is carried out by human beings, and directed by a charismatic human leader who uses rhetorical devices to attribute a "sacred" status to the book and to establish claims to have interpreted the book authoritatively. Likewise, in the biological realm, ontogeny is accomplished not by the genes but by the organism as a whole.

of an ontology based on process, the need for an appeal to external information disappears. The developing embryo extends into the realm of multicellular organisation the basic principle of autopoiesis: the organismic process is autonomous and essentially *organizes itself.* It is certainly significant that Waddington (1956), one of the few major biologists to have actually attempted a substantial theory of ontogenesis as such, extensively used the term "individuation". He used this term in the context of embryonic development to stress the tendency of developing organisms to self-organize into coherent wholes during morphogenesis, despite perturbations. He adopted the term from Jung's use in psychology, which provides an interesting cycle of relationships between biology and cognition, of which Waddington was keenly aware, resulting in part from dialogues he had with Piaget (Goodwin, personal communication).

I hope I have said enough to show that there is a real issue in the epistemological foundations of biology - the structural formalism of Mendelian genetics and neo-Darwinism vs an alternative organismic paradigm in which the notion of process is central; and to suggest that this opposition may not be unrelated to the issue of constructivism. However, this is not quite sufficient to show that objectivism as such is really an issue in biology. To take the argument a step further, I shall now turn to a domain where the issue of objectivism vs constructivism is intrinsic to the scientific object itself. This is the field of cognitive science. I shall then reinforce my argument concerning biology by showing that the opposing paradigms in cognitive science are in strong resonance with those in biology.

III. Cognitive Science.

There is room for reasonable doubt whether the ontological issue of obectivism vs constructivism (von Glaserfeld 1988) has any real impact on the actual practice of natural scientists. As Latour and Woolgar (1979) have perceptively noted: 'Scientists themselves constantly raise questions as to whether a particular statement "actually" relates to something "out there," or whether it is a mere figment of the imagination, or an artefact of the procedures employed..... Depending on the argument, the laboratory, the time of the year, and the currency of controversy, investigators will variously take the stand of realist, relativist, idealist, transcendental relativist, sceptic, and so on'. Given this pragmatic flexibility, what are the grounds for claiming that a whole paradigm may be distinctively "objectivist" or "constructivist"?

In the natural sciences in general, and in physics in particular, the answer to this question is indeed far from obvious. In quantum mechanics, for example, the same mathematical equations, and the same experimental procedures of measurement, are compatible with either an objectivist or a constructivist ontology at the level of interpretation[1]. There is, however, one area where the question of constructivism is directly relevant to the object of science itself: and this is cognitive science. Any scientific theory concerning the nature of knowledge - more generally, of "cognition" - is *necessarily* based on a theory concerning the relationship between the subject and object

[1] Quantum mechanics has been astoundingly successful at the level of operational effectiveness: the agreement between theoretical prediction and empirical observations often extends to many places of decimals. However, it is something of a scandal that over fifty years later, there is no reasonable *interpretation* of this theory, and in particular of what happens when the wave-packet collapses in conjunction with an observation. This is what shocked Einstein, and accounts for his life-long reticence with respect to quantum mechanics. It may be that the issue of objectivism or constructivism has to do primarily with the *meaning* of the science.

of knowledge. There are two major paradigms in cognitive science, and they have contrasting positions on this very point. To summarize briefly, I shall now argue that classical cognitivism, which is based on the computer metaphor, is intrinsically objectivist; whereas the alternative paradigm, which takes living organisms as the paradigm for cognition, is intrinsically constructivist.

III.1. Classical cognitivism.

Historically, the field now known as "cognitive science" first emerged in the 1940's with the cybernetics movement (Gardner 1985); major figures were Turing (the theory of Universal Computation) and von Neumann (the architecture of actual computers). It reached a canonical or "classical" form in the 1970's, notably with Fodor and Chomsky. With respect to psychology, cognitivism was a reaction to behaviourism which had proposed to explain behaviour purely in terms of externally observable stimulus-response patterns, while avoiding any hypotheses as to what went on in the "black box" of the organism itself. Cognitivism proposed to open this "black box", and to posit the existence of "mental states" composed of symbols. "Cognition" is then defined as computation (i.e. syntaxically defined operations based on the pure form of the symbols); the symbols nevertheless gain semantic meaning from the fact that they are representations of a referential state of affairs in the world[1]. On this basis, Fodor (1975) proposed to explain behaviour schematically as follows: the animal has representations:

i) of its situation;

ii) of the various actions it can take in that situation;

iii) of the forseeable consequences of each of these actions;

iv) an ordered list of preferences for these consequences.

On Fodor's view, if an animal behaves in a certain way, it is because it has computed that this action will lead to the most preferable consequences. I shall not enter here into a long discussion of the various merits and demerits of the cognitivist paradigm; I shall focus on just two aspects, both of which relate to the key issue of "representation".

The first point is that the symbolic computations are in themselves purely formal and meaningless; these operations are only "cognitive" and meaningful because the symbols gain semantic content from the fact that they are "representations" of states of affairs in the real world. The symbols are "grounded" on the basis of the precise correspondance relations between each symbol and its reference which is an aspect of reality. Precisely because the computations are purely syntaxic, the referents of the symbols must have an independent existence in order to be able to confer semanticity on symbols which are intrinsically meaningless. In other words, this scheme can only work if the referents - the objects of knowledge - pre-exist and are what they are, quite independently of whether a cognitive subject is there or not, and whether the subject's representations are correct (i.e. in a proper correspondence relation with the referent) or not. This is, rather precisely, a specification of an objectivist ontology; what I wish to emphasize is that it is quite fundamental to the coherence of the whole cognitivist paradigm.

The second point concerns the "symbol grounding problem". For both theoretical and practical reasons, the very project of "grounding" the symbols on the basis of correspondence relations with pre-given referents has proved to be highly problematic. The theoretical problems have been identified from within the cognitivist paradigm itself: Putnam (1981), notably, has shown that a truth-conditional semantics inevitably suffers

---

[1] Lakoff (1987) has explained in a most illuminating way how this scheme is a redeployment of the formalist movement in mathematics initiated by Hilbert.

from indeterminacy of the referent. The practical problems have been revealed most notably in the field of experimental robotics. Engineers have found that it is practically impossible to set up a mobile robot in such a way that its "representation" of its situation (e.g. its position and orientation on a map of its domain) robustly and automatically corresponds to its actual situation[1]. The difficulty, of course, is that the robot has no means of knowing what the engineer takes to be its "true" situation *other* than its own representation; and therefore, if that representation is in error, the robot has no means of correcting it.

My main point is not that the symbol-grounding problem is necessarily insuperable. It has, however, proven sufficiently troublesome to have given rise to a most revealing *rapprochement* between cognitivism and the neo-Darwinian paradigm. The idea is simple and *a priori* attractive. Animals whose representations are not in correspondance with reality are most likely to suffer a decrease in viability, and will therefore be selected against. Conversely, animals whose correspondance relations are good will be favoured. In other words, it is natural selection that will solve the problem by guaranteeing at least adequate grounding of the symbolic representations. Current research from within the cognitivist paradigm is actively exploring this possibility (e.g. Millikan 1984, Dretske 1988, Pacherie 1993).

Lenay (1993) has pointed out, in considerable detail, the remarkable homology between the intellectual structure of cognitivism and that of neo-Darwinian biology. Genes, like symbols, are epistemologically constituted as purely formal entities; both gain their meaning from correspondence relations, genes with phenotypic characters that are "adapted", symbols with referents. But perhaps the most revealing homology of all is the curiously abstract and ultimately empty nature of the correspondance relation between the two terms. In both cases, one goes straight from an abstract entity to its ultimate reference, while totally short-circuiting the actual material mediation of this relationship. In the alliance between cognitivism and neo-Darwinism, the "truth" of a symbolic representation is supposed to be guaranteed by natural selection in the same way that the "fitness" of a genotype is guaranteed. We have already noted the weakness of this argument in biology: to the extent that it is not possible to define the "fitness" of a genotype *other* than by observing, *after the event,* which have survived and which have not, the "survival of the fittest" collapses into a tautology (Saunders 1989). Thus, it is not certain that the appeal of cognitivism to neo-Darwinian biology will really help to solve the symbol-grounding problem; on the contrary, it may just highlight weaknesses that are common to both paradigms.

Be that as it may, however, the main point that I want to make is that there is indeed a deep affinity between the computational theory of mind in cognitive science, and the neo-Darwinian paradigm in biology. And since, as I have argued, the cognitivist paradigm is intrinsically objectivist, this greatly strengthens the connexion between neo-darwinism and objectivism.

III.2. Enaction.

Varela et al (1993) have recounted how the limitations and deficiences of the classical cognitivist paradigm have led, through the intermediate stages of connexionism and emergentism, to the appearance of a new paradigm which they propose to designate by

---

[1] In industrially operational robotics, the only practical solution has been to engineer the robots' environment so that it corresponds to their representation! For a natural system, this is obviously putting the cart before the horse.

the term "enaction". What I wish to emphasize here is that this paradigm bears a privileged relationship to the biology of autopoiesis. Maturana (1980) has recounted his long search for answers to two questions: (i) "What is the organization of the living?" and (ii) "What takes place in the phenomenon of perception?"; and his realization, in 1969, that the answers to the two questions both involved a basic circularity and were, in fact, the same. The term "autopoiesis" was coined by Maturana and Varela (1980), quite explicitly in order to express this basic identity of "life" and of "cognition".

From the point of view of cognitive science, this clearly represents a shift from the cognitivist paradigm, in which the operation of a computer is taken as a root metaphor for cognition, to an alternative paradigm in which living organisms as such are taken as the basic metaphor. Now, in what sense can it be said that *all* living organisms, from the simplest bacteria onwards, "know" something? It is clear that they do not have propositional "knowledge *that"* such-and-such is the case. They have, paradigmatically, implicit tacit "know-*how"* - most fundamentally, knowledge "how" to function in such a way as to keep themselves alive. Now it will be immediately clear that this knowledge is quite inseparable from the existence of the *subject* of the knowledge in question. In other words, it may be taken as the very prototype of a constructivist position with respect to knowledge; the object of knowledge does *not* pre-exist independently of the existence of a cognitive subject.

Taking this as a starting-point does not mean reducing cognitive science to the microbiology of primitive bacteria. On the contrary, the study of cognition becomes the study of biological evolution, or more precisely the evolutionary emergence of successive forms of life; more precisely still, of life-strategies and the accompanying "worlds" (Umwelt) that they bring into existence. As I have argued at length elsewhere (Stewart 1996), each succesive step in this evolution - multicellular organisms, animal communication, hominization and language, the entry into human history with writing, and so on - is clearly an emergent phenomenon, and the nature of the new realm of "reality" that is brought forth each time is manifestly "constructed" through the historical process itself. Thus, contrary to the impression that is sometimes conveyed, it is definitely not the case that the constructivist theory of autopoiesis is appropriate only for lower forms of cognition, whereas objectivist cognitivism is appropriate for "higher" forms of knowledge as expressed in the propositional statements of human language. The thematic opposition between an objectivist and a constructivist paradigm runs throughout the whole of cognition - and the whole of biology - from the lowliest organisms to the "highest".

Putting various pieces of my argument together, this prompts me to point out that the neo-Darwinian concept of the "adaptation" of a primitive bacteria is indeed objectivist, in the sense that this adaptation is conceived of as adaptation to an objective, pre-given environment. In this respect, neo-Darwinism is actually a regression, an impoverishment of Darwin's own conceptions. As Lenay (1999) has pointed out, Darwin was clear that the very existence of a new variation might bring into existence the particular "selective forces" that would subsequently favour it. In the case of Artificial Selection by animal and plant breeders, which played such a seminal role in the genesis of his thought, Darwin pointed out that a breeder could scarcely select a form which was not there; and that conversely the appearance of a surprising and unanticipated new form might give the breeder an idea as to what he could select for. Translating this to the case of "natural selection", this would mean that the definition of what the environment is for the organism, and what the organism is for the environment, are inseparably entwined; and that the organism in some sense "specifies" what will be there in the environment for the organism. This comes very close to constructivism; the irony is, of course, that in this case the sort of variation that can arise will play a major rôle in determining the course of evolution, so that natural selection is no longer all-powerful.

IV. Reflexivity: objectivism vs constructivism in the philosophy of science.

These latter remarks have brought me full circle in my examination of the intertwining threads of biology and cognitive science, objectivism and constructivism. I have noted, in the course of my argument, that cognitive science is perhaps the domain above all others where the issue of objectivism vs constructivism is inescapable, being as such a part of the object of the science in question; and that the thematic opposition between objectivism and a constructivism runs throughout the whole of cognition, extending in particular to that peculiar form of human cognition that we call "science". And since cognitive science is itself a science, the issue of reflexivity is inescapable: a paradigm in cognitive science is bound, willy-nilly, to pronounce on its own epistemological and ontological status.

Now as soon as it is recognized that science is a human social activity, constructivism comes fully into its own. As Bloor (1976) has pointed out, it would be arbitrary and quite out of keeping with the scientific tradition itself to exclude "scientific knowledge" as an object of science; but as soon as this is admitted, the "strong programme" in the sociology of knowledge is unavoidable. Scientific theories which are now held to be "false" (although in their time they were held to be "true") are not different in kind from theories that are currently held to be "true"; they are, alike, constructed[1]. Latour and Woolgar (1979) have provided clear empirical evidence on this point. They note that during the period when scientific knowledge is actually being generated, the hypothetical nature of scientific statements is clearly recognized by all concerned. However, on the rather rare occasions when a consensus forms in the relevant scientific community, and the "factual" status of a statement stabilizes (for a time at least), then two remarkable processes called "splitting" and "inversion" occur. Firstly, the statement "splits" and sends a perfect copy of itself into "the real world out there"; secondly, the relation between the statement and the "copy" is inverted so that the statement is presented, rhetorically, as the "copy" of the "real object out there". Thus, scientific facts have the peculiar property that the final stages of their construction consist of creating a "referential impression" which belies their own nature as social constructions; and the *adequatio rei et intellectus* becomes a miraculous matter for philosophical wonderment, whereas it actually derives quite simply from the fact that during its genesis, the "real object out there" first arose as a perfect carbon-copy of the scientific statement.

---

[1] This is the basis for Bloor's claim that methodologically, in the history and sociology of knowledge, it is appropriate to treat "true" and "false" theories "symmetrically". It is most important, here, to avoid confusion with rank relativism. I am definitely *not* saying that, at a given place and time, the statements "pigs can fly" and "pigs cannot fly" are equally valid; I am *not* saying that there is no difference between theories which have been refuted and theories which, so far, have escaped Popperian falsification. Every working scientist knows that the difference is very real, and indeed painstakingly constructed! The point, rather, is this. Everyone agrees that discredited theories, such as Lysenko's genetics, require a thoroughgoing social constructivist explanation in order to account for their very existence. When Bloor calls for a "symmetrical" treatment, he is claiming that "true" theories *equally* require a thoroughgoing constructivist explanation; they do not "explain themselves" by the simple virtue of being (putatively) "true". This argument becomes cogent when we consider that judgements as to whether a scientific theory is "true" or "false" often change (in both directions) over historical time.

What then is the situation in the confrontation between objectivism and constructivism as competing paradigms in cognitive science? It is a logical consequence of objectivism that if two theories are different , they cannot both be true (they may well of course both be false!). "True", in the context of representationalism, means isomorphism between the theory and referential reality; if both theories were "true" in this sense, they would both be isomorphic with the referent and therefore isomorphic also with each other - i.e. the theories could not be different. Thus, applying reflexivity in the domain of cognitive science, objectivism and constructivism cannot both be true. Now considered as rival theories in the domain of the sociology of science, objectivism and constructivism give rise to contrasting empirical predictions.

According to objectivism, the formation of a scientific consensus sufficiently long-lived to trigger "splitting", "inversion" and the creation of a referential impression is to be explained, very simply, by the circumstance that the scientific statement in question *really does* reflect pre-given objective reality. A necessary consequence of this is that the process can only happen once, and thereafter must be quite irreversible. The empirical prediction is therefore that scientific knowledge can only increase monotonically, as the range of isomorphism between scientific knowledge and reality is extended; once a scientific theory is held to be true, it can never thereafter come to be considered as false.

According to constructivism, the causal relation is the other way round: the formation of a consensus in the relevant scientific community (which is a contingent, psycho-social event) causes the illusory *appearance* that the "theory reflects reality". As in other areas of social life, the consensus may last for varying lengths of time; it may occasionally last for a long time on the scale of human history; but it will never be intrinsically immutable. The prediction is therefore that historically, scientific "facts" will change, sometimes radically; a theory that at one time was held to be true may dissolve into an "artefact" or a mistaken impression.

Put this way, the verdict is as clear as it ever is in scientific debate: the prediction derived from objectivism is decisively refuted by the empirical evidence. Scientific knowledge is remarkably mutable, more so in fact than most other human social institutions. The objectivist position, which in the domain of cognitive science is obliged to engage in reflexivity, is therefore self-refuting.

One might think that this is the end of the argument; but in fact the situation is far more subtly ironical than that. The objectivist paradigm in cognitive science is only self-refuting *on the objectivist premise* that if there are two different scientific paradigms, they cannot both be "true". However, on the constructivist premise this is not the case. "True" in this context means sufficiently coherent, and above all sufficiently productive of interesting scientific "puzzles" (Kuhn 1962), to be "viable" (i.e. to attract a critical mass of scientific resources). Constructivism is strongly sceptical as to whether the confrontation between any two contrasting paradigms can ever be solved by an apodictic, logical appeal to empirical "refutation". The Popperian criterion of "falsifiability", which functions well enough *within* a given paradigm, where the theoretical framework and the experimental methodologies are stabilized and pragmatically given, is quite inoperative when it comes to choosing *between* substantially different paradigms. Observations can *always* be re-interpreted, and the theory adjusted, so as to re-establish minimal coherence. Thus, multiple theories are no more incompatible than multiple species in biological evolution; which is not to say that species never compete for limited resources, nor that they never interfere with each others' eco-systems, in such a way that one of them goes extinct.

The irony, then, is that constructivism actually rehabilitates objectivism as a viable, pragmatic possibility in the field of cognitive science; whereas objectivism, left to its own arrogant devices, self-destructs.

V. Conclusions.

This article has addressed the issue of objectivism vs constructivism in two areas, biology and cognitive science, which are intermediate between the natural sciences such as physics on one hand and the human and social sciences on the other. Radical constructivism is well established in many areas of the human and social sciences, but to date has made little impact in the natural sciences. I have argued that the issues in biology and in cognitive science are intimately related; that in each of these twin areas, the "objectivism vs constructivism" issue is interestingly and rather evenly balanced; and that this opposition gives rise to two contrasting paradigms, each of which has substantial specific scientific content. I have argued that the neo-Darwinian paradigm in biology is closely resonant with the classical cognitivist paradigm in cognitive science, and that both of them are rather intrinsically objectivist. The organismic paradigm in biology, based on the concept of autopoiesis, is clearly consonant with the paradigm of "enaction" in cognitive science indeed in some formulations, the two paradigms are actually identical. I have argued that the latter pair of paradigms are both profoundly constructivist.

It is a part of the peculiar fascination of cognitive science that in this area, the objectivism vs constructivism issue can be thematized as an integral part of the content of the scientific field itself. In this field, then, the dimension of reflexivity is not "metaphysical" (as working natural scientists habitually tend to assume), but becomes *scientifically* relevant. I have argued that at this level, strong ontological objectivism is self-contradictory and therefore untenable. The situation with respect to constructivism is more subtle. I have argued that strong, radical constructivism is self-coherent. However, constructivism *also* rehabilitates a weak form of objectivism as a pragmatically viable alternative. The conclusion of this argument thus tends to promote a spirit of even-handed reciprocity between "objectivist" and "constructivist" perspectives. To end this article, I would like to briefly examine the consequences of this conclusion for fields other than cognitive science.

I shall first look at biology, which I have argued is closely akin to cognitive science and is a major focus of this article. For the neo-Darwinian paradigm, recognizing the *relative* nature of the rehabilitation involves acknowledging that neo-Darwinism has *in principle* a number of blind-spots (from outside the paradigm, a major example is organismic autopoiesis, but of course that is invisible from within). This could usefully serve, for example, to inhibit the temptation to indulge in strong forms of sociobiological reductionism or genetic determinism. Ideas such as: "the organism is just the genes' way of making more genes" are intellectually tenable *if but only if* they are clearly recognized as one possible but rather special way of looking at things among many others. My own opinion is that if such ideas are presented as "amusing, provocatively irrefutable but obviously rather bizarre oddities", which avoid the fatal mistake of taking themselves *too* seriously, their morally objectionable and socially pernicious aspects are significantly attenuated.

In the spirit of even-handedness, what are the consequences of my conclusion for an organismic biology of autopoiesis? Without exhausting the question, I see two possibilities. The first concerns the foundations of the autopoietic paradigm itself. Maturana in particular has a rather distressing tendency to take autopoiesis for granted, and to look only at its consequences (which are indeed considerable), while leaving unexplored the question of what is involved in achieving autopoiesis in the first place. Rosen (1991) has addressed the question of inventing a mathematical formulation of the concept of autopoiesis (or, in his own terms, of "closure under efficient cause); but although his theoretical considerations are admirably profound, they have not (as yet?) come up with a pragmatically operational solution. This undeniable weakness may go far to explain why the concept of autopoiesis has had (as yet?) so little impact on mainstream biology.

The second possibility is that the intellectual dynamic of the neo-Darwinian viewpoint (which is considerable, as witnessed by the energetic fervour of many of its proponents) may be a "blind-spot" from the point of view of the autopoietical paradigm. If this is the case (and a number of animated discussions with students and colleagues incline me to think that it is), then the relational configuration of neo-Darwinism and Autopoiesis is that each stands in the blind-spot of the other.

Moving from biology to physics, the prototypical "queen" of the natural sciences, I see some illustrative consequences of my "even-handed" conclusion as follows. In physics, objectivism is dominant to the point of not even being thematized as such. It seems to me that Prigogine and Stenghers (1979) strike the right note when they point out that the "natural laws" sought by physicist are not the definitive "secrets of the universe" as such; they are only ever *contingent* regularities and stabilities whose basis is obscure ("vague" as Whitehead would say) and ontologically ungrounded. For constructivism, which at the present time is marginal in this field of human activity, the most important "message" is probably an encouragement to take heart. For example, there *must* be at least one elegant solution to the problem of the interpretation of quantum mechanics - a typically constructivist endeavour. And if physicists such as Wheeler are right, this would lead to a much more profound understanding of why the laws of quantum mechanics - and indeed "natural laws" in general - are what they (contingently?) are.

Finally, there is the area of the human and social sciences where constructivism is currently often dominant. Again, I cannot hope to exhaust the question and it would be misplaced even to try. My suggestion here stems from the point that constructivism actually rehabilitates an suitably modest form of objectivism. This could function, usefully, as an antidote to some of the nihilist and relativistic excesses of post-modernism. From a constructivist point of view, it seems rather weird and perverse that so much fashionable attention should be paid to <u>de</u>construction; whereas the fruitful focus, surely, lies in the *positive* processes by which reality is constructed.

Acknowledgements.

This work was supported by Fellowship Praxis XXI/BC/14204/97 from the Portuguese Ministry of Science and Technology. I am grateful to the referees for their constructive comments, many of which are incorporated in this text.

References.

Bloor D. (1976). *Knowledge and Social Imagery.* London, Routledge and Kegan Paul.

Dawkins R. (1976). *The Selfish Gene.* Oxford, Oxford University Press.

Dretske F. (1988). *Explaining behaviour.* MIT Press, Cambridge.

Fodor J. (1975). *The Language of Thought.* Thomas Crowell, New York.

Gardner H. (1985). *The Mind's New Science: a history of the cognitive revolution.* Basic Books, New York.

Ho M.W. (1997). *The Rainbow and the Worm.*

Kuhn T.S. (1962). *The Structure of scientific revolutions.* Chicago University Press, Chicago.

Lakoff G. (1987). *Women, Fire and Dangerous Things : What Categories Reveal About the Mind.* University of Chicago Press, Chicago.

Latour B. & Woolgar S. (1979). *Laboratory Life: the social construction of scientific facts.* Sage, Beverly Hills.

Lenay C. (1999). *Darwin.* Les Belles Lettres, Paris.

Lenay C. (1993). Caractères adaptatifs et représentations symboliques. Intellectica **16**, 209-257.

Lewontin R.C. (1992). The Dream of the Human Genome. New York Review of Books.

Lovelock J. (1988). *The ages of Gaia.* Norton, New York.

Maturana H. & Varela F.J. (1980). *Autopoiesis and cognition: the realization of the living.* Reidel, Boston.

Medawar P. (1956). *The Uniqueness of the Individual.* Methuen, London.

Millikan R.G. (1984). *Language, Thought and Other Biological Categories: New Foundations for Realism.* MIT Press, Cambridge.

Oyama S. (1985). *The ontogeny of information: developmental systems and evolution.* Cambridge University Press, Cambridge.

Pacherie E. (1993). *Naturaliser l'intentionnalité.* PUF, Paris.

Prigogine I. & Stenghers I. (1979). *La nouvelle alliance. Métamorphose de la science.* Gallimard, Paris.

Putnam H. (1981). *Reason, Truth and History.* Cambridge University Press.

Rosen R. (1991). *Life itself: a comprehensive enquiry into the nature, origin and fabrication of life.* Columbia University Press, New York.

Saunders P.T. (1989). The evolution of biological forms. Leonardo **22**, 33-38.

Simondon G. (1964). *L'individu et sa genèse physico-biologique.* PUF, Paris.

Stewart J. (1996). Cognition = Life : Implications for higher-level cognition. Behavioural Processes 35: 311-326.

Varela F., Thompson E. & Rosch E. (1993). *The embodied mind.* MIT Press, Cambridge.

von Glaserfeld E. (1988). Introduction à un constructivisme radical. In: P.Watzlawick Ed., *L'invention de la réalité* . Ed du Seuil, Paris.

Waddington C. (1956). *Principles of Embryology.* Allen and Unwin, London.

Whitehead A.N. (1926). *Science and the modern world.* Cambridge University Press, Cambridge.