Contents

Impressum

Evolution and Cognition: ISSN: 0938-2623 **Published by:** Konrad Lorenz Institut für Evolutions- und Kognitionsforschung, Adolf-Lorenz-Gasse 2, A-3422 Altenberg/Donau. Tel.: 0043-2242-32390; Fax: 0043-2242-323904; e-mail: sec@kla.univie.ac.at; World Wide Web: http://www.univie.ac.at/evolution/kli/ **Editors:** Rupert Riedl, Manfred Wimmer **Layout:** Alexander Riegler **Aim and Scope:** “Evolution and Cognition” is an interdisciplinary forum devoted to all aspects of research on cognition in animals and humans. The major emphasis of the journal is on evolutionary approaches to cognition, reflecting the fact that the cognitive capacities of organisms result from biological evolution. Empirical and theoretical work from both fields, evolutionary and cognitive science, is accepted, but particular attention is paid to interdisciplinary perspectives on the mutual relationship between evolutionary and cognitive processes. Submissions dealing with the significance of cognitive research for the theories of biological and sociocultural evolution are also welcome. “Evolution and Cognition” publishes both original papers and review articles. **Period of Publication:** Semi-annual **Price:** Annuals subscription rate (2 issues): ATS 500; DEM 70, US$ 50; SFr 60; GBP 25. Annual subscriptions are assumed to be continued automatically unless subscription orders are cancelled by written information. **Single issue price:** ATS 300; DEM 43; US$ 30; SFr 36; GBP 15 **Publishing House:** WUV-Universitätsverlag/Vienna University Press, Berggasse 5, A-1090 Wien, Tel.: 0043/1/3105356-0, Fax: 0043/1/3197050 **Bank:** Erste österreichische Spar-Casse, Acct.No. 073-08191 (Bank Code 20111) **Advertising:** Vienna University Press, Berggasse 5, A-1090 Wien. **Supported by Cultural Office of the City of Vienna, Austrian Federal Ministry of Science/ Transportation and the Section Culture and Science of the Lower Austrian State Government.**

James Barham

Biofunctional Realism and the Problem of Teleology

Nothing is of any consequence to the knowing subject by itself; yet the will and the knowing subject are united in the I or ego. In every animal being the will has achieved an intellect, and this is the light by which the will here pursues its ends. — SCHOPENHAUER (1966, p500).

Abstract

Intentionality is a special case of biological functionality; therefore, in order to naturalize the former we must first naturalize the latter. To naturalize biological functionality means above all to account for its teleological character. Neither cybernetics nor natural selection can naturalize teleology because they both presuppose it; therefore, the reigning cybernetic-selectionist worldview *is radically incomplete, and the problem of teleology is alive and well. An alternative* dynamical-emergentist worldview *is outlined based on various principles of contemporary science. Against this background, the teleological and normative character of biological functionality may be viewed as a manifestation of a particular organization of bulk matter, on an ontological par with rigidity, superconductivity, or any other collective physical property (*biofunctional realism*). Thus, the functionalist doctrine of multiple realizability is false. The phenomenon of teleology is revisited accordingly, and found to consist of two intimately related but distinct structural features, namely,* conation *and* cognition. *A model of biological functionality incorporating these two features is sketched using concepts borrowed from nonlinear dynamics, and some of the wider philosophical implications of biofunctional realism are briefly considered.*

Key words

Cognition, conation, emergence, evolution, function, information, intentionality, life, nonlinear dynamics, teleology.

Introduction

Just as intentionality is the mark of the mental, functionality is the mark of the vital. Since minds are attributes of living things, beliefs and desires must be special cases of biological functions. If that is so, then in order to understand the place of intentionality in nature, we must first understand functionality. But what is a biological function?

Ordinarily, it is assumed that a function must be either a subsystem playing a certain sort of causal role within a larger system (CUMMINS 1998) or else a system with a certain kind of causal history (WRIGHT 1998). Since the kind of part-whole relation usually assumed by the former theory is *cybernetic organization*, and the kind of aetiology usually assumed by the latter theory is *natural selection*, I shall refer to these two views as the "cybernetic" and "selectionist" styles of explanation, respectively. While cybernetic and selectionist explanations were formerly assumed to be in competition, an irenic consensus is now emerging which views them as complementary (AMUNDSON/LAUDER 1998; GODFREY-SMITH 1999; MILLIKAN 1999). As ENÇ/ADAMS (1998) have pointed out, this approach to the philosophical understanding of functions dovetails nicely with the biological practice of pairing "proximate" and "ultimate" explanations (MAYR 1988a). By reinforcing each other in this way, these philosophical and biological complementarity principles are widely seen along a broad intellectual front as sanctioning the belief that biological functionality has been fully reduced to mechanistic causation. For convenience, let us call this the *cybernetic-selectionist (CS) worldview*.

So, if intentionality is a species of biological functionality; and if the CS worldview has reduced biological functionality to mechanistic causation; then it would seem that intentionality too has been reduced to mechanism, at least in principle. But now something seems to have gone wrong in our reasoning, for one of the most fundamental properties of intentionality is that of being about, or directed towards, something. This "aboutness" or directedness, in turn, creates conditions of satisfaction according to which mental states may be *evaluated*. Beliefs may be true or false, desires may be satisfied or unsatisfied, intentional states generally are directed towards their object as towards a goal which they may either succeed or fail in attaining. In short, the aboutness of intentionality seems to be essentially connected with *value* or *normativity*. So how can intentionality possibly be reduced to mechanism?

Well, why not? After all, it is generally agreed that normativity must be viewed as an essential feature of functionality as such, not just of intentionality—that is, the idea of a biological function essentially contains the idea of value in some way or other.[1] If that is so, then the problem of the natural ground of normativity cannot be separated from the general problem of *teleology*.[2] But in that case, since the CS worldview teaches us that the appearance of teleology in biological functions is an illusion, why can we not just say that the appearance of normativity in mental states is deceiving in the same way? Of course, there are some who do say just this (CANFIELD 1990; DENNETT 1987). However, it seems that most thinkers, though willing enough to view the behavior of amoebae, or digger wasps, or their own circulatory systems as purely mechanical, balk when it comes to mental states. There is a deep intuition at work here that nonnormative intentionality is a contradiction in terms, and that normativity cannot be reduced to mechanism. On the other hand, it is hard to see on what naturalistic grounds a principled distinction can be drawn between biological functions in general, and beliefs and desires in particular. To all appearances, hearts are just as purposive as minds; how, then, can we reconcile a fictionalist stance towards the one with a realistic stance towards the other? In this paper, I shall argue that the only way out of this dilemma is *to extend the realistic stance to biological functions as such*. This approach—which I call *biofunctional realism*—entails viewing the teleological and normative character of biological functionality as a manifestation of a particular organization of bulk matter, on an ontological par with rigidity, superconductivity, or any other collective physical property. Therefore, biofunctional realism explicitly rejects the functionalist doctrine of the "multiple realizability" of biological functions, including mental states.

Biofunctional realism is akin to, but distinct from, the position of some other recent critics of functionalism (e.g., T. NAGEL 1998; SEARLE 1992). These authors take a realistic stance towards mental states, but still cling to a fictionalist view of lower-level biofunctions. That is, they maintain that, while nonmental functionality is reducible to mechanism, intentionality requires an additional explanatory principle. To their credit, they recognize that the fashionable notion of "supervenience"[3] cannot be that principle. As Thomas NAGEL remarks, "pure, unexplained supervenience is not a solution but a sign that there is something fundamental we don't know" (1998, p344). He is also quite right to insist that what is required is a theory which *explains* (shows the nomological necessity of) the supervenience relation (*ibid.*, p347), and that the envisioned theory be "expansionist", not reductionist or eliminativist in character (*ibid.*, p343). What all of this points to is the need for a new approach to type-identity theory which builds context-sensitivity and normativity directly into brain processes themselves, i.e., which derives the functionality of mental states somehow from the brain's special form of material organization. However, neither NAGEL nor SEARLE gives any hint about how to reconcile such a physicalist yet nonreductionist theory of neural functions in particular with the standard reductionist view of biological functions in general. I believe that, if we are to have any hope of arriving at a coherent view of the mind's place in nature, their viewpoint must be extended beyond the brain to the living state as such. As BEDAU has noted, "[r]esolving how, if at all, life and mind are connected is one of the basic puzzles about life" (1998, p135). The nature of this fundamental connection between life and mind is the principal concern of this paper.

Understanding the nature of this connection depends, in turn, on arriving at a more adequate naturalistic conception of teleology which somehow transcends the mechanistic perspective. This does not imply a return to vitalism. In the words of E.S. RUSSELL: "I do not propose to revive the dead controversy between the mechanistic and the vitalistic theories of the living organism. *Neither point of view is satisfactory … Some quite different point of view is required*" (1946, p1, emphasis added).

The paper is organized into four main sections. In the first section, I begin by spelling out exactly what I take the "problem of teleology" to be. Next, I argue that neither cybernetics nor natural selection can account for the teleological character of biological functionality, since they both presuppose it. It follows that the problem of teleology is alive and well. Of course, it is not enough to demonstrate the inadequacy of the CS worldview; one must have something to put in its place. Therefore, in the third and longest section I attempt to outline an alternative *dynamical-emergentist (DE) worldview*, drawing on a number of empirical results and theoretical constructs from contemporary science. Finally, in the last section I revisit the problem of teleology, sketching a model of biological functionality based on concepts borrowed from nonlinear dynamics (see, also, BARHAM 1996), and reviewing very briefly some of the wider philosophical implications of the biofunctional realist perspective (BARHAM 1990, 1992).

The Problem of Teleology

By "teleology" I mean that property or power of living organisms by means of which mechanical or "efficient" causes are coordinated and made to cohere together so that a particular result, specifiable in advance, may be achieved. The problem, of course, is to give an adequate naturalistic account of this power. ARISTOTLE puts his finger on the heart of the matter in a well-known passage in which he refutes the claim of DEMOCRITUS and other *physiologoi* that living things can be completely accounted for by means of efficient causation alone (*De part. animal.* I.1, 640^b36–641^a15): "…a hand constituted in any and every manner, e.g., a bronze or wooden one, is not a hand except in name; and the same applies to a physician depicted on canvas, or a flute carved in stone. None of these can perform the functions appropriate to the things that bear those names … DEMOCRITUS's statement, therefore, needs to be qualified, or a carpenter might as well claim that a hand made of wood really was a hand. The physiologers, however, when they describe the formation and the causes of the shape of animal bodies, talk in this selfsame vein. Suppose we ask the carver "By what agency was this hand fashioned?" Perhaps his answer will be "By my axe" or "By my auger", just as if we ask the physiologer "By what agency was this body fashioned?" he will say "By air" and "By earth". But of the two the craftsman will give a better answer, because *he will not feel it is sufficient to say merely that a cavity was created here, or a level surface there, by a blow from his tool*. He will state the cause on account of which, and the purpose for the sake of which, he made the strokes he did; and that will be, in order that the wood might finally be formed into this or that shape. (ARISTOTLE 1961, pp67–69, emphasis added)

Just as the blows of the carpenter's tools are not haphazard, but are rather organized in a certain way, as the means for achieving a certain end—namely, the wooden hand—so too, ARISTOTLE saw, the efficient causes which produce the flesh-and-blood hand must be similarly organized. Of course, he had no way of knowing what those efficient causes might be; but, whatever they were, they would have to be organized in an analogous fashion, as means to end. How did he know this? Because, even less than a wooden hand, is it rational to consider a flesh-and-blood hand the accidental effect of unorganized causes.

The intuition that there is something special about living matter—something which no appeal to chance can account for no matter how many particles collide for how many aeons—remains as strong for us today as it was for ARISTOTLE. Indeed, we are better placed than he was to appreciate the reason for the peculiar power of the intuition of the "specialness" of living matter. Today, with the help of the analytical tools provided by probability theory and statistical mechanics, the notion of "specialness" can be made quantitatively precise. When even the simplest biological systems are analyzed using these techniques, they are found to be very special indeed in the sense that they occupy a small or "condensed" volume of their phase space (SKLAR 1993). It is easy to show that the probability of such systems' arising through stochastic processes alone is so minuscule as to amount to effective impossibility.[4] Of course, while statistical considerations can tell us *that* a particular system is special ("nonergodic"), they cannot explain *how* this specialness has come about—how it is physically possible. In order to do that, we must investigate the particular forces acting on and within the system in question.

In general, nonergodic behavior comes about through the entrainment of many individual particles into a collective or coherent mode as a result of the operation of one or more of the fundamental forces of nature. The different types of order we observe in the world are due to the interplay of different sets of constraints always acting so as to dissipate energy (i.e., minimize potentials/maximize entropy) in accordance with the second law of thermodynamics. By way of example, one may cite processes like phase transitions (condensation, crystallization) and vortex formation (whirlpools, hurricanes). However, while such dissipative processes are clearly a necessary condition for the existence of living things, they are not sufficient. We must be careful to distinguish true goal-directed processes from other superficially similar kinds of nonergodic processes. What is the difference?

First and foremost, the difference lies in a property of functional systems often referred to as *orthogonality,* which Ernest NAGEL has called "a formal criterion for distinguishing processes that are goal-directed from those which are commonly held not to be such" (1998, p211). "Orthogonality" refers to the way the different efficient causes acting within a functional system are related to each other. Causes are said to be orthogonal with respect to one another if they are independent of each other insofar as the laws of physics are concerned—that is, if the existence of one cause does not necessitate the existence of the other. The importance of this basic insight has been acknowledged by numerous authors using various terminologies. For example, MONOD (1972) had much the same thing in mind when he spoke of the "gratuity" of the relationship between the chemistry and the physiology of biological macromolecules, as did POLANYI (1969) when he maintained that the operation of "living mechanisms" is constrained by "boundary conditions" which transcend the laws of physics and chemistry. (See, also, SWENSON 1998.) If two variables of a functional system are orthogonal to each other, then any value of one variable is physically compatible with any value of the other variable, which means that there is something about the means-end relationship of goal-directed behavior that transcends forces and energy potentials. Because of the orthogonality requirement, it is necessary to look beyond applied forces, the second law of thermodynamics, or the principle of least action to fully explain functionality.

It is often observed in this connection that because living things are thermodynamically open systems there is no violation of the second law of thermodynamics, since any entropy decrease internal to the system is more than balanced by an entropy increase in the environment. But to point out that living things do not violate the second law is like observing that birds do not violate the law of gravity. Birds may not violate the letter of the law of gravity, but they violate its spirit. In the same way, organisms violate the spirit of the second law. Just as a complete explanation of flight required the development of a science of aerodynamics in addition to knowledge of the law of gravity, so too a complete explanation of biological functionality will require the development of a science of teleology in addition to knowledge of the second law. In short, the problem of teleology is the problem of finding *the biological equivalent of lift.*

The Poverty of the Cybernetic-Selectionist Worldview

The question is, then, whether cybernetics and natural selection together constitute a theoretical structure capable of solving this problem. Let us attempt to answer this question by examining each pillar of the CS worldview in turn, beginning with natural selection.

Natural selection

Given a population of organisms with a set of stochastically variable, heritable biological functions ("traits"), the theory of natural selection tells us that any trait-token which confers some competitive advantage on its bearer (organism-token) relative to a given environment will tend to proliferate within the population in succeeding generations until it becomes the dominant trait-type for that organism-type. This process is deemed to explain how trait-types come to be coordinated with the environmental conditions for which they are appropriate, thus providing a mechanistic explanation of a seemingly teleological phenomenon. But have we really succeeded in reducing teleology to mechanism, or have we only smuggled the teleology in surreptitiously somewhere along the way?

Selection-style explanations have two fundamental failings with respect to the problem of teleology. The first is obvious. Selection theory always presupposes the existence of something to select; therefore, it cannot constitute the complete explanation of a trait all by itself. Of course, we can always simply *assume* that biological functions arise via some unknown mechanistic process or other, but then it is our assumption which has accomplished the "reduction", not the selection process itself. The second failing is that there is nothing within selection theory as such which explains what a "comparative advantage" is. It is no use attempting to explain the comparative advantage of traits in terms of differential reproduction, for that would be viciously circular. Rather, we must explain differential reproduction in terms of the comparative advantage conferred by traits understood as intrinsic dispositions (the "propensity theory of fitness"). What this means, in practice, is relative success in *functioning* in an "ecological" or "engineering design" sense (BRANDON 1990; VAN DER STEEN 1994—see, also ROSE/LAUDER 1996; SCHMIDT-NIELSEN 1997; WAINWRIGHT/REILLY 1994; WEIBEL, 2000; WEIBEL et al. 1998). As one of the architects of the propensity theory has written, "[t]he manner in which any type of organism achieves high fitness is ultimately a matter of the physiological, anatomical, and behavioral traits that underlie its viability and fertility and in turn underlie its overall descendant contribution ability" (BEATTY 1992, p117). Why is this a failing of selection theory? (The propensity theorists themselves do not see it that way.) The trouble is, simply, that *the problem of teleology has been begged*. That is, there is nothing which has explained how biological traits differ from other kinds of "mechanisms". Selection theory does nothing to help us understand what it is about functions that makes it appropriate to speak of their "advantage", "benefit", "utility", etc. for their bearers. Natural selection is like a conveyor belt which *transmits* a biological trait or function from one generation to the next. If one assumes that a trait-token is mechanistic to begin with, then the trait-type will be mechanistic at the end—mechanism in, mechanism out. But, as we have already seen, in that case it is the presupposition that is doing all the work of mechanization. Natural selection has no power to magically *transform* a mere effect into a true function. At best, selection theory explains how one kind of success (relative functioning) gets turned into a different kind of success (differential reproduction). But natural selection cannot explain how the *capacity* of biological functions for success or failure arose out of physics in the first place, for the simple reason that the selection process has no hand in constituting biological traits *as* functions. Furthermore, selection's accomplice—"random variation"—is no help either, since even the simplest organic macromolecule is so thermodynamically improbable that invoking "chance" to explain it is tantamount to invoking a miracle (FRY 1995).

But if selection theory cannot account for teleology on its own, perhaps it can do so in tandem with a suitable "proximate" or physiological theory of biological functionality. In fact, several prominent evolutionary theorists have argued that natural selection must be harnessed to the science of cybernetics in order to carry the reduction of teleology through to completion (e.g., MAYR 1988b; MONOD 1972; G.C. WILLIAMS 1992). So let us take a closer look at the other pillar of the CS worldview.

Cybernetics

In order to understand the cybernetic style of explanation, let us look again at the analysis of Ernest NAGEL (1998). (For further discussion, see especially BENNETT 1990, and WOODFIELD 1976; for entrée to the older literature, see GEORGE/JOHNSON 1985.) According to NAGEL, there are two fundamental aspects of cybernetic organization, i.e., two requirements for a mechanical system to qualify as being functionally organized (E. NAGEL 1998, pp208–213 *et passim*). The first condition is that the mechanical causes which constitute the system must be *directively organized*. The notion of "directive organization" refers to the fact that the system has a special state, specifiable in advance, which the efficient causes operating in the system collectively produce. This directive organization is further characterized by the properties of "plasticity" and "persistence"—that is, the efficient causes are capable of converging on the special state by "following alternative paths or starting from different initial positions" (*ibid*., pp208–209), and by "compensat[ing] for any disturbances taking place (provided these are not too great) either within or external to the system, disturbances which, were there no compensating changes elsewhere, would prevent the realization of the goal" (*ibid*., p209). According to this view, then, it is the directive organization of efficient causes which constitutes a particular physical state of a system *as* a goal state; goals have no reality apart from directive organization. The other requirement for functional organization, as we have seen, is the *orthogonality* of the efficient causes with respect to each other. In essence, the orthogonality requirement means that in order for a system to count as genuinely functional, its directive organization cannot be explained by reference to the laws of nature (*ibid., p*211). But in that case, how *is* directive organization to be explained?

One way the cyberneticist might attempt to answer this question is by invoking the concept of *information*. A goal-directed system is one which is capable of being "informed" about its own states via feedback, and of acting on the basis of such information so as to maintain the goal state. With the help of information, then, efficient causes can be made to converge on a particular goal state even while being orthogonal to each other insofar as the laws of physics are concerned. Unfortunately, this move only succeeds in replacing the mystery of the source of directive organization with an even deeper mystery: What is information?

"Information" is one of the most vexed concepts on the contemporary intellectual scene. There are two main ways in which the term is abused, with equally deleterious results. One way reifies classical SHANNON information by elevating it to a fundamental physical principle on a par with matter and energy (FRIEDEN 1998; STONIER 1990; J. A. WHEELER 1990). The trouble with this tactic is that SHANNON information can only be operationally defined by attaching a semantic interpretation to it, whether explicitly or implicitly. That is to say, no physical pattern or structure can properly be called "information" except insofar as it is *meaningful*. In other words, nothing *counts as* information except in relation to the goals or interests of some agent. Thus, information is itself an intrinsically teleological concept; as such, it can scarcely serve as one of the basic building blocks of the universe, at least within a naturalistic perspective. The other way in which the notion of information is regularly abused is by anthropomorphizing it, whether by restricting its scope unduly to human beings (SEARLE 1992), or by reducing it to a fiction or projection of the human mind (DENNETT 1987) or a "language game" (CANFIELD 1990). The trouble with this tactic is that it ignores the fact that information—in some sense of the term—is intimately involved in all living processes at all levels (BHALLA/IYENGAR 1999; BRAY 1995; LOEWENSTEIN 1999; MIKULECKY 1996; PATTEE 1986; WENG et al. 1999; YATES 1997; YATES/KUGLER 1984), and thus has no essential connection to human interests.[5] Does this not contradict what was just said about information's being relative to an agent's goals? It does not. The reason is that, while information is essentially *agent*-relative, this does not mean that it is also *observer*-relative.

Equivocation on this point has given rise to much confusion in the literature. In order to attain clarity, I believe it is necessary to distinguish carefully between two senses of the word "information" based on the recognition that, while meaning is always relative to the goals of an agent, the agent need not be human. That is to say, we must make a distinction between, on the one hand, information *for* an organism that is meaningful with respect to the goals of that organism itself, and, on the other hand, information *about* an organism that is meaningful with respect to the goals of a human observer (or another organism). For convenience, I shall call the first, intrinsic sort of information *autotelic,* and the second, extrinsic sort *allotelic.* While allotelic information is indeed observer-relative, and hence subjective, autotelic information is completely independent of any outside observer's interests, and so is fully objective.

Now, the information utilized by biofunctions is internal or intrinsic to those functions themselves, hence autotelic. Therefore, any theory of functional organization must be able to give some account of how autotelic information arises out of mere physics. But here we have reached the limits of the usefulness of cybernetics. Cybernetics cannot possibly hope to explain the source of information in the autotelic sense, since one of its fundamental postulates (the orthogonality requirement) states that there is no intrinsic connection between cybernetic organization and physics. This means that for cybernetics information can only be understood in the extrinsic or allotelic sense. Which is why every effort to naturalize intentionality in the manner of computationalist cognitive science inevitably commits the homunculus fallacy. As SEARLE puts it, "[w]ithout the homunculus, there is no computation, just an electronic circuit" (1992, p221). (See, also, BICKHARD/TERVEEN 1995; HORST 1996; PATTEE 1993; R. ROSEN 1991.) Nor are the other cybernetic concepts encountered in the literature (e.g., "feedback", "control") of any avail, since they also presuppose an independently determined goal state. As WIMSATT has pointed out, "it cannot automatically be assumed that feedback is a single precise theoretical concept and an objective property of a certain class of systems" (1985, p179). In cybernetic thinking, the notions of feedback and information are interdefined and mutually support each other; neither is more basic than the other.

The only way to break out of this vicious circle of teleological concepts is to put information back into its biological context—in the form of autotelic information that is intrinsically meaningful for the organism. As CAO has observed, "the essence of neural activity is meaning rather than information" (1998, p44). We may summarize the foregoing argument in the form of a maxim: *no meaning, no information*. It is for this reason that cybernetics begs the question of teleology—because it ignores the problem of the meaning of information.

What is true of mental states is also true of other biological functions. To see that this is so, let us consider a concrete case—that staple of the philosophical literature, the heart. The heart is, of course, a pump.[6] Pumps are machines. And if the heart is a machine, is that not an existence proof of the reducibility of teleology to mechanism? It is not, since a mechanical pump has no intrinsic goal-directedness and can only be called a "heart" relative to the human being whose interests it serves. To be sure, the parts of the pump may be directively organized in order to produce the goal state "automatically", but there is nothing intrinsic in the pump itself which supports this directive organization. It is we who impose the directive organization on the efficient causes operating within the pump, we who determine its goal state. Or, rather, it is we who determine what particular concatenation of efficient causes shall *count as* a goal state. Therefore, we have things exactly backwards when we attempt to explain biological functions by analogy with machines. Far from the biofunction's being a natural machine, it is really the machine that is an artificial function—that is, organisms are not so much mechanical as machines are biomimetic. Since machines are nothing but congeries of mechanical causes directively organized by organisms for their own purposes, it is circular to invoke machines in the attempt to reduce the teleological character of organisms to efficient causes.

The fundamental difficulty we face is this. According to cybernetics, the normativity inherent in biological functionality has no reality apart from the directive organization of efficient causes, and yet *there is nothing in cybernetic theory itself which explains the source of directive organization.* We have already seen that it does not help the cyberneticist to invoke "information" here, since for him all information is allotelic, hence question-begging. There is no getting around the fact that organisms are organized somehow from within, while machines are organized from without. This crucial difference is sometimes summarized by saying that the directive organization of organisms is "adaptive" or "robust", while that of machines is "nonadaptive" or "brittle" (on the difference between organisms and machines, see DREXLER 1989; FONTANA et al. 1995; R. ROSEN 1993). The brittleness of machines is ultimately traceable to the second law of thermodynamics. By this I mean that the pump, for example, has no intrinsic ability to oppose the inevitable erosion of the coherence and coordination of its efficient causes, since its functional integrity is merely imposed on it from without as a set of boundary conditions. The heart, in contrast, actively combats such erosion, since all the efficient causes in it are continually cooperating to preserve its functional integrity from within. From the point of view of philosophical mechanism—with its vision of the world as a ramshackle Rube GOLDBERG contraption—such behavior is utterly mysterious. In order to understand the robustness of biofunctions, we must delve much more deeply into the physics of the living state.

If the machine analogy fails absent an explanation for the directive organization in organisms—if, in essence, the concept of directive organization simply begs the question of value or normativity—then what is to be done? Ernest NAGEL has a neat solution to the problem. He writes that "[g]oal-directed processes in living systems are patently programmed, containing 'instructions' for the development (among other things) of 'feed-back' subsystems; *and the origins of the programs are left to be explained by evolutionary theory*" (E. NAGEL 1998, p210, emphasis added). Thus does cybernetics pass the teleological buck back to natural selection. Nor is NAGEL alone in performing this sleight-of-hand trick. Unfortunately, it has become the stock-in-trade of contemporary naturalistic philosophy. For example, consider the ongoing debate between two of our leading philosophers of mind—selectionist Ruth MILLIKAN and cyberneticist Jerry FODOR. MILLIKAN accuses FODOR of begging the question of normativity, noting that it is only "through confusion of a biological norm with a simple disposition or fact [that] causal/informational theories achieve their initial plausibility" (1993, p7). For his part, FODOR tartly observes, in commenting on MILLIKAN's selectionist account of intentionality, that "DARWIN has nothing to say to BRENTANO; the whole point of DARWIN's enterprise was to get biology out of BRENTANO's line of work" (1990, p79). Well, MILLIKAN and FODOR are both right—that is, they are both wrong. Neither natural selection nor cybernetics can naturalize normativity for the simple reason that they both presuppose it (cf. MANNING 1997). But how can cybernetics and selection theory explain together what neither of them can explain separately? The CS worldview as a whole only maintains its façade of plausibility by operating an intellectual shell game: No matter which theory you look under—cybernetics or natural selection—the final reduction of teleology to mechanism always occurs under the other one.

Towards a Dynamical–Emergentist Worldview

We are faced with a seemingly insoluble problem. The essence of functionality lies in orthogonality—that is, in the irreducibility of the directive organization of efficient causes to natural law. And yet to naturalize something would appear to consist precisely in showing how it acts in accordance with natural law. I believe that *the only way out of this impasse is to relax the implicit identification of natural law with mechanistic causation.* This is admittedly a radical step, but it is one for which there are many indications that the time is now ripe.

On the one hand, one of our most distinguished philosophers of mind recently expressed his conviction that "the [mind-body] problem goes deep, deep into our fundamental metaphysical views about ourselves and the world we live in, and that we need to make fairly drastic adjustments if we are serious about coming to terms with the problem" (KIM 1998, p59). On the other hand, the great theoretical biologist Robert ROSEN not long before his recent death observed that there is "a real sea change in science today; a general increase in conceptual temperature which is liquefying outmoded doctrines which have hovered around absolute zero for the past half-century or more" (R. ROSEN 1996b), while a prominent contemporary theoretical physicist has written that "[t]o find a period previously accepted assumptions were questioned as deeply as they are now, one must go back four centuries" (ROVELLI 1997, p180). This happy conjunction of a growing philosophical recognition of the need to overhaul our fundamental metaphysical categories and a very high level of ferment in both physics and biology presents us with a golden opportunity to canvass new scientific insights for ways to make the needed philosophical adjustments. (For a penetrating, scientifically and philosophically sophisticated analysis of our present historical situation, see LECLERC 1984.)

At this point, I would like to make explicit certain metaphysical principles which underpin biofunctional realism. First is a rough-and-ready, commonsense *realism* about the objective existence of the world (knowing is a species of being, and not the other way around). This is really an absolute presupposition necessarily shared by all living things for which it would be otiose to adduce arguments (SANTAYANA 1955). However, there are three other principles which will be more controversial, but which I believe must be granted if we are to arrive at a proper understanding of our place in nature: namely, *temporalism*, *emergentism*, and *multicausalism*.

By *temporalism* I mean the idea that the universe as it is in itself is in process of becoming, i.e., that the arrow of time is objectively real and not merely a human projection. Although the belief that "the irreversibility within the cosmos is fully objective in the sense of being independent of man's presence" (DENBIGH/DENBIGH 1985, p118) would appear to be an essential element of our animal-faith realism, temporalism remains contentious because it conflicts with the theoretical structure of fundamental physics, most of whose formalisms are time-symmetric. It is true that some tentative proposals for eliminating this paradox have been made (e.g., PETROSKY/PRIGOGINE 1990; PRIGOGINE 1997), but there is as yet no consensus on how to construct an integrated theory of reversible and irreversible processes. Therefore, while an eventual solution to this problem is indispensable to a coherent DE worldview, there is little we can do at present except acknowledge the problem, go with common sense, and move on.

Emergentism is the idea that over time the universe has acquired a hierarchical structure of levels with qualitatively novel properties. If emergentism is true, then lower-level laws are in principle insufficient to account for higher-level phenomena (reductionism is false). Emergentism is widely known and discussed, but closely related to it is a much less widely known but equally important principle which I shall dub *multicausalism*—the idea that the novel properties of emergent levels give rise to fundamentally different kinds of causal processes, such that the universe contains a richer variety of causes at later times than in earlier epochs. If multicausalism is true, then other causal modalities may exist in nature besides the linear or mechanistic type (determinism is false).

While there is now a sizable philosophical literature on emergence in which many variants of the basic concept are distinguished and their history traced in detail (BLITZ 1992; MCLAUGHLIN 1992; STEPHAN 1992), I shall here be more concerned with the scientific foundations of a dynamical and emergentist worldview. As MEEHL and SELLARS have remarked, "the question whether the world is to be conceived along emergentist lines is a scientific question which cannot be settled on a priori grounds" (1956, p239). The question is not so much whether an adequate conceptual analysis of emergence can be found, as whether contemporary science sanctions an emergentist-multicausalist or a reductionist–determinist perspective. Therefore, I will review a number of scientific findings which I believe show that the old faith in reductionism and determinism is no longer justified. Though by now fairly well established, some of these developments are not as widely known as they should be. By presenting them here, I do not mean to suggest that their philosophical interpretation is settled, only that they are suggestive of a new DE worldview which provides a more congenial metaphysical framework for making sense of biological functionality. I will now review these considerations in turn as they relate to the phenomena of emergence and multicausality.

Emergence

It is by no means the case that all practicing chemists and condensed-matter physicists believe that their objects of study are epiphenomena which might in principle be "reduced" to fundamental physics; on the contrary, expressions of support for a broadly emergentist position are quite common. For example, FISHER points out that "we see in polymeric matter new, subtle and universal behavior which we have succeeded in understanding theoretically. But quantum mechanics has had essentially nothing to say about the problem!" (FISHER 1988, p74). In a similar vein, DRESDEN writes that "the behavior of large aggregates of elementary atoms should not just be understood in terms of a simple extrapolation of the properties of the system with just a few atoms. There seem to be many levels of complexity, and at each such level entirely new properties begin to appear" (DRESDEN 1974, p161). ANDERSON goes still farther, dubbing the principle of emergence "the fundamental philosophical insight of twentieth century science" (ANDERSON 1995, p2020), and insisting that "everything we observe emerges from a more primitive substrate, in the precise meaning of the term 'emergent', which is to say obedient to the laws of the more primitive level, but not conceptually consequent from that level ... philosophically such a structure as the Standard Model, or the laws of chemical bonding, breaks the chain of reductionism and makes further delving into the underlying laws somewhat irrelevant to higher levels of organization" (*idem*). Of course, such views might be regarded merely as expressions of current limitations on our knowledge. In that case, no ontological inferences could be drawn from them. However, in recent years some new developments in fundamental physics itself have lent support to an ontological interpretation of the emergent hierarchy.

One of these is the prominence accorded the notions of symmetry and spontaneous symmetry breaking in quantum field theory. The deep connection between symmetry[7] and order was already clearly enunciated more than a century ago by Pierre CURIE, who wrote that "...certains éléments de symétrie peuvent co-exister avec certains phénomènes, mais ils ne sont pas nécessaires. Ce qui est nécessaire, c'est que certains éléments de symétrie n'existent pas. *C'est la dissymétrie qui crée le phénomène*" (1908, p127, emphasis added). CURIE's idea was taken up a half-century later by physicists working on the theory of condensed matter and the foundations of quantum field theory (BROWN/CAO 1991). Today, the concept of spontaneous symmetry breaking has become a cornerstone of fundamental physics and cosmology, and has given rise to a grand vision of cosmic evolution as the gradual emergence of the hierarchical structure we observe today. (For a summary of the current view of cosmic evolution, see HOGAN 1998.) As CAO and SCHWEBER have put it:

"The profound understanding of these implications of [spontaneous symmetry breaking] has provided the strong impetus to search for an ultimate unified description of nature, in which natural laws with different invariance properties, symmetrical theories, and asymmetrical physical states all emerge from the highest symmetry that characterizes physics under the conditions present in the early universe, passing through a sequence of phase transitions as the temperature decreases while the universe expands until it reaches the state described by [quantum chromodynamics] and the electroweak theory." (CAO/SCHWEBER 1993, p57)

The concept of spontaneous symmetry breaking not only unites elementary particle physics with cosmology, it also connects both of these with condensed-matter physics[8] since many analogous symmetry breakings occur at lower temperatures which give rise to the macroscopic collective behavior of matter in the solid and liquid phases. Some of these phase transitions are familiar features of everyday life (evaporation and condensation, melting and freezing), while others are the province of the laboratory (lasers, liquid crystals, superfluids, BOSE–EINSTEIN condensates). Phase transitions and symmetry breaking have also played an important role in a number of recent theories of the selforganization of complex systems, from inorganic coherent structures (HAKEN 1983) to the origin of life (KAUFFMAN 1993), protein folding (WOLYNES 1991), metabolic function (DEL GIUDICE et al. 1989), embryonic development (WEBSTER/GOODWIN 1996), and brain function (KELSO 1995). Thus, the notion of symmetry breaking provides us with a fundamental framework for understanding the growth of order in the universe as ultimately driven by the cosmic expansion (LAYZER 1990). It is a "punctuated equilibrium" model of universal structuration in which continuous changes in variables at a given length scale lead to the discontinuous creation of new macroscopic variables at longer length scales. From crystallography to quantum field theory, one of the chief lessons of contemporary physics is that quantitative change gives rise to qualitative novelty, or that—in ANDERSON's (1994) well-known maxim—"more is different".

Another recent development in fundamental physics that lends support to the emergentist point of view is the rise of the effective field theory viewpoint. In its original version, quantum field theory was conceived of as a single, all-embracing theory like the older quantum mechanics which might provide the basis for a universal reductionism, at least in principle. However, in recent years a number of ideas (e.g., scale invariance, the decoupling theorem, cutoff, and the renormalization group) have given rise to a very different interpretation that has come to be known as the "effective field theory" program (CAO 1997). An effective field theory is a quantum field theory tailored for a particular length and energy scale that is semi-autonomous with respect to physics at higher energies and shorter lengths (GEORGI 1989). For each such theory there is a cutoff defined beyond which the theory breaks down and new physics is assumed to occur requiring a new effective field theory. Collectively, the various effective field theories thus partition reality into a "tower" of qualitatively distinct physical regimes. At each level of the tower macroscopic variables arise specific to that length and energy scale which decouple from the physics of the other regimes. These dimensional parameters have to be determined empirically and put into the equations by hand. At the same time, however, all effective field theories are also unified in the sense that they are invariant under a group of scale transformations of the dynamical or field variables (the renormalization group transformations). In this way, the principle of scale invariance connects the various regimes together again into a single coherent conceptual structure.

The support afforded to emergentism by the effective field theory program is especially strong due to the fact that many of its fundamental concepts were originally imported into elementary-particle physics from statistical mechanics and condensed-matter physics (K. G. WILSON 1983). CAO has explained this connection in the following passage:

"In statistical physics, the renormalization group approach is powerful in establishing connections between physics at different scale levels, achieving conceptual unification of various complicated systems, such as those of elementary excitations (quasiparticles) and collective ones (phonons, plasmons, spin-waves), understanding the universality of various types of critical behavior, and calculating order parameters and critical components, by scaling out the irrelevant short-range correlations and finding a stable infrared fixed point. In [quantum field theory], the same approach can be used to suppress the irrelevant low-energy degrees of freedom, and to discover a stable ultraviolet fixed point. In both cases, the essence of the approach … is to concentrate on the relevant degrees of freedom for a particular problem, and the goal is to find fixed point solutions of the renormalization group equations." (CAO 1993, pp114–15)

Needless to say, like any scientific theory, effective field theory is susceptible to an instrumentalist interpretation. However, by the same token, a realistic interpretation is equally available, and I believe preferable on both physical and general philosophical grounds. As AUYANG has observed, "*The world exhibits many levels of scale and complexity that require radically different descriptions*. The classical and the quantum are two such levels, and there are many more beyond physics. Ontologically, we can agree that all systems, no matter how large and complicated, are made up of subatomic constituents. However, this does not imply that the theory for the subatomic constituents is applicable to all systems as integral units. Those who assume that it does have neglected the effect of composition and confused the properties of the parts with that of the whole … There is no universally unified theory that accounts for all levels of complexity" (1995, p83, original emphasis). It is hard to believe that the striking success of this approach across so many different length and energy scales, from quarks to crystals, is an accident. As always, the most plausible explanation for the success of our theories is realism. I believe that we are entitled to conclude that effective field theories probably have a grip on the way the world really is in itself.

The effective field theory program, realistically interpreted, serves as an excellent foundation for the most plausible of the various versions of emergentism—namely, "weak emergence". I mean by this that it gives us a sense of how it is possible for the levels of reality to be both causally interconnected and yet partially autonomous. As the distinguished historian and philosopher of physics, Sylvan SCHWEBER, has put it:

"The hierarchical picture of the physical world implied by effective field theories explains why the description at any one level is so stable and is not perturbed by whatever happens at higher energies, and thus justifies the use of such description. In this hierarchical depiction each level is populated by its own ontology, yet the picture recognizes connections between levels through what are known as the renormalization group equations. It, thus, allows the possibility for the emergence of complexity and novelty without rejecting the possible description of the mechanism of emergence in terms of component parts." (SCHWEBER 1997, pp179–180).

The doctrine of weak emergence opposes the radical reductionism of strict logical entailment of higher- from lower-level events and properties—thus affirming the reality of creativity in the universe—while at the same time recognizing the causal role of lower-level influences in producing the definiteness of novel higher-level phenomena (BEDAU 1997; BUNGE 1991; WIMSATT 1997). If we take radical reductionism to be the thesis that knowledge of the parts is both necessary and sufficient for understanding of the whole, and holism to be the converse thesis that it is neither necessary nor sufficient, then weak emergence "occupies a middle position between these two extremes with its thesis that knowledge of the parts is necessary but not sufficient for the comprehension of the whole" (BLITZ 1992, p177). Weak emergence is more or less identical to what is coming to be called "global reductionism" in the field of integrative biology (RIPOLL et al. 1998), although I personally prefer to eschew the word "reductionism" altogether, due to its deep-seated mechanistic connotations.

To be sure, emergentism in whatever form remains highly controversial insofar as it is advanced as a general metaphysical thesis. Critics sometimes dismiss the very idea of an emergent property as "magic" or "pixie dust". However, VAN CLEVE is no doubt right when he claims that such attitudes are more often than not the result of misunderstanding:

"I suspect that a good deal of opposition to emergent properties as inexplicable is simply the product of terminological confusion. People slide in their thinking from "emergent" in the sense of "not following logically from any properties of the parts" to "emergent" in the sense of "not following in *any* manner from properties of the parts", and that makes emergence seem like magic. But there is no reason why emergent properties can't follow with *nomological* necessity from properties of the parts, and in that case they would not be inexplicable for anyone but a causal rationalist". (VAN CLEVE 1990, p218, original emphasis)

Understood in this way, weak emergence is not far removed from that older form of thought which went under the rubric of "integrative levels". As NOVIKOFF, one of the foremost exponents of the integrative levels viewpoint, explained: "Knowledge of the general qualities of development common to all levels of organization of matter will aid in the analysis and description of the concrete attributes of each level. But it can not be a substitute for such analysis or for the determination of the qualitative uniqueness of each level and the characteristic laws which govern it" (1945, p210).[9]

It is hard-won scientific insight, not conceptual analysis, that is our surest guide to reality. In the final analysis, then, the argument in support of an emergentist worldview comes down to this: If we are inclined to realism about any of the theoretical constructs of natural science, then there seems no good reason why we should make an exception in the case of emergent levels. Many new developments in contemporary science—especially spontaneous symmetry breaking and the effective field theory program—indicate that emergent levels are just as real as atoms or molecules or crystal lattices, and that, indeed, the latter can only be fully comprehended in terms of the former. If that is so, then we are entitled to conclude, with REDHEAD, that "[t]he reductive hierarchy founding the whole of science on the properties of individual elementary particles [has been] shown to be mistaken" (1995, p62). (For further discussion, see CAO 1998; K. G. DENBIGH 1975; PRIMAS 1991; THIRRING 1995. See, also, BATTERMAN 2000, a very important work of which I have become aware only when the present paper was already in the proof-stage.)

Multicausality

Quantum field theory provides us with a framework for understanding the emergence of order or stability at lower levels of the emergent hierarchy. However, there is as yet no effective field theory of life; therefore, we must content ourselves for now with whatever bodies of knowledge we can find which throw light on the relevant phenomena at the mesocosmic length scales and energies characteristic of living things. If we turn our attention to the emergent level at which the transition from nonlife to life occurred, there are two basic questions we may ask ourselves. First: What was the nature of the previous level which must have formed the immediate causal background out of which life emerged? That is, what is the generic form of stability which life possesses? Second: What is the nature of the threshold level itself? What is the differentia that marks off life as a distinct species within its genus (what we called above the "biological equivalent of lift")?[10] Let us turn now to the first of these questions, saving the second question for the next section below.

To begin with, it is essential to keep in mind that living things belong to the class of thermodynamically open systems. An open system is one which exchanges both matter and energy across the boundary separating it from its surround. One of the fundamental insights gained in recent decades from work in the discipline of nonequilibrium thermodynamics is that energy fluxes that are constrained in such a way that throughput exceeds relaxation rate will cause the system in question to "selforganize", where "selforganization" means the spontaneous appearance of material cycles or oscillations in a dynamical steady state (so-called "dissipative structures"). The underlying physical reason for the selforganization of dissipative structures is the fact that, given the constraints imposed by the material constitution of the system, energy dissipation occurs faster via coherent cycling than it would by means of incoherent thermal diffusion (KONDEPUDI/PRIGOGINE 1998; MOROWITZ 1979). Commonly cited examples of dissipative structures include hurricanes, candle flames, the Red Spot of Jupiter, BÉNARD cells, and the BELOUSOV-ZHABOTINSKY reaction. The oscillatory cycles that are ubiquitous in living things (GOLDBETER 1996) are also widely viewed as belonging to the same class of phenomena. This tendency of constrained energy fluxes to spontaneously selforganize is such a profound aspect of nature that some theorists in the field have taken to referring to it as the "fourth law of thermodynamics" (KAUFFMAN 1996, *passim*; MOROWITZ 1992, p77).[11]

Against this background, we can begin to see the generic form of stability which biological functions must possess. Rather than "mechanisms", biofunctions are better conceived of as *nonlinear oscillators* (COHEN/RICE 1996, p242). Drawing on the conceptual apparatus of nonlinear dynamics,[12] we may define a nonlinear oscillator as a steady-state dissipative structure whose behavior may be described by a *phase-space trajectory* of a certain kind: either a *limit-cycle* or a more complicated, higher-dimensional *attractor*. Such attractors possess the mathematical property of *equifinality*—i.e., there exists a many-to-one mapping of virtual initial states of the system onto a single final state. Such attractors are robust, in that perturbations within a certain range of magnitude will be spontaneously damped. I shall refer to this robust form of stability inherent in nonlinear oscillators as *dynamical stability*, following ABRAHAM (1985) and YATES/KUGLER (1984). How does the mere fact that a process is *nonlinear* give rise to dynamical stability? WEST gives a lucid explanation of the effects of nonlinearity in the following discussion of solitons (coherent vibratory modes or resonances found in certain condensed-matter systems):

"The strength of the nonlinear terms in the field equations is such as to exactly balance the effect of linear dispersion. This delicate balance of linear dispersion driving the component waves apart and nonlinear interactions pulling them together provides for coherent structures such as solitons and generally inhibits the tendency of systems to separate into their incoherent linear components. Thus the physical effects of linear dispersion and nonlinear interactions can reach a dynamic balance to form a soliton … This dynamic balance is one kind of mechanism that is intrinsically nonlinear and could not have been predicted from any extrapolation of the linear world view." (WEST 1985, p80)

In addition to providing a natural way of understanding certain forms of macroscopic coherence, nonlinear dynamics has the added advantage of being tailor-made for modeling emergent phenomena. After all, the essence of emergence lies in the familiar saying that "the whole is greater than the sum of its parts". This is also the essence of the mathematical notion of nonlinearity. A "linear" system is one whose behavior is characterized by a strict proportionality between impressed forces and system responses. That is to say, there is a one-to-one correspondence between causes and effects, or between possible initial states and final states of the system. Such a relationship might be graphed as a straight line (hence the name). A "nonlinear" system, then, is one whose behavior cannot be graphed as a straight line—that is, one in which a many-to-one (or one-to-many) correspondence exists between initial states and final states, or in which there is a disproportionality between causes and effects. As WEST has put it: "Perhaps we can now better appreciate what it is we deny when we contend that this or that phenomenon is dominated by its nonlinear character. We abandon the *proportionate response* of the linear world view. A small change in the input does *not* imply a correspondingly small change in the output; instead a nonlinear system (process) may have a *disproportionate response* in the output, ranging from no effect (if the change is below some threshold value) to an overwhelming instability" (WEST 1990, pp275–276, original emphasis). Similarly, SCOTT rhetorically asks: "What does it mean to say that the dynamics of a system are nonlinear?" and replies: "*In its deepest sense, this is a statement about the nature of causality* … the key feature of nonlinear dynamics is to create something ($E_{1,2}$) that differs from an aggregation of its parts ($E_1 + E_2$)" (SCOTT 1996, p485, emphasis added).

The discipline of nonlinear dynamics, in general, and the concept of dynamical stability, in particular, provide us with powerful tools for understanding the physics out of which the peculiar causal powers of living things emerged. (For a defense of this view of causality—i.e., that the various causal powers in the world derive from the real natures of things and constitute natural kinds—see HARRÉ/MADDEN 1975.) SAVAGEAU has argued that "nonlinear dynamics is central, not peripheral, to understanding integrated molecular networks. It is the nonlinear dynamics that give rise to the most characteristic behaviors associated with living systems … Any attempt to understand integrated behavior in terms of underlying molecular mechanisms must be based on a formalism that includes nonlinear dynamics from the start" (1996, pp121–122). It is not possible here to survey the various dynamical approaches that have appeared in the literature in recent years, whether to metabolism and physiology (HARRISON 1993; MIKULECKY 1995; SAVAGEAU 1996; YATES 1994), morphogenesis (CUMMINGS 1994; HAROLD 1990; NEWMAN 1994; WEBSTER/GOODWIN 1996), evolution (FLYVBJERG et al. 1995; STADLER 1995), or all of the above (BAK 1996; BALL 1999; KAUFFMAN 1993). However, it is important to indicate at least in a qualitative way what is meant by a "dynamical" model of biological functionality.

Organisms consist of congeries of biological functions, each of which may be considered as a nonlinear oscillator functioning as an atomic unit or node within a vast, densely connected network of other oscillators. Each such network may be viewed, in turn, as constituting an individual node within a still larger network, and so on. These networks of networks may be conceived of as extending both vertically across length scales (hierarchical nesting) and horizontally across space at a single scale (heterarchical coupling). The behavior of a given function may usually be represented both as a phase-space trajectory across multiple basins of attraction at one level, and as an attractor in its own right at a higher level. YATES paints a lucid picture of this conception of functional organization in the following passage:

"In any persistent system, whose operations are sustained over periods of time very long compared to the characteristic process and interactional times within it, cyclic energy transformations must be present. Certain processes must occur again and again if the system is to persist. Otherwise we would observe only relaxational trajectories to equilibrium death. Thus, limit cycle-like, nearly periodic, oscillatory behavior is the signature of energy transformations in open, complex, thermodynamic systems obeying both the First and Second Laws (as all real, complex systems must do) … The dynamic structure of a complex system is thus a collection of loosely coupled, limit cycle-like oscillators … In summary, the important characteristic that distinguishes complex atomisms in a field is that their interactions do not rapidly equipartition energy among the accessible translational and internal degrees of freedom. Instead, very significant time delays appear in the distribution of energy among the internal degrees of freedom of the system. A complex field can be thought of as a cooperative, in which the chief processes are fluid-mechanical, gel-like, dissipative—very unlike the more spring-like, conservative interactions of simple, idealized statistical mechanical systems. The net effect of such complexity is to make an account of motion by translational momentum (i.e., by NEWTON's laws of motion) inappropriate. Instead, one must integrate over a time much longer than the relaxation times of translational interactions in order to close the thermodynamic books on energy and entropy changes. This is the process cycle time in which action modes characteristic of the field emerge. Complex systems are thus 'soft' systems without many tight, direct, or 'hard' causalities or couplings." (YATES 1994, pp62–63)

YATES's view of the organism, then, is of a dense network of hierarchically nested and heterarchically loosely coupled, dynamically stable, nonlinear oscillators. The significance of this vision for the philosophical problem of teleology has been pointed out by DELATTRE (1986), who to my knowledge was the first to suggest that we employ the conceptual apparatus of nonlinear dynamics as a framework for understanding goal-directed, functional action (or "finality", as it is referred to in the French literature—see, also, FAVRE et al. 1995). On this view, a given biofunction may be identified with a nonlinear oscillator, a functional action modeled by the associated phase-space trajectory or attractor, and the *success* of the action equated with the *preservation of the dynamical stability of the oscillator*. By adopting the vocabulary of nonlinear dynamics in this way, we free ourselves from enslavement to mechanistic causation defined as a linear, or one-to-one, mapping of initial states (causes) onto final states (effects). In so doing, we also obtain a powerful heuristic for giving a more adequate scientific account of biological functionality.

Needless to say, this way of reading metaphysical lessons from nonlinear dynamics has not won universal acceptance. For example, BRICMONT (1997) claims that dynamicists fail to distinguish between the failure of long-term predictability in nonlinear dynamical systems (due to the property known as "sensitivity to initial conditions") and the determinism of the trajectories themselves implied by the differential equations employed in making the predictions. That is, he accuses dynamicists of confusing epistemology with ontology, and even charges them with "postmodern" irrationalism on that account. But if our warrant for believing in metaphysical determinism in the first place—namely, the apparent universality of NEWTONIAN mechanics—has been overthrown, how can it be irrational to question a belief now shown to be unwarranted in the first place? As ANSCOMBE points out, "[i]t was the impression made on HUME and later philosophers by [NEWTON's] mechanics, that gave them so strong a conviction of the iron necessity with which everything happens" (1993, p95). ANSCOMBE, writing in 1971 long before the recent flurry of interest in "chaos theory", goes on to point out with remarkable prescience that the phenomenon of sensitivity to initial conditions and that of quantum indeterminacy, taken together, show that the very notion of the infinitely precise determination of the position of a real object is incoherent, quite apart from any issues concerning measurement (1993, pp95–96; see, also, KELLERT 1993). This is an *ontological* point. Only fully integrable systems support the inference to metaphysical determinism, but, as CUSHING has recently observed, "such integrable systems turn out to be very special. Our intuition, or general picture of the world, was based on a poor induction from too narrow a range of systems. For nearly 300 years we thought we understood classical mechanics, but we didn't" (1998, p172). On the other hand, the general case—that of nonintegrable, nonlinear, complex systems—does not support metaphysical determinism. As FAVRE and coworkers have put it, "The postulates of LAPLACIAN determinism (if one knew perfectly the positions and velocities of bodies at a given instant…) appear to derive, even in the classic domain, from a philosophy out of touch, except ideally or even ideologically, with the realities of physical systems" (1995, p111, ellipsis in original). Once this fact is admitted, then the notion of multicausality begins to seem decidedly less fantastic.

What, exactly, does this idea amount to? First of all, according to the standard model of cosmic evolution, the four fundamental forces of nature recognized by contemporary physics came into existence piecemeal over time through the process of successive symmetry breakings, as discussed above. As THIRRING observes: "The hierarchy of laws has evolved together with the evolution of the universe. The newly created laws did not exist at the beginning as laws but only as possibilities" (1995, p132). Second, as CAMPBELL (1985) has pointed out, the specific causal powers of things derive from their particular form of material organization. Therefore, I believe it is legitimate to generalize THIRRING's viewpoint beyond the standard model. Since the ways in which matter is organized have changed over the course of cosmic evolution, so too must the causal powers of things have changed over time, or, as THIRRING writes, "[a]s the universe evolved, the circumstances created their own laws" (1995, p135). On this basis, we may distinguish three principal causal modalities corresponding to three "epochs" in the history of the universe. The first modality—which would have been the only form of causation for a short time following the big bang—is the pure spontaneity associated with quantum mechanical systems. The second modality—which may be said to date from the "freezing out" of matter within the primordial fireball—is a kind of quasi-determinism. Processes under the sway of this causal modality appear deterministic over some definite time interval, but begin slowly to deviate from determinism over longer intervals (the length of the characteristic time depending on the type of system). Kenneth DENBIGH paints a picture that is helpful for thinking about the quasi-deterministic modality:

"Consider an analogy. A rigid steel rod of, say, 1 cm diameter is almost unbendable if it is only 10 cm long. But suppose it is 10^n miles long? Very minute impacts would cause it to wave about like a blade of grass. The actual constitution of matter is not such that there is complete rigidity over indefinite lengths; rigidity is an idealization. Similarly, I suggest, the 'determination' of one state of affairs by another is also an idealization and is only to be taken as a good approximation to the extent that, in any actual instance, prediction over finite time intervals is found in fact to be a good approximation." (DENBIGH 1981, p85)

Finally, there is the teleological-functional modality, dating from the time of the origin of life. On this view, the later modalities did not supplant the former ones, but rather supplemented them. Each modality continues to hold sway in its own domain, that is, within its characteristic form of the organization of matter.

In summary, biological functions may be viewed as belonging to a class of systems (nonlinear oscillators) which possess the property of dynamical stability—that is, they are systems whose material organization somehow endows them with the ability to behave in a goal-directed or teleological manner. From this point of view, the selectionist slogan "survival of the fittest" begs the question of teleology not once but twice, for not only fitness but *survival* itself is a teleological notion that implies an active struggle against the second law of thermodynamics.

However, hurricanes and candle flames are nonlinear oscillators and yet they do not engage in any such active struggle against the second law—they are not alive. What is the differentia of life, within the class of dissipative structures? That is, What else, in addition to dynamical stability, is required for life? In short, What is the biological equivalent of lift? This is the question to which we must now turn.

The Problem of Teleology Revisited

According to Ernest NAGEL's (1998) cybernetic interpretation, the two marks of functionality are directive organization and orthogonality, as we saw above. Directive organization, in turn, has the features of persistence and plasticity. There has been an unfortunate tendency on the part of some adherents of the functionalist school to conflate the orthogonality requirement with the plasticity feature of directive organization. This confusion is very likely one source of the pernicious doctrine of the "multiple realizability" of functions in arbitrary material "substrates" (e.g., ENÇ/ADAMS 1998, p389). If there is one place more than any other where our thinking about functions has gone off the track, it is here. As NAGEL himself pointed out, even inanimate systems can have "plastic" behavior in the sense that they may tend towards a given final state by different routes. In short, there is more to orthogonality than plasticity, and neither one necessarily implies multiple realizability.

What the orthogonality requirement does imply is that a truly functional or goal-directed system must be loosely coupled to local external thermodynamic flows; otherwise, it could not help being driven to the local energy minimum. So functionality necessarily implies a degree of thermodynamic decoupling. We also observe that in order for a living system to preserve itself in existence for an extended period of time, it must be capable of acting against local energy gradients on occasion, if these tend to undermine its dynamical stability. Since all action requires expenditure of energy, this implies that a functional system must have internal or "on-board" energy stores which it can draw upon in order to act against local gradients (SWENSON 1998). Therefore, functionality also implies energy autonomy.

But notice that there is a problem here. The problem may be expressed by the following question: What good is the possibility of action against local energy gradients, which partial thermodynamic decoupling and energy autonomy make possible, without some means of coordinating such action with the appropriate external conditions? That is, independence of action will not help a functional system to preserve its dynamical stability *unless the system can somehow distinguish those external conditions which will support that stability from those which will not*. Put another way, the problem of coordination is essentially one of timing: of selecting the right moment at which to act. YATES has emphasized the importance of this fundamental point as follows: "…in the case of a crystal lattice the spatial order is best expressed by the presence of correlations among the positions of equal atoms … In functional order the correlations must be formed among the *times* at which different events occur" (YATES 1993, pp190–191, original emphasis). So, it seems that in order for genuinely functional or normative action to be possible, there must be a means of coordinating actions with external conditions. In short, *the possibility of normativity implies the recognition of its own conditions of satisfaction*.

According to this analysis, then, functional or teleological action consists of two closely related but distinguishable features. First, there must be a definable goal state which is achieved through the coordination of efficient causes through some means other than energy minimization. All nonergodic behavior is "special" in a statistical-mechanical sense, but true functional behavior transcends this "specialness" because its coherence cannot be accounted for by the (known) laws of physics. This "special specialness", so to speak, is really nothing else than *normativity*, or so I wish to claim. However, I would like to introduce a technical term here for the sake of clarity. Let us call this power of striving towards a goal state against local energy gradients—a power inherent in all living things—the *conative* aspect of functionality, or simply *conation*. Then, as we have seen, in order for individual efficient causes to collectively bring about a goal state, each one must play its part in just the right way at just the right time. That is, each cause must act in a way that is appropriate to the overall task. This ability of living things to coordinate their actions with external circumstances in order to achieve their ends is one of the most mysterious aspects of functionality. Since oscillation, or cyclicity, is an essential property of biological functionality (GOLDBETER 1996), there is no need to invoke "backwards causation". However, due to the loose thermodynamic coupling between system and surround, there does appear to be a problem in understanding how the correct adjustments occur at the right time—

namely, the problem of accounting for the appearance of a *preestablished harmony*, so to speak. For present purposes, let us call this ability of living things to respond appropriately to external circumstances the *cognitive* aspect of functionality, or simply *cognition*.

Now, we saw above that it makes no sense to speak of information except in reference to the goals or interests of an agent, or, as we might say in the form of a maxim: *no cognition without conation*. Furthermore, we have just seen that the idea of normativity implies the capacity to recognize its own conditions of satisfaction. Therefore, we can equally well say: *no conation without cognition*. In this way, conation and cognition appear to be the jointly necessary and sufficient conditions for biological functionality—that is: *teleology = conation + cognition*. Moreover, we now have the answer to the question we posed above about the additional feature (besides nonergodicity) that is required for functional action: *Cognition is the biological equivalent of lift.*[13]

Recalling DELATTRE's idea that the phenomenon of teleology might be usefully modeled by means of the notion of the dynamical stability of a nonlinear oscillator, it is now possible to give a more specific and detailed account of the dual conative-*cum*-cognitive nature of teleology within the framework of nonlinear dynamics. This is what I shall attempt to do in the next section (see, also, BARHAM 1990, 1996).

A dynamical model of biological functionality

Let us suppose that every biofunction contains within it a subsystem such that (1) the subsystem is capable of undergoing a physical interaction with certain low-energy exogenous constraints which are highly correlated with that part of the external surround which supports the high-energy functional action of the system as a whole; and (2) the interaction of the subsystem with the low-energy constraints causes a state transition which acts as a trigger for the functional action. In this way, the timing of the action of a biofunction (considered as a nonlinear oscillator) may come to be correlated with the presence of just those external conditions which permit the success of the action (where "success" is understood as the oscillator's dynamical stability's being preserved). I call the postulated subsystem the *epistemon* and the interaction it undergoes with the low-energy constraints the *epistemic interaction*. In this scenario, the low-energy constraints constitute *information* in the autotelic sense explained above.

This postulated fundamental interaction between system and surround, mediated by the epistemon, may be further analyzed into two sets of four basic elements. First, we have a set of four basic *entities*: (1) a function *F* (conceived of as a nonlinear oscillator); (2) an epistemon *E* (a component or subsystem of the function); (3) information *I* (a set of low-energy exogenous constraints); and (4) a surround *S* (a set of high-energy exogenous constraints). Second, we have a set of processes or *interactions* among these four entities: (i) some physical process correlating the surround with some information (*S-I*); (ii) the epistemic interaction between the information and the epistemon (*I-E*), resulting in a state transition in the latter; (iii) the triggering process such that the state transition in the epistemon initiates the oscillation of the function (*E-F*); and (iv) the functional action itself, consisting of a high-energy interaction between the overall function and the surround (*F-S*), thus closing the cycle. In this way, the preestablished harmony problem may be understood as the mirroring of the surround-information correlation (*S-I*) in the world by the epistemon-function (*E-F*) correlation in the organism.

This model gives us for the first time a basis for understanding how information can be at the same time physical and meaningful. Information is physical insofar as it is a low-energy[14] constraint on a biofunction capable of interacting with the function's epistemon, i.e., capable of taking part in an epistemic interaction. Information is meaningful insofar as the epistemic interaction correlates functional action with those environmental conditions that will support it. Thus, on this view, we may say that *the meaning of information is the prediction of successful functional action*. For this reason, I have dubbed the fundamental process of cognition embodied in the tetradic model outlined above *prorrhesis* ("predicting"), as opposed to the more familiar triadic model of *semiosis* ("signalling"). Semiosis begs the question of meaning by positing the "interpretant" as an unanalyzed primitive concept. To be sure, prorrhesis also presupposes teleology, both in the notion of the dynamical stability of a nonlinear oscillator and in the preestablished harmony between function and surround. But, unlike semiotics, "prorrhetics" is anchored in a specific body of scientific theory—namely, nonlinear dynamics. If the dynamical model of biological functionality has not completely exorcised the homunculus, it has at least confined him inside a theoretical structure where he may be subjected to further scientific analysis.

I believe that the prorrhesis model has a significant bearing on a whole range of philosophical problems. Indeed, this is what one would expect of any general theory of functionality, for, as BENNETT has noted, "[t]here is a line of conceptual dependence from 'language' to 'meaning', to 'intention' and 'belief', to 'organism', to 'physical object', and philosophical problems abound all the way down" (1990, p24). Here I can do no more than briefly sketch a few of the most important of these wider problems (for further discussion, see BARHAM 1990, 1992, 1995).

Evolution

Biofunctional realism throws light on a number of problems in the philosophy of biology. First of all, it provides a new framework for thinking about adaptation. Instead of fitness, we have function, grounded in the concept of dynamical stability. In place of the selection for/selection of distinction, we might postulate something like a central/peripheral distinction to mark the fact that some biofunctions are more important to an organism than others. Since we do not attempt to define "proper functions" in terms of evolutionary history (MILLIKAN 1984, 1999), we may make free use of ordinary engineering criteria instead in order to evaluate the relative importance (degree of centrality) of a given function. In this way, we may dispense with Just-So Stories and the whole scholastic vocabulary of "preadaptations", "exaptations", "aptations", etc. According to the dynamical model, the only difference between an adaptation, on the one hand, and a spandrel (a structurally determined, nonselected trait) or a vestige (a formerly selected trait no longer actively selected for), on the other, is the degree of centrality of the function in question for the dynamical stability of one or more higher-order oscillators of which it is a part (up to the organism as a whole). On this view, every organic "trait" is functional by definition, but functionality comes in different "strengths", so to speak. One structure (e.g., the appendix) may be quite peripheral (with reference, say, to the whole organism), while another (the heart) may be absolutely central. In this way, functionality ceases to be a metaphysically mysterious property deriving from the history of an organism instead of its present causal powers (MILLIKAN 1996), and becomes once again a real physiological property amenable to experimental investigation. Simply remove the appendix from one animal, the heart from another, and compare the results. Even the hoary question of whether the function of the heart is to pump blood or to make beating sounds is decidable according to this model. Open the chest cavities of two similar experimental animals. Wrap one heart in acoustic insulation; sever the other one's connections to its circulatory system. Then, observe the effect of each alteration on the dynamical stability of oscillators at various hierarchical levels (e.g., the heart itself, the circulatory system, and the whole organism).

Biofunctional realism also provides a better framework for understanding the relationship between evolution and development. We begin by distinguishing between two very different properties of nonlinear dynamical systems: *equifinality* and *metastability*. The ontogenetic development of multicellular organisms is clearly a goal-directed, i.e., a many-to-one (equifinal) process. This means, as we have seen, that the system strives to remain within its basin of attraction even after perturbation (within limits). As is well known, developmental processes often compensate for even quite large insults and may arrive at the pre-determined form via alternative routes. Contrary to a well-known view (SALTHE 1993) that is in danger of becoming confounded in some quarters with the notion of selforganization itself, evolution does *not* exemplify a developmental process of this sort. On the contrary, phylogenesis can be traced to an entirely different property of nonlinear dynamical systems: namely, metastability. This term refers to the propensity of a nonlinear oscillator to find a new dynamically stable regime when forced out of its original basin of attraction by a too-great perturbation (an event known as a "bifurcation"). On this view, then, the evolutionary process may be modeled by a sequence of bifurcations leading from one attractor to another. Each individual bifurcation event may be represented as a one-to-many mapping of the single actual initial state onto a multitude of virtual final states, while the whole sequence of bifurcations may be viewed as a trajectory through a higher-order phase space. According to this way of looking at things, then, an instance of phylogenetic adaptation may be viewed as a bifurcation event through which a population (i.e., an organism-type) finds a new dynamically stable attractor, while an instance of ontogenetic learning represents a similar bifurcation event in an organism-token. That is to say, ontogenetic development and functional action are both special cases of the fundamental dynamical property of equifinality, whereas phylogenetic adaptation and learning are examples of the closely related but conceptually distinct property of metastability.[15]

Multiple realizability vs. dynamical stability

Goal-directed systems are by definition plastic, meaning that they will strive to reach their goal states by alternative routes if necessary. Plasticity may be conceptualized as the many-to-one mapping of possible initial states of a system onto an actual final state (the mathematical property of equifinality). Functionalism equates plasticity with the doctrine of multiple realizability (SOBER 1999, p545). To say that a function is "multiply realizable" is to say that there is no intrinsic connection between its material substance and its behavior. The functions of *machines*, clearly, are multiply realizable in this sense. If organisms are machines, then their functions must be multiply realizable, too. But what if organisms are not machines? In that case, how do we preserve the grain of truth in functionalism (the recognition of the plasticity of functions), while rejecting the universal mechanistic causality it presupposes? The answer lies in the concept of dynamical stability. The dynamical stability of a system is a robust, global property that is insensitive to the tokens "instantiating" the system (i.e., the details of the system's microstructure), but is nevertheless causally dependent on a class or type of physical structure constituting a natural kind. Since robust plastic behavior clearly exists, and since it is inherently irreducible to brittle linear causality, something like the notion of dynamical stability is logically required to save the phenomena (as has become increasingly apparent to a few philosophers in recent years—see, e.g., BATTERMAN 2000, BECHTEL/MUNDALE 1999; M. WILSON 1993). In addition, this way of looking at biofunctions is also increasingly supported by empirical evidence. For example, a number of recent studies have demonstrated the robustness of functional properties at the cellular level in the face of alterations of various molecular parameters (ALON et al. 1999; BARKAI/LEIBLER 1997; BHALLA/IYENGAR 1999). One remarkable study (HEALD et al. 1996) has even revealed the plasticity of the selforganization of mitotic spindles in the eggs of the tree frog *Xenopus* in the absence of many of the usual components believed mechanistically necessary (e.g., centrosomes, kinetochores). Just as dissociated myocardial cells spontaneously seek to form a functioning heart (see Note 6), so too do microtubules dissociated from their normal molecular matrix still strive to form a functioning mitotic spindle. These experiments show that it is highly misleading to think of biological functions as "finely-tuned machines" (or, indeed, as machines of any sort). At the same time, though, the causal powers of biofunctions must derive somehow from the properties of the special kind of matter of which they are made. The notion of dynamical stability can help us to bridge this conceptual gap by showing how functional behavior might simultaneously be associated in lawlike fashion with a class of physical structures and yet be relatively insensitive to microstructural details (cf. the related notion of "ordered heterogeneity" in ELSASSER 1998). In this way, it can serve also as the foundation for a new way of conceiving the type-type identity relation between "mental" and "physical" states (or, better, events).[16] On this view, mental events are identical to physical events in the brain, but the physical events are themselves irreducible to mechanistic causation.

Malfunction and error

An important test of any theory aiming to naturalize teleology is accounting for malfunction and error. The present model accomplishes this by grounding function in nonlinear dynamics, as opposed to mechanistic causation. Once the rigid link between cause and effect is broken, the difficulty in seeing how actions elicited by environmental triggers can ever be *mistaken* vanishes. For example, take the fundamental unit of biological function, the enzyme (DRESSLER/POTTER 1991; STRYER 1989; see, also, FRUTON 1999). The function of an enzyme is to make and break high-energy (covalent) bonds between its ligands (substrates). It does this by means of a prior low-energy (noncovalent) interaction between its own active site and a ligand. This prior low-energy interaction at the active site induces a conformational change in the enzyme as a whole, which in turn produces the high-energy functional action. If this whole process occurs as it is supposed to, the ligands will detach from the enzyme, and the latter will return to its initial state, ready for another cycle. Now, according to the prorrhesis model, we may view the active site as an *epistemon*. Noncovalent bonding with a particular ligand, then, is the *epistemic interaction*, and is tantamount to the prediction that a successful covalent bonding will follow. The way many drugs work is by imitating an enzyme's substrate so as to cause this prediction to fail. For example, the cell wall in many bacteria consists of a macromolecule called *peptidoglycan* (STRYER 1989, pp195–197). This molecule is constructed, in part, by the enzyme *transpeptidase*. Transpeptidase's job is to join its normal substrate (D-alanine) with glycine residues in the growing strand of a peptidoglycan molecule. Penicillin works by imitating D-alanine insofar as the preliminary noncovalent bonding at transpeptidase's active site is concerned. In this way, transpeptidase is fooled into undergoing its conformational change in the wrong context. That this is indeed a *mistake* can be seen by its consequences: When transpeptidase undergoes its conformational change, penicillin bonds covalently with transpeptidase itself, instead of linking up with glycine and releasing the enzyme, as D-alanine would do. Thus, the functional action *fails* in the sense that transpeptidase loses its dynamical stability and ceases to function as an oscillator (the prediction implied by the initial noncovalent bonding is proven wrong). Notice that this way of looking at things avoids the "disjunction problem" that has plagued causal theories of intentionality (see FODOR 1990; Chap. 3, pp51–87). There is no need to worry that transpeptidase's substrate might have to be interpreted as "D-alanine or penicillin" (or a frog's food as "flies or BBs"), since we are able to distinguish between the act of perception (the epistemic interaction) and the criterion according to which the act is evaluated (whether or not it predicts successful functional action). According to the prorrhesis model, the correct functioning of the system as a whole is the criterion for assessing the correctness of the information delivered to it via the epistemon. Disjunction is only a problem so long as we lack a normative criterion that is intrinsic to the function itself. As soon as we regard biofunctions realistically as possessing intrinsic normativity, then we are able to distinguish informational inputs according to their *meaning* (intension or aspect) for the system, and the disjunction problem disappears.

Nonrepresentationalism and proto-intentionality

Biofunctional realism has a number of other implications for the philosophy of mind. The most important of these may well be to help us tease apart which causal powers the human mind enjoys merely by virtue of being biological, and which it owes to its species-specific properties (primarily language). Biofunctional realism can aid in this effort by contributing to the research program already well under way which is endeavoring to understand brains, not as computational devices employing symbolic representation, but rather as dynamical systems coupled to their environments (e.g., BEER 1995; FREEMAN 1995A; JUARRERO 1999; KELSO 1995; PORT/VAN GELDER 1995; SHAW/TURVEY 1999; TURVEY/SHAW 1999; VAN GELDER 1998). For instance, it has been shown (FREEMAN 1992; SKARDA/FREEMAN 1987—see, also, NICOLIS/TSUDA 1999) that particular chaotic attractors in the patterns of electrical activity of nerve cell assemblies in the olfactory cortex of rabbits can be reliably correlated with particular categories of odorants the rabbits have learned to recognize. Such attractors may be deemed "concepts". However, concepts so understood do not "represent" the world; rather they constitute biological functions that *interact* with the world (BEER 1997; R. A. BROOKS 1999; FREEMAN 1995B; KEIJZER 1998; PFEIFER/SCHEIER 1999; VAN GELDER 1995).[17]

By assimilating this body of work to the biofunctional–realist perspective, we can begin to see how the full-fledged intentionality of the human mind is grounded in the proto-intentionality of biological functionality generally. In this way, we also acquire the basis for a general theory of *voluntary action* (FREEMAN 1999; HO 1996). Such a theory would view the desire/belief distinction at the level of the mind as a reflection of the more fundamental conation/cognition distinction at the level of the biofunction. That is, conation may be viewed as proto-desire, or world-to-agent fit, while cognition constitutes proto-belief, or agent-to-world fit. This general approach can also help us see how both sides in the internalism/externalism debate have hold of part of the truth. Meanings are partly "in the head" (that is, mental states have "narrow content") since cognition is in part intensional or aspectual (relative to categories conceived of as nonlinear attractors internal to the system). However, meanings are also partly "in the world" (mental states have "broad content") since cognition is in part extensional (functional action being triggered by information conceived of as low-energy constraints external to the system). On this view, then, all biological functions are constituted by both intrinsic *and* extrinsic properties, all possess both essential natures *and* context-sensitivity; so far as brains are concerned, the seeming paradoxes of the internalism/externalism debate disappear as soon as one recognizes that the "here-and-now" causality within the brain is in fact a dynamical system coupled to an external informational field (SWENSON/TURVEY 1991).

Where human mentation differs profoundly from other biofunctions is in the additional looseness of fit (nonlinearity) we enjoy between our internal functional categories, on the one hand, and external information, on the other. This additional flexibility is likely almost wholly attributable to language, which is itself a form of social rule-following. That is to say, rational thought is firmly grounded in pan-functional proto-intentionality which encompasses both aboutness (teleology/conation) and intensionality (meaning/cognition); however, human reason transcends proto-intentionality through the power of symbolic representation. In true (i.e., language-mediated) representation, the relationship between signs and meanings really is arbitrary, since it is determined by socially imposed rules (SEARLE 1995) and not the causal powers of living matter.[18] It is convention-based representation which makes possible full-fledged intentionality, with the possibility of nonexistent reference, in addition to aboutness and intensionality. This is due to the fact that a linguistic symbol (e.g., "unicorn") is capable of creating concepts (i.e., eliciting neural attractors) in individual brains, even if there is no referent for the symbol outside of the language itself. On this view, cognition in the biofunctional sense arose through an initial partial decoupling between system and surround at the time of the origin of life, whereas rational thought came about through a second decoupling between brains and their environment at the time of the origin of language.

This view of the relationship between biological function, in general, and human reason, in particular, has a number of further philosophical implications. For one thing, the conception of teleology as conation-*cum*-cognition provides a much firmer foundation for evolutionary epistemology. By decoupling functional success from whole-organism survival, we preserve what is true in pragmatism (cognition is essentially connected to functional action), while avoiding DARWINIAN irrationalism (truth is whatever works) (for details, see BARHAM 1990, 1992). For another, this perspective has important implications for the contemporary debate on the objectivity of knowledge. Although human linguistic meanings are multiply realizable (instantiated in different grammars and phonologies) due to their grounding in social conventions, they are nevertheless universal, in the sense of being potentially learnable and mutually intertranslatable, due to the grounding of social praxis itself in universal biological functionality (see BARHAM 1995). (In the future, I hope to show that biofunctional realism can provide the basis for achieving a scientific understanding of man's place in nature that is far more adequate to the full range of human experience than that offered by the currently fashionable DARWINIAN "evolutionary psychology"—see BARHAM, in preparation.)

Conclusion

As MILLIKAN has rightly observed, "Meaning and truth cannot be naturalized without a theory that naturalizes norms generally" (1991, p151). The chief aim of this paper has been to argue that biofunctional realism shows promise of finally succeeding in this endeavor where the DARWINIAN approach MILLIKAN herself favors has failed. For too long, naturalistic philosophers have distorted the problem of teleology to fit the PROCRUSTEAN bed of an outdated cybernetic-selectionist worldview. The new dynamical-emergentist worldview sanctioned by nonlinear dynamics, nonequilibrium thermodynamics, and condensed-matter physics provides a much better fit—a far more natural way of conceiving of the phenomenon of teleology—and for that reason alone deserves the serious attention of philosophers.

Many will no doubt dismiss biofunctional realism as a return to "vitalism". I believe that this is mainly a matter of semantics. Historically, vitalism has come in a number of "strengths", the weakest of which is "the view that there is a distinctive organization among living things" (BECHTEL/RICHARDSON 1998, p639). By this definition, biofunctional realism does indeed qualify as a form of vitalism. So be it. Life *is* different. That is an undeniable fact. What, then, could be more natural than to conclude that this difference is the result of a "distinctive organization" of matter itself? The important thing is to ask whether biofunctional realism is compatible with *physicalism*. Here, it is crucial to distinguish "weak" and "strong" forms of physicalism, as well, for which purpose we may draw on the well-known distinction introduced by MEEHL/SELLARS (1956). According to their scheme, events or entities are physical in the strong sense ("physical_2") if they are "definable in terms of theoretical primitives adequate to describe completely the actual states though not necessarily the potentialities of the universe before the appearance of life" (MEEHL/SELLARS 1956, p252), and physical in the weak sense ("physical_1") if they "[belong] in the space-time network" (*idem*). By these definitions, there is no doubt that biofunctional realism is a form of physicalism in the weak sense, though not in the strong sense (for further discussion of the meaning of physicalism, see POLAND 1994).

The question is not whether living matter is different from nonliving matter. There can be no doubt that it is very different indeed. The question is, rather, in what the difference consists, and whether our current concepts can adequately account for it.

As KAUFFMAN has observed, "Living systems unequivocally demonstrate that certain forms of matter and energy, displaced from thermodynamic equilibrium, emerged spontaneously and coevolved to form the biosphere. *The emergence and evolution of life must be a natural expression of a properly defined class of matter and energy*" (1994, p86, emphasis added). Getting clear about the phenomenon of emergence is key here. As EMMECHE and coworkers have observed, "the concept of emergence is exactly that reasonable aspect of vitalism which is worth [maintaining]" (EMMECHE et al. 1997, p86). Emergentism, on this view, is a position halfway between reductionist monism and "egalitarian pluralism" (CRANE/MELLOR 1995, p88). The postmodern pluralists (e.g., DUPRÉ 1993; GALISON/STUMP 1996; ROSENBERG 1994) deny the metaphysical unity of the world in order to avoid "privileging" any one level or science over any other. And yet, if one is a realist (that is, if one believes that the cosmos has given rise to human beings, and not the other way around), and if one is a temporalist (if one believes that time and change are real, and not figments of the human mind), then it is impossible to avoid the conclusion that the very process which has brought forth the present highly differentiated cosmos out of an original featureless chaos itself constitutes a kind of unifying metaphysical principle. That is, the DE cosmology offered here is a *dynamical monism*. If traditional monism sought to impose unity on the world by reducing the many to the one (*e pluribus unum*, so to say), dynamical monism seeks to explain how in the fullness of time the one has given rise to the many (*ex uno plura*).

But that still leaves the question, How should the emergent level constituting life be defined? Admittedly, the dynamical model of biological functionality presented here is phenomenological in character; one would clearly like to have a deeper understanding of the physical processes involved. It is, of course, impossible to say at present exactly what form such a theory might take; however, it is easy to say where we should look for it. Above all, what is wanted is a theory of *proteins*,[19] which are the primary locus of biological agency and which constitute a separate and distinct state of matter (FRAUENFELDER et al. 1991a). Such a theory might conceivably take a number of forms, and several proposals have in fact been placed on the table, both classical (HARRISON 1993; MIKULECKY 1995; SAVAGEAU 1996; YATES 1994) and quantum (DEL GIUDICE et al. 1988; FRÖHLICH 1988; HO 1993; WELCH 1992) in character. (See, also, the more abstract, but highly suggestive, topological approaches of HYDE et al. 1997, and THOM 1990b.) Ideally, what is ultimately required is a theory of the living state with the scope and rigor, if not the precise mathematical form, of quantum field theory. While a lively debate on the best way to proceed towards such a theory is under way (see HAMEROFF/SCOTT 1998; VITIELLO 1998), we are still very far from a consensus.[20] So, for now, something along the lines of the qualitative approach outlined here remains indispensable. The essential point is to recognize that "the complex systems that constitute living organisms obey general laws at the macroscopic level. This way of thinking is still alien to many biologists, but not to physicists" (RIPOLL et al. 1998, p20). In order to understand the mind's place in nature, biologists and philosophers alike are going to have to learn to think more like physicists.

Functionalist cognitive science teaches that there is no essential connection between biofunctions and the matter of which they are composed. A certain biofunction has the causal powers that it does by virtue of falling under a particular biological category. Thus, the capacity to *pump blood* does not proceed from the powers inherent in myocardial tissue; rather, myocardial tissue has the capacity to pump blood because it is a *heart*. In other words, the properties of the token derive from those of the type, and not the other way around. A more profoundly *anti-naturalistic* way of thinking is hard to imagine. Functionalism is anti-naturalistic above all because it is anthropocentric. It refuses to distinguish human artifacts, which are multiply realizable (because it is we who determine what counts as a goal state for them), from biofunctions, which are not. By conflating these two essentially different kinds of teleological phenomena, we end up viewing them both as equally dependent upon human intentionality, and hence ultimately as subjective fictions. But demanding a single theory for all "functions", whether natural or artificial, is senseless. To expect a single theory of "hearts" to explain both living organs and mechanical pumps makes as much sense as expecting a single theory of "gold" to explain both mineral ores and wedding bands. As LAMBERT has observed (1995, p256): "Function as we know it today represents a concept gone wrong. But more than that, it is a particular manifestation of a broad trend in a biology gone wrong, one in which explanation resides, not in the logic of underlying processes, but in historically imagined events which lead to an 'adapted' world … In contrast, we need to develop an alternative biology in which structures are again active and one in which their generative processes represent the driving force of nature". In short, in order to understand teleology as an objective natural phenomenon, there is no alternative to some form of biofunctional realism.

Konrad LORENZ once remarked that the theory of natural selection could be said to naturalize KANT's synthetic *a priori*—what is *a priori* for the individual is *a posteriori* for the species (LORENZ 1962). Certainly, DARWINISM has traditionally regarded KANT's anthropomorphic view of teleology as most congenial (DENNETT 1987). Functionalist biology and philosophy of mind remain mired in KANTIAN subjectivism—and ultimately, despite all protestations to the contrary, CARTESIAN dualism (M. WHEELER 1997), and thus incoherence—precisely because they invariably presuppose a "principle of an original *organization*, a principle that is inscrutable to us" (KANT 1987, p311, original emphasis). Of course, KANT was quite right to declare a *mechanistic* science of biology self-contradictory (1987, pp282–283). But that only shows that the possibility of a return to pre-KANTIAN realism and objectivity in science requires the overthrow of the CS worldview. There is every reason to believe that there may be a "NEWTON of a blade of grass" some day, but DARWIN was not that person. We cannot know in advance what a general theory of biological functionality will look like, but one thing we may be sure of is that it will transcend the CS worldview. If and when such a theory is developed, whether in accordance with the DE worldview sketched in this paper or some other principles currently unimaginable, then we will have succeeded at last in naturalizing, not so much KANT's synthetic *a priori*, as LEIBNIZ's preestablished harmony. This may sound like a grandiose ambition, but it is a necessary one, for we will never understand the mind's place in nature until we have first solved the problem of teleology.

Acknowledgments

I would like to thank staffs of the Ephrata Public Library, the Lancaster Area Library Interlibrary Loan Department, and the Follett Bookstore at Franklin & Marshall College for their aid in obtaining research materials for this project. In addition, I wish to thank the following individuals for their invaluable assistance: Jesse ELLIS, Sharon GORDON, Jim GROFF, Vince HAAG, Chris PAUL, and Connie WEBSTER. I am also grateful to Mark BICKHARD, Simon CONWAY MORRIS, William DEMBSKI, Elias KHALIL, Kalevi KULL, Paul LEVINSON, Pedro MARIJUÁN, Koichiro MATSUNO, Alexander MEBANE, Howard PATTEE, Glenn ROSS, Rod SWENSON, Roger D. K. THOMAS, Roger K. R. THOMPSON, and F. Eugene YATES, who have all given to me generously of their time and their expertise. None of them, of course, should be held responsible for any of the ideas presented here (indeed, most of them will dissent vigorously!). It is a pleasure, as well, to be able to thank my friends Stephen J. ANDERSON, Mark McCULLEY, and Aron ZYSOW for their support, both moral and material, over the years. Finally, I wish to express my deepest sense of gratitude to Ellen F. HALL.

I would like to dedicate this work to the memory of Donald T. CAMPBELL and Edward S. REED, who both disagreed profoundly with the ideas contained herein, but did everything in their power to help them see the light of day, anyway.

Author's address

James Barham, 192 Gregg Circle, Ephrata, PA 17522, USA
Email: jamesbarham@supernet.com

Notes

1 There are two closely connected issues here. The first is whether value is essentially connected with minds (BEDAU 1990; BENNETT 1990). The second is whether, even it is not, it makes sense to attribute value to non-mental biological functions. Most philosophers are now willing to speak *as if* biological functions possessed intrinsic value, since most agree that it is incoherent to do otherwise. That is, in order for a biological "trait" to count as a function, there must be something that it is *supposed* to do. This means that the concept of function contains as an essential element the notion of a criterion or standard, in relation to which either *success* or *failure* (*malfunction*, *error*) is possible. But that is not to say that there is agreement about whether this appearance of value is subjective or objective, or, if it is objective, how it is to be understood in relation to the rest of nature. Far from it. It is not possible to provide here a detailed exegesis of the spectrum of opinions on this question, but for a representative sample of the literature, see BEDAU (1992), BENNETT (1990), DAVIES (2000), DRETSKE (1988), KITCHER (1998), MILLIKAN (1984), PRICE (1995), WACHBROIT (1994), WOODFIELD (1976), and WOUTERS (1995). (For an introductory overview of this literature, see BULLER 1999; for a comprehensive review, see NISSEN 1997.) Many of these authors signal the as-if normative character of biological functions by referring to "proper functions" or "teleofunctions". However, since one of the main theses of this paper is that the normativity of biological functions is entirely objective, and the idea of a nonnormative (mechanistic) biofunction incoherent, I shall speak here of "functions" *tout court*.

2 In this paper, the term "teleology" will always be used in the "internal" or "immanent" sense, and not the "external" or "transcendent" sense (LENNOX 1992). That is to say, natural or "original" teleology is always conceived here as arising spontaneously from within a particular kind of concrete material system, never as imposed arbitrarily on heterogeneous systems from without.

3 Supervenience, of course, is the covariance, or token-identity, relation between hierarchical levels, most often invoked in connection with mental states and their corresponding neural "substrates". As KIM has pointed out, "mind-body supervenience itself is not an *explanatory theory* ... it is only a 'phenomenological' relation about patterns of property covariance, patterns that possibly are manifestations of some deeper dependence relationships. If this is right, mind-body supervenience *states* the mind-body problem—it is not a solution to it" (1998, p14, original emphasis). This is as it must be if multiple realizability is true, since the latter doctrine is just the denial of a systematic, or type-identity, causal relation between levels. KIM himself (if I read him right) draws a reductionist/fictionalist lesson from the vacuousness of supervenience; in my view, his arguments are better viewed as a *reductio* of the doctrine of multiple realizability. For further discussion (and a penetrating critique) of the explanatory value of the notion of supervenience, see HORGAN (1993).

4 For the relevant calculations, see YOCKEY (1992, pp246–257); see, also, ELSASSER's notion of "immense" numbers (1998, p50). For an insightful discussion of the interpretation of such calculations, see FRY (1995). While the same statistical-mechanical considerations would be applicable to the question of the role of chance in the origin of any novel biological structure, the problem is frequently discussed in relation to the origin of life. This is the truly vital piece of the puzzle that is still missing, without which our picture of life and mind is bound to remain radically incomplete. As WÄCHTERSHÄUSER has noted, "the life sciences are the arena where mechanistic explanation and teleological understanding come into close encounter. Nowhere is this encounter in sharper focus than in the problem of the origin of life" (1997, p483). (The best point of entry into the vast and contentious field of origin of life research is LAHAV 1999; for the cosmic, geological, and chemical contexts in which life has evolved, see WILLIAMS/FRAÚSTO DA SILVA 1996—a fascinating text that brings L. J. HENDERSON's notion of the "fitness of the environment" up to date.) There are really two problems here: (1) getting clear about what we are trying to explain; and (2) explaining it. In regard to (1), the point of view of the present paper is consistent with the idea that domain formation (LUISI 1993; NORRIS/RAINE 1998) and a minimal self-sustaining metabolism (DYSON 1999; MOROWITZ 1999; SEGRÉ/LANCET 1999) must have preceded gene-mediated replication. There are two powerful reasons for preferring the "cells first" point of view to the more fashionable "RNA world" scenario: a logical one and a statistical-mechanical one. The logical reason is that without phase separation (i.e., a proto-cell) and a dynamically stable energetics capable of maintaining it away from equilibrium, there is no living entity and hence nothing to replicate. The statistical-mechanical reason is that lipids and proteins are simpler than nucleic acids, and we must assume that less thermodynamically improbable structures preceded more improbable ones. The present viewpoint is also opposed to the "strong artificial life" program that views life as essentially an abstract relation rather than a physical process (BODEN 1999). (Note that strong A-Life is just the doctrine of functionalism taken to its logical conclusion. For an attempt to reorient A-Life thinking along more physically realistic—i.e., non-functionalist—lines, see M. WHEELER 1997.) In regard to (2)—the *explanation* of the origin of proto-cells—I believe that YOCKEY is correct when he opines that that is far beyond our current capability, and that there is little to be gained from persisting with our present lines of inquiry (1992, p289). As VITIELLO has observed, biology today is like chemistry before the creation of quantum mechanics—in an essentially pretheoretical state (1998, p194). If that is so—if some fundamentally new ideas are required in order to understand how life is possible—then we must be prepared to fundamentally rethink the physics of the living state. For various speculations along these lines, see CONRAD (1997 1998), DEL GIUDICE et al. (1988), FRÖHLICH (1988), HO (1993), KAUFFMAN (1993, 1996), MATSUNO (1989), and WELCH (1992).

5 See LEVINTHAL (1991) and SARKAR (1996) on the ways in which the concept of information has been used and misused in molecular biology. I agree with SARKAR that "there is no clear technical notion of 'information' in molecular biology. It is little more than a metaphor that masquerades as a theoretical concept" (1996, p187). However, that is not to say that information can simply be dispensed with—a glance at any textbook reveals a plethora of such closely related terms as "recognition", "regulation", "signals", "receptors", "second messengers", etc. The reason for the ubiquity of such seemingly anthropomorphic concepts, of course, is their usefulness; without such an informal teleological vocabulary it would be impossible to make sense of molecular biology. SARKAR himself recognizes this when he remarks that "[p]erhaps lurking behind this usefulness there is some insight to be grabbed, which the conventional information-based account of molecular biology has grasped even if, so far, very shakily" (*ibid.*, p219). It is interesting that Sarkar calls for a new dynamical approach towards molecular biology that will go beyond conventional information theory and cybernetics (*idem*), while despairing of the possibility of a "general account of information" (*ibid.*, p218). In this paper, I shall argue that an adequate theory of biological functionality and a theory of semantic information must go hand in hand (see, also, BARHAM 1996). Some readers may object at this point that we already possess such a theory—namely, "biosemiotics". It is true that by emphasizing questions relating to the *meaning* of signs, or information, biosemiotics appears to get much closer to the heart of the matter than conventional cybernetics. The problem is that writers working in this tradition for the most part take the normativity inherent in sign use, or "semiosis", for granted. If they deal with the question of the natural ground of normativity at all, they inevitably refer the problem back to some combination of cybernetics and selection theory (see, e.g., FALK 1995; HOFFMEYER 1996). For this reason, I do not view biosemiotics as a genuine advance beyond cybernetics.

6 It is well to remember what kind of "pump" the heart really is. First of all, it is self-regulating in the sense that it will adjust to the body's metabolic needs very well even with all external nervous connections severed (BERNE/LEVY 1997, p95). Second, it not only goes on pumping if removed from the body, it even continues to self-regulate if attached to an artificial circulatory system (*ibid.*, pp96–97). To be sure, artificial hearts might be given similar capabilities; more difficult to imagine is how a pump could be made to adapt to unanticipated metabolic demands by reorganizing itself in the way that hearts do (by increasing their mass) in well-

trained athletes or in populations living at high altitudes. Then, too, there is the fact that hearts are not constructed by any outside agency; rather, they spontaneously selforganize during embryonic development, a feat well beyond the capability of any pump yet conceived. Finally, there is the fact that the heart is composed of cells which are themselves capable of independent functional action, as is evidenced by the properties of automaticity and rhythmicity intrinsic to myocardial tissue (*ibid.*, p26), and still more vividly by the fact that dissociated cardiomyocytes placed in a suitable medium will spontaneously reaggregate into a rhythmically pulsating mass (RENSBERGER 1996, p17; see, also, MOSCONA 1959). On the complex relationship between the whole heart and its parts, see NOBLE (1998).

7 For an introduction to the mathematics of symmetry, see J. ROSEN (1995). For the role of symmetry and symmetry breaking in modern physics, see ICKE (1995). For applications to biology, as well as some philosophical implications, see HAHN (1998), HAHN/WEIBEL (1996), and MAINZER (1996).

8 For a review of condensed-matter physics, see CHAIKIN/LUBENSKY (1995). For a popular account, see DE GENNES/BADOZ (1996). While these works deal largely with inorganic matter, see the excellent text by GROSBERG/KHOKHLOV (1997) for a discussion of biological macromolecules from the perspective of condensed-matter physics.

9 Another closely related concept is that of "hierarchy". There is now a large and growing literature on "hierarchy theory" which in effect brings the doctrine of integrative levels up to date by giving that qualitative notion some quantitative rigor (e.g., AHL/ALLEN 1996; BAAS 1994; PETTERSSON 1996; SALTHE 1985). Still another related notion is that of "complexity"; on this contentious issue, see BARHAM (1999), and the references therein.

10 Many readers will no doubt feel uncomfortable with the viewpoint advocated here, which may seem tantamount to seeking the "essence of life"—a program that has been in bad odor for a long time. On the mainstream DARWINIAN view, of course, life is all a matter of "bricolage", "contingency", and "frozen accidents"; if there can be by definition no essence to find, then it is a waste of time looking for it. However, such a reflexive DARWINIAN response ignores the profound biochemical and molecular unity of all living things (GILBERT et al. 1996; SKULACHEV 1992). As CONWAY MORRIS (2000, p8) has pointed out, the "diversity of life is, in molecular terms, little more than skin deep". The real question is not whether life constitutes a genuine natural kind, but rather what its essential characteristics are. This question is now being addressed by a growing number of investigators (BEDAU 1996; DYSON 1999; FONTANA/BUSS 1994a, 1996; HO 1993, 1994; KAUFFMAN 1993, 1995, 1996; MARGULIS/SAGAN 1995; MORÁN et al. 1997; NORRIS et al. 1996; R. ROSEN 1991, 1996a; YATES 1994). In case anyone should wonder about the absence of any reference here to MATURANA and VARELA's notion of "autopoiesis", let me say that I agree with MAHNER and BUNGE that "[autopoiesis] seems to be nothing but a fancy synonym for both 'selforganization' and 'self-maintenance'" (1997, p144). It is imperative that we pose the question of the essence of life if we are to have any hope of understanding the place of mind in nature. However, it is equally important that we carefully distinguish the scientific motivation behind this endeavor from the irrationalist and subjectivist philosophical motivation of much of the contemporary discussion of autopoiesis and similar notions in postmodern and New Age circles.

11 It has to be acknowledged that there is still a good deal of controversy surrounding the thermodynamics of selforganization. In addition to the more or less mainstream accounts cited above, a number of more radical views have been advanced in recent years in an attempt to clarify the vexed cluster of concepts comprising selforganization, complexity, entropy, and information (BROOKS/WILEY 1988; ELITZUR 1994; MATSUNO 1984; SCHNEIDER/KAY 1994; SWENSON 1998; WICKEN 1987). Unfortunately, many of the details of these proposals are at variance with one another. Another problem is that many of these authors claim a close connection between the notions of information and entropy, if not their outright identification. However, as ELSASSER (1998, p46) has pointed out, this makes as much sense as identifying light with water just because the same formalism applies to wave phenomena in both media (see, also, DENBIGH/DENBIGH 1985). Nevertheless, even if our understanding of the thermodynamics of selforganization is not as secure as one would like, the existence of the phenomenon itself is not in dispute.

12 While the field of nonlinear dynamics is still in its formative stages and remains controversial (LEWIN 1999), it has finally begun to win a measure of mainstream acceptance (GALLAGHER/APPENZELLER 1999). The literature is by now immense, is still growing exponentially, and is widely dispersed across a dozen different disciplines. Here, I can do no more than indicate a few good points of entry. For the general mathematical background, see ABRAHAM/SHAW (1992), JACKSON (1991), and G. NICOLIS (1995). For some recent biological applications, see BAR-YAM (1997), SCHWEITZER (1997), and ZAK et al. (1997). Of the many popular accounts now available, two of the best are COVENEY/HIGHFIELD (1995) and MAINZER (1997). On the historical roots of the recent renaissance of nonlinear dynamics, especially in the work of Henri POINCARÉ, see BARROW-GREEN (1997), DIACU/HOLMES (1996), and GOROFF (1993).

13 I am not the first to propose that functions, understood normatively, imply the cognitive capacity to recognize the conditions of their own satisfaction. This idea may be found on occasion in the philosophical literature (BREWER 1995; MARTIN/PFEIFER 1986). The closely related notion that life and cognition are intimately connected at the most fundamental level may also be encountered in the scientific literature from time to time (BEDAU 1998; HESCHL 1990; STEWART 1996). However, none of these authors links conation to cognition via nonlinear dynamics in the way that I do here. On the other hand, the handful of authors who do acknowledge the potential importance of nonlinear dynamics for our understanding of life and mind (CHRISTENSEN 1996; WEBER/DEPEW 1996) remain intent on assimilating it to the CS worldview.

14 Obviously, the notion that information consists of "low-energy" constraints on a system cannot be given an absolute sense, but must be relativized to the type of "high-energy" interaction the function as a whole takes part in. The distinction is perhaps most apparent in the case of vision. When I look to see that the coast is clear before making a left turn in traffic, I am taking advantage of the fact that the photons correlated with the oncoming automobiles have a much lower kinetic energy than the machines themselves do. On the other hand, if I grope with my hands in a dark room, it is not so much a question of an energy differential as of a difference in vulnerability to kinetic energy of different parts of my body. Perhaps the best one can do by way of a general definition is to say that a "low-energy" or informational constraint is anything that does not in itself directly threaten the dynamical stability of the function.

15 It is of course true that the integration of development and evolution has become a major theme of evolutionary thought in recent years (GERHART/KIRSCHNER 1997; B.K. HALL 1998; RAFF 1996). However, it is equally true that a large number of recent studies of evolutionary dynamics—from comparative anatomical (B.K. HALL 1994; MOORE/WILLMER 1997; SANDERSON/HUFFORD 1996), paleontological (CONWAY MORRIS 1998; MCMENAMIN 1998), comparative molecular (MOORE/WILLMER 1997; Nagy 1998; Tabin et al. 1999), experimental (LENSKI/TRAVISANO 1994; TRAVISANO et al. 1995), and theoretical (FONTANA/BUSS 1994b) perspectives—all provide abundant evidence of the existence of inherent constraints on phylogenetic change via the canalization of development (see, also, ALBERCH 1989;CUMMINGS 1994; B. K. HALL 1996; HAROLD 1990; NEWMAN 1994; WEBSTER/GOODWIN 1996). The fact of convergent evolution has always represented one of the strongest empirical challenges to conventional selectionist thinking, and continues to do so today. In short, the fundamental tension between chance and necessity that has always plagued DARWINISM remains very much unresolved. Recently, there has been a renaissance in the theoretical study of biological form (MCGHEE 1999; MEINHARDT 1995; MURRAY 1997; PRUSINKIEWICZ/LINDENMAYER 1990; SCHMIDT-KITTLER/VOGEL 1991; THOMAS/REIF 1993), amounting to a revival of the "rational morphology" tradition of GOETHE, VON BAER, GEOFFROY SAINT-HILAIRE, OWEN, BATESON, THOMPSON, NEEDHAM, and WADDINGTON. (For the history of this alternative tradition of biological thought in the eighteenth and nineteenth centuries, see ASMA 1996, LENOIR 1989, and RUSSELL 1982 [though originally published in 1916, this work remains indispensable]; for the twentieth century, see BOWLER 1983, HARAWAY 1976, and REID 1985.) My contention is that the dynamical approach to function advocated herein holds the key to finally reconciling the rationalist and transformist traditions in biology (cf. WEBSTER/GOODWIN 1996; see, also, RESNIK 1994). In my view, nonlinear dynamics will achieve this, not by "deepening and renewing" DARWINISM (DEPEW/WEBER 1995, pxii), but by replacing it altogether with a comprehensive physical theory of the living state. The problem with selectionism is that it is superficial; it begs all of the really important questions. As Ball has put it, "as an explanation for natural form, natural selection is not entirely satisfying. Not because it is wrong, but because it says nothing about mechanism" (1999, p6). If and when a genuinely theoretical understanding of living things is achieved, it will become obvious in retrospect that the theory of natural selection which held us in its grip for so long was always just a phenomenological description—not an explanation—of the evolutionary process (BRADY 1982; SHIMONY 1993; THOM 1990a).

16 This is, of course, predicated on the assumption that classes of mental functions considered realistically as nonlinear oscillators constitute natural kinds. However, RUEGER/SHARP (1998) have challenged this idea by arguing that the dynamical concept of structural stability (as they say) should supersede that of natural kind. I certainly agree with them that the notion of dynamical stability (as I prefer to say) is absolutely crucial to articulating a position capable of navigating safely between the Scylla of reductionist atomism and the Charybdis of subjectivist holism. However, I think that the conflict between their position and the realistic view of biological functions as natural kinds is more apparent than real. I would simply note that the examples they adduce—lasers and solitons in optical fiber cables—are *artificial* structures. If mental kinds, in contrast, comprise *natural* solitons or other coherent structures in living matter, then there is no reason why dynamical stability cannot be associated with real natures.

17 I do not wish to imply that rejection of representationalism logically entails rejection of philosophical mechanism; indeed, I doubt that any of the authors cited would be prepared to abandon the CS worldview altogether. Even so, an increasing number of mainstream investigators with an explicit commitment to the CS worldview now recognize that pure representationalism is a dead-end, and that cognition must be viewed as a form of dynamical interaction (e.g., BICKHARD/TERVEEN 1995; CLARK 1997; HENDRIKS-JANSEN 1996). These authors are seeking a compromise position halfway between the CS and DE worldviews, as it were. Whether a coherent position in this location in fact exists, however, is another matter.

18 Needless to say, there are many problems to be worked out here. In order to discover what is distinctive about human, language-based societies, we must first achieve a deeper understanding of the dynamics of animal aggregations in general (see BONABEAU et al. 1999; PARRISH/EDELSTEIN-KESHET 1999). For this purpose, the distinction between inorganic and organic nonlinear oscillators (hurricanes vs. cells) is no longer adequate. In addition, we must now distinguish between nonteleologically organized aggregations of individual agents (e.g., ecosystems, markets), on the one hand, and teleologically organized collective agents (multicellular organisms, business firms), on the other. (See KHALIL 1990, which is one of the few works to give proper weight to this distinction.) Unfortunately, the difficulty of this task is compounded by the fact that some collectivities of individual agents are capable of switching back and forth between the teleological and the nonteleological forms of organization (e.g., populations of free-ranging *Dictyostelium* cells vs. fruiting bodies; hives and herds vs. swarms and stampedes; states at peace vs. states mobilized for war; etc.).

19 How do genes fit into this picture? First of all, it may be recalled that genes are relatively inert; it is enzymes that manipulate genes, not the other way around (TJIAN 1995). Second, the main purpose of gene-enzyme interactions is to regulate the functioning of the enzymes themselves. Thus, genes may be regarded as essentially constraints on enzyme function (NEWMAN 1994, p483; WEBSTER/GOODWIN 1996, pp209–211). As R. ROSEN has pointed out, since genes regulate the rates of enzyme activity (1996a, p184), one might say that they are "second-order enzymes". WÄCHTERSHÄUSER has even gone so far as to call nucleic acids "glorified coenzymes" (1997, p492). According to KAUFFMAN, a particular type of cell is "*…an attractor*, in the integrated dynamical behavior of a coupled system made up of thousands of genes and their products" (1993, p442, original emphasis). That is to say, cell functioning, or metabolism, may be thought of as a vast and densely interconnected regulatory network of tightly linked genes and proteins. (See, also, COHEN/RICE 1996; SAVAGEAU 1996). Furthermore, the recent demonstration of both the functional plasticity over the course of evolution of highly conserved homeogene "cassettes" (that is, coherent genetic regulatory modules – see EIZINGER et al. 1999; NAGY 1998; TABIN et al. 1999) and the convergence of nonhomologous genes on similar functions (BESCHIN et al. 1999; LEE 1999), as well as evidence for the real (if controversial) possibility that genetic mutation and recombination may themselves be under metabolic control (B.G. HALL 1997; JABLONKA/LAMB 1995; STEELE et al. 1998), also tends to reinforce the dynamical point of view. As a leading geneticist has noted, "Thinking about genetic change as a regulated biological function is fundamentally different from thinking about genetic change as the stochastic, accidental result of replication errors and physicochemical insults … Knowing about mechanisms for multiple coordinated changes in the genome, and about the potential for biological feedback onto genome restructuring, forces us to think of evolutionary genetic change in cell biological terms, not as fundamentally different from other kinds of cellular biochemistry" (SHAPIRO 1997, p103). (See, also, CONWAY MORRIS, 2000; GILBERT et al. 1996.) In general, there is now a growing recognition that "[m]any claims about genes that pass for truisms in the broader intellectual community are unsustainable. DNA does not contain a program for development. Genes are not 'self-replicating' … Genes are not the sole or main units of selection" (GRIFFITHS/KNIGHT 1998, pp254–255). (See, also, LEWONTIN, 2000; NIJHOUT 1990; SARKAR 1996.) In short, contemporary mainstream genetics research itself increasingly demands a dynamical interpretation of gene function. It is interesting to note that this view of genes as second-order enzymes is consistent with DYSON's (1999) theory that genes originated as metabolic byproducts (ATP fragments) which became first parasites, then symbionts, of the "homeostatic" dynamics (as DYSON puts it) of early cells.

20 The first step, of course, will be to reach an objective understanding of the transition from the quantum to the classical domain in general. Here, the decoherence approach looks promising (BUB 1999; OMNÈS 1999). (On the historical background, see CUSHING 1994, and WICK 1996.) AMANN's attempt to develop an "individual" quantum mechanics for chemistry in which "'symmetry breaking' … corresponds to generation of attractors in the phase space" (1996, p91) is particularly interesting from an emergentist perspective. Of course, only time will tell whether this or any of the other approaches currently envisaged is correct. But even assuming that quantum field theory can ultimately be expanded to account for the macroscopic world in general, there still remains the problem of living matter in particular. The most difficult challenge here is clearly what to make of the orthogonality requirement. Clearly no theory involving an ordinary action principle will pass muster. Actually, this is already evident on empirical grounds alone, since a functionally active, fully folded protein is not at an energy minimum, but rather has a highly degenerate ground state consisting of numerous isoenergetic "conformational substates" (FRAUENFELDER et al. 1991b). It has been suggested that this property of proteins arises from an interplay of physical forces analogous to "frustration" in certain amorphous solids known as "spin glasses" (AUSTIN/CHEN 1992). The significance of this is that in such systems "the shapes of the interacting elements and/or the form of their interactions are sufficiently complicated that it is very difficult—usually impossible, in fact—for them to find the true minimum of the free energy, and so they have to be content with sliding from one metastable state to another" (LEGGETT 1987, p134). In contrast with inorganic spin glasses, however, in proteins the transitions between conformational substates are not purely stochastic—that is, they cannot be explained in purely statistical mechanical terms (FRAUENFELDER/WOLYNES 1994, p61). Such nonergodic transitions between isoenergetic substates in proteins have been termed "functionally important motions" (FRAUENFELDER 1987, 1988, 1995). Such a phenomenon is unheard-of in nonbiological materials. As AUSTIN and CHEN have amusingly remarked, "No one has ever spoken about a functionally important motion in a [nonbiological] spin glass, and probably would be driven from the high holy temple of condensed matter physics if they did" (1992, p214). And yet such motions clearly exist in proteins. This means that the functional behavior of proteins is not fully explicable in terms of the known laws of physics. Where might we turn to seek an explanation of this puzzling behavior in terms of an *unknown* law of physics? FRAUENFELDER and WOLYNES have hypothesized that proteins may satisfy a "principle of minimal frustration" (1994, p61). This proposal seems to me to be of the utmost importance. Perhaps one might contemplate combining it with one of the various quasiparticle or soliton models of coherent electron transfer in living matter (for instance, the kink-like domain wall model of INSINNA et al. 1996). (See, also, LOMDAHL 1990; for a survey of similar suggestions, see MILLER 1991.) In this way, functionally important motions in proteins might be interpreted as quasiparticles or collective modes with a conserved quantum number ("topological charge") corresponding to frustration. In stressing the paramount importance of the "protein state of matter", I do not of course mean to imply that proteins should be considered alive in and of themselves. Ultimately, the living state must be understood as a whole in terms of the physical coupling of each class of macromolecule with each of the others, and with water. It is doubtful whether the behavior of any of the cellular components can be fully understood in isolation from the rest. (For example, in the real living cell, membranes are active structures exhibiting functional behavior—a far cry from the idealization of the inert amphiphilic bilayer; see LEIBLER 1991.) But if the cell must ultimately be understood as a whole, the surest avenue of progress towards that understanding would nevertheless seem to be via the physics of the protein state of matter.

References

Abraham, R. H. (1985) Is there chaos without noise? In: Fischer, P./Smith, W. R. (eds) Chaos, fractals, and dynamics. Marcel Dekker: New York, pp. 117–121.

Abraham, R. H./Shaw, C. D. (1992) Dynamics—The geometry of behavior, 2nd ed. Addison-Wesley: Redwood City CA.

Ahl, V./Allen, T. F. H. (1996) Hierarchy theory. Columbia University Press: New York.

Alberch, P. (1989) The logic of monsters: Evidence for internal constraint in development and evolution. Géobios, mémoire spécial 12: 21–57.

Alon, U./Surette, M. G./Barkai, N./Leibler, S. (1999) Robustness in bacterial chemotaxis. Nature 397: 168–171.

Amann, A. (1996) Can quantum mechanics account for chemical structures? In: Gans, W./Amann, A./Boeyens, J. C. A. (eds) Fundamental principles of molecular model-

ing. Plenum: New York, pp. 55–97.
Amundson, R./Lauder, G. V. (1998) Function without purpose: The uses of causal role function in evolutionary biology. In: Allen, C./Bekoff, M./Lauder, G. (eds) Nature's purposes. Bradford Books/MIT Press: Cambridge MA, pp. 335–369. (Originally published in 1994.)
Anderson, P. W. (1994) More is different. In: Anderson, P. W., A career in theoretical physics. World Scientific: Singapore, pp. 1–4. (Originally published in 1972.)
Anderson, P. W. (1995) Historical overview of the twentieth century in physics. In: Brown, L. M./Pais, A./Pippard, B. (eds) Twentieth century physics, vol. 3. Institute of Physics Publishing: Bristol UK/American Institute of Physics: New York, pp. 2017–2032.
Anscombe, G. E. M. (1993) Causality and determination. In: Sosa, E./Tooley, M. (eds) Causation. Oxford University Press: Oxford, pp. 88–104. (Originally published in 1971.)
Aristotle (1961) Parts of animals. Harvard University Press: Cambridge MA.
Asma, S. T. (1996) Following form and function. Northwestern University Press: Evanston IL.
Austin, R. H./Chen, C. M. (1992) The spin-glass analogy in protein dynamics. In: Stein, D. L. (ed) Spin glasses and biology. World Scientific: Singapore, pp. 179–223.
Auyang, S. Y. (1995) How is quantum field theory possible? Oxford University Press: New York.
Baas, N. A. (1994) Emergence, hierarchies, and hyperstructures. In: Langton, C. G. (ed) Artificial life III. Addison-Wesley: Reading MA, pp. 515–537.
Bak, P. (1996) How nature works. Copernicus/Springer: New York.
Ball, P. (1999) The self-made tapestry. Oxford University Press: Oxford.
Barham, J. (1990) A Poincaréan approach to evolutionary epistemology. Journal of Social and Biological Structures 13: 193–258.
Barham, J. (1992) From enzymes to E=mc2: A reply to critics. Journal of Social and Evolutionary Systems (continues Journal of Social and Biological Structures) 15: 249–305.
Barham, J. (1995) Ni Darwin ni Derrida. Journal of Social and Evolutionary Systems 18: 277–308.
Barham, J. (1996) A dynamical model of the meaning of information. BioSystems 38: 235–241.
Barham, J. (1999) On the objectivity of the scala naturae. Evolution and Cognition 5(1): 2–11.
Barham, J. (in preparation) Neither ghost nor machine.
Barkai, N./Leibler, S. (1997) Robustness in simple biochemical networks. Nature 387: 913–917.
Barrow-Green, J. (1997) Poincaré and the three body problem (History of mathematics, vol. 11). American Mathematical Society: Providence RI/London Mathematical Society: London.
Bar-Yam, Y. (1997) Dynamics of complex systems. Addison-Wesley: Reading MA.
Batterman, R. W. (2000) Multiple realizability and universality. British Journal for the Philosophy of Science 51: 115–145.
Beatty, J. (1992) Fitness: Theoretical contexts. In: Keller, E. F./Lloyd, E. A. (eds) Keywords in evolutionary biology. Harvard University Press: Cambridge MA, pp. 115–119.
Bechtel, W./Mundale, J. (1999) Multiple realizability revisited: Linking cognitive and neural states. Philosophy of Science 66: 175–207.
Bechtel, W./Richardson, R. C. (1998) Vitalism. In: Craig, E. (ed) Routledge encyclopedia of philosophy, vol. 9. Routledge: London, pp. 639–643.
Bedau, M. (1990) Against mentalism in teleology. American Philosophical Quarterly 27: 61–70.
Bedau, M. (1992) Where's the good in teleology? Philosophy and Phenomenological Research 52: 781–806.
Bedau, M. A. (1996) The nature of life. In: Boden, M. A. (ed) The philosophy of artificial life. Oxford University Press: New York, pp. 332–357.
Bedau, M. A. (1997) Weak emergence. In: Tomberlin, J. E. (ed) Philosophical perspectives, vol. 11. Blackwell: Malden MA, pp. 375–399.
Bedau, M. A. (1998) Four puzzles about life. Artificial Life 4: 125–140.
Beer, R. D. (1995) A dynamical systems perspective on agent-environment interaction. Artificial Intelligence 72: 173–215.
Beer, R. D. (1997) The dynamics of adaptive behavior: A research program. Robotics and Autonomous Systems 20: 257–289.
Bennett, J. (1990) Linguistic behaviour. Hackett: Indianapolis IN. (Originally published in 1976.)
Berne, R. M./Levy, M. N. (1997) Cardiovascular physiology, 7th ed. Mosby Year Book: St. Louis.
Beschin, A./Bilej, M./Brys, L./Torreele, E./Lucas, R./Magez, S./De Baetselier, P. (1999) Convergent evolution of cytokines. Nature 400: 627–628.
Bhalla, U. S./Iyengar, R. (1999) Emergent properties of networks of biological signaling pathways. Science 283: 381–387.
Bickhard, M. H./Terveen, L. (1995) Foundational issues in artificial intelligence and cognitive science. Elsevier: Amsterdam.
Blitz, D. (1992) Emergent evolution. Kluwer Academic: Dordrecht, Holland.
Boden, M. A. (1999) Is metabolism necessary? British Journal for the Philosophy of Science 50: 231–248.
Bonabeau, R./Dorigo, M./Theraulaz, G. (1999) Swarm intelligence. Oxford University Press: New York.
Bowler, P. J. (1983) The eclipse of Darwinism. Johns Hopkins University Press: Baltimore.
Brady, R. H. (1982) Dogma and doubt. Biological Journal of the Linnean Society 17: 79–96.
Brandon, R. N. (1990) Adaptation and environment. Princeton University Press: Princeton NJ.
Bray, D. (1995) Protein molecules as computational elements in living cells. Nature 376: 307–312.
Brewer, B. (1995) Compulsion by reason. Proceedings of the Aristotelian Society, supp. 69: 237–253.
Bricmont, J. (1997) Science of chaos or chaos in science? In: Gross, P. R./Levitt, N./Lewis, M. W. (eds) The flight from science and reason. Johns Hopkins University Press: Baltimore, pp. 131–175.
Brooks, D. R./Wiley, E. O. (1988) Evolution as entropy, 2nd ed. University of Chicago Press: Chicago.
Brooks, R. A. (1999) Cambrian intelligence. Bradford Books/MIT Press: Cambridge MA.
Brown, L. M./Cao, T. Y. (1991) Spontaneous breakdown of symmetry: Its rediscovery and integration into quantum field theory. Historical Studies in the Physical and Biological Sciences 21: 211–235.
Bub, J. (1999) Interpreting the quantum world (corrected paperback edition). Cambridge University Press: Cambridge.
Buller, D. J. (1999) Introduction: Natural teleology. In: Buller, D. J. (ed) Function, selection, and design. State University of New York Press: Albany NY, pp. 1–27.
Bunge, M. (1991) The power and limits of reduction. In: Agazzi, E. (ed) The problem of reductionism in science. Kluwer

Academic: Dordrecht, Holland, pp. 31–49.

Campbell, J. H. (1985) An organizational interpretation of evolution. In: Depew, D. J./Weber, B. H. (eds) Evolution at a crossroads. Bradford Books/MIT Press: Cambridge MA, pp. 133–167.

Canfield, J. V. (1990) The concept of function in biology. Philosophical Topics 18(2): 29–53.

Cao, T. Y. (1993) New philosophy of renormalization: From the renormalization group equations to effective field theories. In: Brown, L. M. (ed) Renormalization. Springer: New York, pp. 87–133.

Cao, T. Y. (1997) Conceptual developments of 20th century field theories. Cambridge University Press: Cambridge.

Cao, T. Y. (1998) Monism, but not through reductionism. In: Cohen, R. S./Tauber, A. I. (eds) Philosophies of nature: The human dimension. Kluwer Academic: Dordrecht, Holland, pp. 39–51.

Cao, T. Y./Schweber, S. S. (1993) The conceptual foundations and the philosophical aspects of renormalization theory. Synthese 97: 33–108.

Chaikin, P. M./Lubensky, T. C. (1995) Principles of condensed matter physics. Cambridge University Press: Cambridge.

Christensen, W. (1996) A complex systems theory of teleology. Biology and Philosophy 11: 301–320.

Clark, A. (1997) Being there. Bradford Books/MIT Press: Cambridge MA.

Cohen, J./Rice, S. H. (1996) Where do biochemical pathways lead? In: Collado-Vides, J./Magasanik, B./Smith, T. F. (eds) Integrative approaches to molecular biology. MIT Press: Cambridge MA, pp. 239–251.

Conrad, M. (1997) Origin of life and the underlying physics of the universe. BioSystems 42: 177–190.

Conrad, M. (1998) Quantum gravity and life. BioSystems 46: 29–39.

Conway Morris, S. (1998) The crucible of creation. Oxford University Press: New York.

Conway Morris, S. (2000) Evolution: Bringing molecules into the fold. Cell 100: 1–11.

Coveney, P./Highfield, R. (1995) Frontiers of complexity. Fawcett Columbine: New York.

Crane, T./Mellor, D. H. (1995) There is no question of physicalism. In: Moser, P. K./Trout, J. D. (eds) Contemporary materialism. Routledge: London, pp. 65–89. (Originally published in 1990.)

Cummings, F. W. (1994) Aspects of growth and form. Physica D 79: 146–163.

Cummins, R. (1998) Functional analysis. In: Allen, C./Bekoff, M./Lauder, G. (eds) Nature's purposes. Bradford Books/MIT Press: Cambridge MA, pp. 169–196. (Originally published in 1975.)

Curie, P. (1908) Sur la symétrie dans les phénomènes physiques, symétrie d'un champ électrique et d'un champ magnétique. In: Oeuvres de Pierre Curie. Gauthier-Villars: Paris, pp. 118–141.

Cushing, J. T. (1994) Quantum mechanics: Historical contingency and the Copenhagen hegemony. University of Chicago Press: Chicago.

Cushing, J. T. (1998) Philosophical concepts in physics. Cambridge University Press: Cambridge.

Davies, P. S. (2000) The nature of natural norms: Why selected functions are systemic capacity functions. Noûs 34: 85–107.

de Gennes, P.-G./Badoz, J. (1996) Fragile objects. Copernicus/Springer: New York. (Originally published in French in 1994.)

Delattre, P. (1986) An approach to the notion of finality according to the concepts of qualitative dynamics. In: Diner, S./Fargue, D./Lochak, G. (eds) Dynamical systems: A renewal of mechanism. World Scientific: Singapore, pp. 149–154.

Del Giudice, E./Doglia, S./Milani, M./Vitiello, G. (1988) Structures, correlations and electromagnetic interactions in living matter: Theory and applications. In: Fröhlich, H. (ed) Biological coherence and response to external stimuli. Springer: Berlin, pp. 49–64.

Del Giudice, E./Doglia, S./Vitiello, G. (1989) Self-organization and symmetry breaking in living matter. In: Gruber, B./Iachello, F. (eds) Symmetries in science III. Plenum: New York, pp. 379–387.

Denbigh, K. G. (1975) An inventive universe. George Braziller: New York.

Denbigh, K. G. (1981) Three concepts of time. Springer: Berlin.

Denbigh, K. G./Denbigh, J. S. (1985) Entropy in relation to incomplete knowledge. Cambridge University Press: Cambridge.

Dennett, D. C. (1987) The intentional stance. Bradford Books/MIT Press: Cambridge MA.

Depew, D. J./Weber, B. H. (1995) Darwinism evolving. Bradford Books/MIT Press: Cambridge MA.

Diacu, F./Holmes, P. (1996) Celestial encounters. Princeton University Press: Princeton NJ.

Dresden, M. (1974) Reflections on 'fundamentality and complexity'. In: Enz, C. P./Mehra, J. (eds) Physical reality and mathematical description. Reidel: Dordrecht, Holland, pp. 133–166.

Dressler, D./Potter, H. (1991) Discovering enzymes. Scientific American Library: New York.

Dretske, F. (1988) Explaining behavior. Bradford Books/MIT Press: Cambridge MA.

Drexler, K. E. (1989) Biological and nanomechanical systems: Contrasts in evolutionary capacity. In: Langton, C. G. (ed) Artificial life. Addison-Wesley: Redwood City CA, pp. 501–519.

Dupré, J. (1993) The disorder of things. Harvard University Press: Cambridge MA.

Dyson, F. (1999) Origins of life, 2nd ed. Cambridge University Press: Cambridge.

Eizinger, A./Jungblut, B./Sommer, R. J. (1999) Evolutionary change in the functional specificity of genes. Trends in Genetics 15: 197–202.

Elitzur, A. C. (1994) Let there be life: Thermodynamic reflections on biogenesis and evolution. Journal of Theoretical Biology 168: 429–459.

Elsasser, W. M. (1998) Reflections on a theory of organisms. Johns Hopkins University Press: Baltimore. (Originally published in 1987.)

Emmeche, C./Køppe, S./Stjernfelt, F. (1997) Explaining emergence: Towards an ontology of levels. Journal for General Philosophy of Science 28: 83–119.

Enç, B./Adams, F. (1998) Functions and goal-directedness. In: Allen, C./Bekoff, M./Lauder, G. (eds) Nature's purposes. Bradford Books/MIT Press: Cambridge MA, pp. 371–394. (Originally published in 1992.)

Falk, A. E. (1995) Essay on nature's semeiosis. Journal of Philosophical Research 20: 296–348.

Favre, A./Guitton, H./Guitton, J./Lichnerowicz, A./Wolff, E. (1995) Chaos and determinism. Johns Hopkins University Press: Baltimore. (Originally published in French in 1988.)

Fisher, M. E. (1988) Condensed matter physics: Does quantum mechanics matter? In: Feshbach, H./Matsui, T./Oleson, A. (eds) Niels Bohr: Physics and the world. Harwood Academic Publishers: Chur, Switzerland, pp. 65–115.

Flyvbjerg, H./Bak, P./Jensen, M. H./Sneppen, K. (1995) A self-organized critical model for evolution. In: Mosekilde, E./Mouritsen, O. G. (eds) Modelling the dynamics of biological systems. Springer: Berlin, pp. 269–288.

Fodor, J. A. (1990) A theory of content and other essays. Bradford Books/MIT Press: Cambridge MA.

Fontana, W./Buss, L. W. (1994a) 'The arrival of the fittest': Towards a theory of biological organization. Bulletin of Mathematical Biology 56: 1–64.

Fontana, W./Buss, L. W. (1994b) What would be conserved if 'the tape were played twice'? Proceedings of the National Academy of Sciences USA 91: 757–761.

Fontana, W./Buss, L. W. (1996) The barrier of objects: From dynamical systems to bounded organizations. In: Casti, J. L./Karlqvist, A. (eds) Boundaries and barriers. Addison-Wesley: Reading MA, pp. 55–115.

Fontana, W./Wagner, G./Buss, L. W. (1995) Beyond digital naturalism. In: Langton, C. G. (ed) Artificial life. Bradford Books/MIT Press: Cambridge MA, pp. 211–227. (Originally published in 1994.)

Frauenfelder, H. (1987) Function and dynamics of myoglobin. In: Koslow, S. H./Mandell, A. J./Shlesinger, M. F. (eds) Perspectives in biological dynamics and theoretical medicine (Annals of the New York Academy of Sciences, vol. 504), New York Academy of Sciences: New York, pp. 151–167.

Frauenfelder, H. (1988) Protein complexity. In: Markus, M./Müller, S. C./Nicolis, G. (eds) From chemical to biological organization. Springer: Berlin, pp. 177–181.

Frauenfelder, H. (1995) Proteins—Paradigms of complex systems. Experientia 51: 200–203.

Frauenfelder, H./Wolynes, P. G. (1994) Biomolecules: Where the physics of complexity and simplicity meet. Physics Today 47(2): 58–64.

Frauenfelder, H./Chu, K./Philipp, R. (1991a) Physics from proteins. In: Peliti, L. (ed) Biologically inspired physics. Plenum: New York, pp. 1–14.

Frauenfelder, H./Sligar, S. G./Wolynes, P. G. (1991b) The energy landscapes and motions of proteins. Science 254: 1598–1603.

Freeman, W. J. (1992) Tutorial on neurobiology: From single neurons to brain chaos. International Journal of Bifurcation and Chaos 2: 451–482.

Freeman, W. J. (1995a) Societies of brains. Lawrence Erlbaum Assocs.: Hillsdale NJ.

Freeman, W. J. (1995b) The creation of perceptual meanings in cortex through chaotic itinerancy and sequential state transitions induced by sensory stimuli. In: Kruse, P./Stadler, M. (eds) Ambiguity in mind and nature. Springer: Berlin, pp. 421–437.

Freeman, W. J. (1999) Consciousness, intentionality and causality. Journal of Consciousness Studies 6: 143–172.

Frieden, B. R. (1998) Physics from Fisher information. Cambridge University Press: Cambridge.

Fröhlich, H. (1988) Theoretical physics and biology. In: Fröhlich, H. (ed) Biological coherence and response to external stimuli. Springer: Berlin, pp. 1–24.

Fruton, J. S. (1999) Proteins, enzymes, genes. Yale University Press: New Haven CT.

Fry, I. (1995) Are the different hypotheses on the emergence of life as different as they seem? Biology and Philosophy 10: 389–417.

Galison, P./Stump, D. J. (eds) (1996) The disunity of science. Stanford University Press: Stanford CA.

Gallagher, R./Appenzeller, T. (1999) Beyond reductionism. Science 284: 79. (Introduction to a special section on complex systems comprising ten articles by various authors, pp. 80–109.)

George, F./Johnson, L. (eds) (1985) Purposive behaviour and teleological explanations. Gordon and Breach: New York.

Georgi, H. M. (1989) Effective quantum field theories. In: Davies, P. (ed) The new physics. Cambridge University Press: Cambridge, pp. 446–457.

Gerhart, J./Kirschner, M. (1997) Cells, embryos, and evolution. Blackwell Science: Malden MA.

Gilbert, S. F./Opitz, J. M./Raff, R. A. (1996) Resynthesizing evolutionary and developmental biology. Developmental Biology 173: 357–372.

Godfrey-Smith, P. (1999) Functions: Consensus without unity. In: Buller, D. J. (ed) Function, selection, and design. State University of New York Press: Albany NY, pp. 185–197. (Originally published in 1994.)

Goldbeter, A. (1996) Biochemical oscillations and cellular rhythms. Cambridge University Press: Cambridge.

Goroff, D. L. (1993) Introduction: Henri Poincaré and the birth of chaos theory. In: Poincaré, H. New methods of celestial mechanics, vol. 1, American Institute of Physics: New York, pp. I1–I107.

Griffiths, P. E./Knight, R. D. (1998) What is the developmentalist challenge? Philosophy of Science 65: 253–258.

Grosberg, A.Yu./Khokhlov, A. R. (1997) Giant molecules. Academic Press: San Diego CA. (Revised edition of a work originally published in Russian in 1989.)

Hahn, W. (1998) Symmetry as a developmental principle in nature and art. World Scientific: Singapore. (Originally published in German in 1989.)

Hahn, W./Weibel, P. (eds) (1996) Evolutionäre Symmetrietheorie. Hirzel: Stuttgart.

Haken, H. (1983) Synergetics, 3rd ed. Springer: Berlin.

Hall, B. G. (1997) On the specificity of adaptive mutations. Genetics 145: 39–44.

Hall, B. K. (ed) (1994) Homology. Academic Press: San Diego CA.

Hall, B. K. (1996) Baupläne, phylotypic stages, and constraint: Why there are so few types of animals. In: Hecht, M.K./MacIntyre, R.J./Clegg, M.T. (eds) Evolutionary biology, vol. 29. Plenum Press: New York, pp. 215–261.

Hall, B. K. (1998) Evolutionary developmental biology, 2nd ed. Chapman & Hall: London.

Hameroff, S. R./Scott, A. C. (1998) A Sonoran afternoon. In: Hameroff, S. R./Kaszniak, A. W./Scott, A. C. (eds) Toward a science of consciousness II. Bradford Books/MIT Press: Cambridge MA, pp. 635–644.

Haraway, D. J. (1976) Crystals, fabrics, and fields. Yale University Press: New Haven CT.

Harold, F. M. (1990) To shape a cell: An inquiry into the causes of morphogenesis of microorganisms. Microbiological Reviews 54: 381–431.

Harré, R./Madden, E. H. (1975) Causal powers. Blackwell: Oxford.

Harrison, L. G. (1993) Kinetic theory of living pattern. Cambridge University Press: Cambridge.

Heald, R./Tournebize, R./Blank, T./Sandaltzopoulos, R./Becker, P./Hyman, A./Karsenti, E. (1996) Self-organization of microtubules into bipolar spindles around artificial chromosomes in Xenopus egg extracts. Nature 382: 420–425.

Hendriks-Jansen, H. (1996) Catching ourselves in the act. Bradford Books/MIT Press: Cambridge MA.

Heschl, A. (1990) L=C: A simple equation with astonishing consequences. Journal of Theoretical Biology 145: 13–40.

Ho, M.-W. (1993) The rainbow and the worm. World Scientific: Singapore.

Ho, M.-W. (1994) What is (Schrödinger's) negentropy? In: Gnaiger, E./Gellerich, F. N./Wyss, M. (eds) What is controlling life? (Modern Trends in BioThermoKinetics, vol. 3). Innsbruck University Press: Innsbruck, Austria, pp. 50–61.

Ho, M.-W. (1996) The biology of free will. Journal of Consciousness Studies 3: 231–244.

Hoffmeyer, J. (1996) Signs of meaning in the universe. Indiana University Press: Bloomington IN.

Hogan, C. J. (1998) The little book of the big bang. Copernicus/Springer: New York.

Horgan, T. (1993) From supervenience to superduppervenience: Meeting the demands of a material world. Mind 102: 555–586.

Horst, S. W. (1996) Symbols, computation, and intentionality. University of California Press: Berkeley CA.

Hyde, S./Andersson, S./Larsson, K./Blum, Z./Landh, T./Lidin, S./Ninham, B. W. (1997) The language of shape. Elsevier: Amsterdam.

Icke, V. (1995) The force of symmetry. Cambridge University Press: Cambridge.

Insinna, E. M./Zaborski, P./Tuszynski, J. (1996) Electrodynamics of microtubular motors: The building blocks of a new model. BioSystems 39: 187–226.

Jablonka, E./Lamb, M. J. (1995) Epigenetic inheritance and evolution. Oxford University Press: Oxford.

Jackson, E. A. (1991) Perspectives of nonlinear dynamics (corrected paperback edition), 2 vols. Cambridge University Press: Cambridge.

Juarrero, A. (1999) Dynamics in action. Bradford Books/MIT Press: Cambridge MA.

Kant, I. (1987) Critique of judgment. Hackett: Indianapolis IN. (Originally published in German in 1790.)

Kauffman, S. A. (1993) The origins of order. Oxford University Press: New York.

Kauffman, S. A. (1994) Whispers from Carnot: The origins of order and principles of adaptation in complex nonequilibrium systems. In: Cowan, G. A./Pines, D./Meltzer, D. (eds) Complexity. Addison-Wesley: Reading MA, pp. 83–160.

Kauffman, S. A. (1995) 'What is life?': Was Schrödinger right? In: Murphy, M. P./O'Neill, L.A.J. (eds) What is life? The next fifty years. Cambridge University Press: Cambridge, pp. 83–114.

Kauffman, S. A. (1996) The nature of autonomous agents and the worlds they mutually create. http://www.santafe.edu/sfi/People/kauffman/Investigations.html.

Keijzer, F. A. (1998) Doing without representations which specify what to do. Philosophical Psychology 11: 269–302.

Kellert, S. H. (1993) In the wake of chaos. University of Chicago Press: Chicago.

Kelso, J.A.S. (1995) Dynamic patterns. Bradford Books/MIT Press: Cambridge MA.

Khalil, E. L. (1990) Natural complex vs. natural system. Journal of Social and Biological Structures 13: 11–31.

Kim, J. (1998) Mind in a physical world. Bradford Books/MIT Press: Cambridge MA.

Kitcher, P. (1998) Function and design. In: Allen, C./Bekoff, M./Lauder, G. (eds) Nature's purposes. Bradford Books/MIT Press: Cambridge MA, pp. 479–503. (Originally published in 1993.)

Kondepudi, D./Prigogine, I. (1998) Modern thermodynamics. John Wiley & Sons: Chichester UK.

Lahav, N. (1999) Biogenesis. Oxford University Press: New York.

Lambert, D. M. (1995) Biological function: Two forms of explanation. In: Lambert, D. M./Spencer, H. G. (eds) Speciation and the recognition concept. Johns Hopkins University Press: Baltimore, pp. 238–259.

Layzer, D. (1990) Cosmogenesis. Oxford University Press: New York.

Leclerc, I. (1984) The philosophy of nature. Catholic University of America Press: Washington.

Lee, M.S.Y. (1999) Molecular phylogenies become functional. Trends in Ecology and Evolution 14: 177–178.

Leggett, A. J. (1987) The problems of physics. Oxford University Press: Oxford.

Leibler, S. (1991) Recent developments in the physics of fluctuating membranes. In: Peliti, L. (ed) Biologically inspired physics. Plenum: New York, pp. 81–94.

Lennox, J. G. (1992) Teleology. In: Keller, E. F./Lloyd, E. A. (eds) Keywords in evolutionary biology. Harvard University Press: Cambridge MA, pp. 324–333.

Lenoir, T. (1989) The strategy of life. University of Chicago Press: Chicago. (Originally published in 1982.)

Lenski, R. E./Travisano, M. (1994) Dynamics of adaptation and diversification: A 10,000-generation experiment with bacterial populations. Proceedings of the National Academy of Sciences USA 91: 6808–6814.

Levinthal, M. (1991) The use and misuse of metaphor in biology: Studies of the semantics of molecular evolution. In: Anderson, M./Merrell, F. (eds) On semiotic modeling. Mouton de Gruyter: Berlin, pp. 167–184.

Lewin, R. (1999) Complexity, 2nd ed. University of Chicago Press: Chicago.

Lewontin, R. (2000) The dream of the human genome. In: Lewontin, R. It ain't necessarily so. New York Review Books: New York, pp. 133–186. (Originally published in 1992.)

Loewenstein, W. R. (1999) The touchstone of life. Oxford University Press: New York.

Lomdahl, P. S. (1990) Soliton models of protein dynamics. In: Fordy, A. P. (ed) Soliton theory. Manchester University Press: Manchester UK, pp. 208–232.

Lorenz, K. (1962) Kant's doctrine of the a priori in the light of contemporary biology. General Systems 7: 23–35. (Originally published in German in 1941.)

Luisi, P. L. (1993) Defining the transition to life: Self-replicating bounded structures and chemical autopoiesis. In: Stein, W. D./Varela, F. J. (eds) Thinking about biology. Addison-Wesley: Reading MA, pp. 17–39.

Mahner, M/Bunge, M. (1997) Foundations of biophilosophy. Springer: Berlin.

Mainzer, K. (1996) Symmetries of nature. Walter de Gruyter: Berlin.

Mainzer, K. (1997) Thinking in complexity, 3rd ed. Springer: Berlin. (Originally published in German in 1988.)

Manning, R. N. (1997) Biological function, selection, and reduction. British Journal for the Philosophy of Science 48: 69–82.

Margulis, L./Sagan, D. (1995) What is life? Simon & Schuster: New York.

Martin, C. B./Pfeifer, K. (1986) Intentionality and the non-psychological. Philosophy and Phenomenological Research 46: 531–554.

Matsuno, K. (1984) Open systems and the origin of protoreproductive units. In: Ho, M.-W./Saunders, P. T. (eds) Beyond neo-darwinism. Academic: London, pp. 61–88.

Matsuno, K. (1989) Protobiology. CRC Press: Boca Raton FL.

Mayr, E. (1988a) Cause and effect in biology. In: Mayr, E. Toward a new philosophy of biology. Harvard University Press: Cambridge MA, pp. 24–37. (Originally published in

1961.)
Mayr, E. (1988b) The multiple meanings of teleological. In: Mayr, E. Toward a new philosophy of biology. Harvard University Press: Cambridge MA, pp. 38–66. (Originally published in 1974.)
McGhee, G. R. (1999) Theoretical morphology. Columbia University Press: New York.
McLaughlin, B. P. (1992) The rise and fall of British emergentism. In: Beckermann, A./Flohr, H./Kim, J. (eds) Emergence or reduction? Walter de Gruyter: Berlin, pp. 49–93.
McMenamin, M.A.S. (1998) The garden of Ediacara. Columbia University Press: New York.
Meehl, P. E./Sellars, W. (1956) The concept of emergence. In: Feigl, H./Scriven, M. (eds) Minnesota studies in the philosophy of science, vol. 1. University of Minnesota Press: Minneapolis MN, pp. 239–252.
Meinhardt, H. (1995) The algorithmic beauty of sea shells. Springer: Berlin.
Mikulecky, D. C. (1995) Topological determinants of dynamic molecular properties: A network thermodynamic approach. Journal of Molecular Structure (Theochem) 336: 279–292.
Mikulecky, D. C. (1996) Complexity, communication between cells, and identifying the functional components of living systems: Some observations. Acta Biotheoretica 44: 179–208.
Miller, D. A. (1991) Useful perspective on the relation between biological and physical descriptions of phenomena. Journal of Theoretical Biology 152: 341–355.
Millikan, R. G. (1984) Language, thought, and other biological categories. Bradford Books/MIT Press: Cambridge MA.
Millikan, R. G. (1991) Speaking up for Darwin. In: Loewer, B./Rey, G. (eds) Meaning in mind. Blackwell: Oxford, pp. 151–164.
Millikan, R. G. (1993) White Queen psychology and other essays for Alice. Bradford Books/MIT Press: Cambridge MA.
Millikan, R. G. (1996) On swampkinds. Mind and Language 11: 103–117.
Millikan, R. G. (1999) An ambiguity in the notion ‘function’. In: Buller, D. J. (ed) Function, selection, and design. State University of New York Press: Albany NY, pp. 115–121. (Originally published in 1989.)
Monod, J. (1972) Chance and necessity. Vintage Books/Random House: New York. (Originally published in French in 1970.)
Moore, J./Willmer, P. (1997) Convergent evolution in invertebrates. Biological Reviews of the Cambridge Philosophical Society 72: 1–60.
Morán, F./Moreno, A./Minch, E./Montero, F. (1997) Further steps towards a realistic description of the essence of life. In: Langton, C. G./Shimohara, K. (eds) Artificial life V. Bradford Books/MIT Press: Cambridge MA, pp. 255–263.
Morowitz, H. J. (1979) Energy flow in biology. Ox Bow Press: Woodbridge CT. (Originally published in 1968.)
Morowitz, H. J. (1992) Beginnings of cellular life. Yale University Press: New Haven CT.
Morowitz, H. J. (1999) A theory of biochemical organization, metabolic pathways, and evolution. Complexity 4(6): 39–53.
Moscona, A. A. (1959) Tissues from dissociated cells. Scientific American 200(5): 132–144.
Murray, J. D. (1997) Mathematical biology, 2nd ed. Springer: New York.
Nagel, E. (1998) Teleology revisited. In: Allen, C./Bekoff, M./Lauder, G. (eds) Nature’s purposes. Bradford Books/MIT Press: Cambridge MA, pp. 197–240. (Originally published in 1977.)
Nagel, T. (1998) Conceiving the impossible and the mind-body problem. Philosophy 73: 337–352.
Nagy, L. (1998) Changing patterns of gene regulation in the evolution of arthropod morphology. American Zoologist 38: 818–828.
Newman, S. A. (1994) Generic physical mechanisms of tissue morphogenesis: A common basis for development and evolution. Journal of Evolutionary Biology 7: 467–488.
Nicolis, G. (1995) Introduction to nonlinear science. Cambridge University Press: Cambridge.
Nicolis, J. S./Tsuda, I. (1999) Mathematical description of brain dynamics in perception and action. Journal of Consciousness Studies 6: 215–228.
Nijhout, H. F. (1990) Metaphors and the role of genes in development. BioEssays 12: 441–446.
Nissen, L. (1997) Teleological language in the life sciences. Rowman & Littlefield: Lanham MD.
Noble, D. (1998) Reduction and integration in understanding the heart. In: Bock, G. R./Goode, J. A. (eds) The limits of reductionism in biology. John Wiley & Sons: Chichester UK, pp. 56–72.
Norris, V./Raine, D. J. (1998) A fission-fusion origin for life. Origins of Life and Evolution of the Biosphere 28: 523–537.
Norris, V./Madsen, M. S./Freestone, P. (1996) Elements of a unifying theory of biology. Acta Biotheoretica 44: 209–218.
Novikoff, A. B. (1945) The concept of integrative levels and biology. Science 101: 209–215.
Omnès, R. (1999) Quantum philosophy. Princeton University Press: Princeton. (Originally published in French in 1994.)
Parrish, J. K./Edelstein-Keshet, L. (1999) Complexity, pattern, and evolutionary trade-off in animal aggregation. Science 284: 99–101.
Pattee, H. H. (1986) Universal principles of measurement and language functions in evolving systems. In: Casti, J. L./Karlqvist, A. (eds) Complexity, language, and life. Springer: Berlin, pp. 268–281.
Pattee, H. H. (1993) The limitations of formal models of measurement, control, and cognition. Applied Mathematics and Computation 56: 111–130.
Petrosky, T. Y./Prigogine, I. (1990) Laws and events: The dynamical basis of selforganization. Canadian Journal of Physics 68: 670—682.
Pettersson, M. (1996) Complexity and evolution. Cambridge University Press: Cambridge.
Pfeifer, R./Scheier, C. (1999) Understanding intelligence. Bradford Books/MIT Press: Cambridge MA.
Poland, J. (1994) Physicalism. Clarendon Press/Oxford University Press: Oxford.
Polanyi, M. (1969) Life’s irreducible structure. In: Polanyi, M. Knowing and being. University of Chicago Press: Chicago, pp. 225–239. (Originally published in 1968.)
Port, R. F./van Gelder, T. (eds) (1995) Mind as motion. Bradford Books/MIT Press: Cambridge MA.
Price, C. (1995) Functional explanations and natural norms. Ratio 8 (n.s.): 143–160.
Prigogine, I. (1997) The end of certainty. Free Press: New York. (Originally published in French in 1996.)
Primas, H. (1991) Reductionism: Palaver without precedent. In: Agazzi, E. (ed) The problem of reductionism in science. Kluwer Academic: Dordrecht, Holland, pp. 161–172.
Prusinkiewicz, P./Lindenmayer, A. (1990) The algorithmic beauty of plants. Springer: New York.

Raff, R. A. (1996) The shape of life. University of Chicago Press: Chicago.

Redhead, M. (1995) From physics to metaphysics. Cambridge University Press: Cambridge.

Reid, R.G.B. (1985) Evolutionary theory: The unfinished synthesis. Cornell University Press: Ithaca NY.

Rensberger, B. (1996) Life itself. Oxford University Press: New York.

Resnik, D. (1994) The rebirth of rational morphology: Process structuralism's philosophy of biology. Acta Biotheoretica 42: 1–14.

Ripoll, C./Guespin-Michel, J./Norris, V./Thellier, M. (1998) Defining integrative biology. Complexity 4(2): 19–20.

Rose, M. R./Lauder, G. V. (eds) (1996) Adaptation. Academic Press: San Diego CA.

Rosen, J. (1995) Symmetry in science. Springer: New York.

Rosen, R. (1991) Life itself. Columbia University Press: New York.

Rosen, R. (1993) Bionics revisited. In: Haken, H./Karlqvist, A./Svedin, U. (eds) The machine as metaphor and tool. Springer: Berlin, pp. 87–100.

Rosen, R. (1996a) The Schrödinger question: What is life? fifty years later. In: Buckley, P./Peat, F.D (eds) Glimpsing reality. University of Toronto Press: Toronto, pp. 168–190.

Rosen, R. (1996b) Personal communication.

Rosenberg, A. (1994) Instrumental biology, or the disunity of science. University of Chicago Press: Chicago.

Rovelli, C. (1997) Halfway through the woods: Contemporary research on space and time. In: Earman, J./Norton, J. D. (eds) The cosmos of science (Pittsburgh-Konstanz series in the philosophy and history of science). University of Pittsburgh Press: Pittsburgh PA, pp. 180–223.

Rueger, A./Sharp, W. D. (1998) Metaphysical presuppositions of scientific practice: 'Atomism' vs. 'wholism'. Canadian Journal of Philosophy 28: 1–20.

Russell, E. S. (1946) The directiveness of organic activities. Cambridge University Press: Cambridge.

Russell, E. S. (1982) Form and function. University of Chicago Press: Chicago. (Originally published in 1916.)

Salthe, S. N. (1985) Evolving hierarchical systems. Columbia University Press: New York.

Salthe, S. N. (1993) Development and evolution. Bradford Books/MIT Press: Cambridge MA.

Sanderson, M. J./Hufford, L. (eds) (1996) Homoplasy. Academic Press: San Diego CA.

Santayana, G. (1955) Scepticism and animal faith. Dover: New York. (Originally published in 1923.)

Sarkar, S. (1996) Biological information: A skeptical look at some central dogmas of molecular biology. In: Sarkar, S. (ed) The philosophy and history of molecular biology. Kluwer Academic: Dordrecht, Holland, pp. 187–231.

Savageau, M. A. (1996) A kinetic formalism for integrative molecular biology: Manifestation in biochemical systems theory and use in elucidating design principles for gene circuits. In: Collado-Vides, J./Magasanik, B./Smith, T. F. (eds) Integrative approaches to molecular biology. MIT Press: Cambridge MA, pp. 115–146.

Schmidt-Kittler, N./Vogel, K. (eds) (1991) Constructional morphology and evolution. Springer: Berlin.

Schmidt-Nielsen, K. (1997) Animal physiology, 5th ed. Cambridge University Press: Cambridge.

Schneider, E. D./Kay, J. J. (1994) Life as a manifestation of the second law of thermodynamics. Mathematical and Computer Modelling 19:25–48.

Schopenhauer, A. (1966) The world as will and representation, vol. 2. Dover: New York. (Originally published in German in 1844.)

Schweber, S. S. (1997) The metaphysics of science at the end of a heroic age. In: Cohen, R. S./Horne, M./Stachel, J. (eds) Experimental metaphysics. Kluwer Academic: Dordrecht, Holland, pp. 171–198.

Schweitzer, F. (ed) (1997) Self-organization of complex structures. Gordon and Breach: Amsterdam.

Scott, A. (1996) On quantum theories of the mind. Journal of Consciousness Studies 3: 484–491.

Searle, J. R. (1992) The rediscovery of the mind. Bradford Books/MIT Press: Cambridge MA.

Searle, J. R. (1995) The construction of social reality. Free Press: New York.

Segré, D./Lancet, D. (1999) A statistical chemistry approach to the origin of life. Chemtracts—Biochemistry and Molecular Biology 12: 382–397.

Shapiro, J. A. (1997) Genome organization, natural genetic engineering and adaptive mutation. Trends in Genetics 13: 98–104.

Shaw, R. E./Turvey, M. T. (1999) Ecological foundations of cognition: II. Degrees of freedom and conserved quantities in animal-environment systems. Journal of Consciousness Studies 6: 111–123.

Shimony, A. (1993) The non-existence of a principle of natural selection. In: Shimony, A. Search for a naturalistic worldview II: Natural science and metaphysics. Cambridge University Press: Cambridge, pp. 228—252. (Originally published in 1989.)

Skarda, C. A./Freeman, W. J. (1987) How brains make chaos in order to make sense of the world. Behavioral and Brain Sciences 10: 161–195.

Sklar, L. (1993) Physics and chance. Cambridge University Press: Cambridge.

Skulachev, V. P. (1992) The laws of cell energetics. European Journal of Biochemistry 208: 203–209.

Sober, E. (1999) The multiple realizability argument against reductionism. Philosophy of Science 66: 542–564.

Stadler, P. F. (1995) Towards a theory of landscapes. In: López-Peña, R./Capovilla, R./García-Pelayo, R./Waelbroeck, H./Zertuche, F. (eds) Complex systems and binary networks. Springer: Berlin, pp. 78–163.

Steele, E. J./Lindley, R. A./Blanden, R. V. (1998) Lamarck's signature. Helix/Perseus: Reading MA.

Stephan, A. (1992) Emergence—A systemic view on its historical facets. In: Beckermann, A./Flohr, H./Kim, J. (eds) Emergence or reduction? Walter de Gruyter: Berlin, pp. 25–48.

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

Stonier, T. (1990) Information and the internal structure of the universe. Springer: Berlin.

Stryer, L. (1989) Molecular design of life. W.H. Freeman & Co.: New York.

Swenson, R. (1998) Spontaneous order, evolution, and autocatakinetics: The nomological basis for the emergence of meaning. In: Van de Vijver, G./Salthe, S. N./Delpos, M. (eds) Evolutionary systems. Kluwer Academic: Dordrecht, Holland, pp. 155–180.

Swenson, R./Turvey, M. T. (1991) Thermodynamic reasons for perception-action cycles. Ecological Psychology 3: 317–348.

Tabin, C. J./Carroll, S. B./Panganiban, G. (1999) Out on a limb: Parallels in vertebrate and invertebrate limb patterning and the origin of appendages. American Zoologist 39: 650–663.

Thirring, W. (1995) Do the laws of nature evolve? In: Murphy, M. P./O'Neill, L. A. J. (eds) What is life? The next fifty years. Cambridge University Press: Cambridge, pp. 131–136. (Originally published in German in 1994.)

Thom, R. (1990a) Darwin cent ans après. In: Thom, R. Apologie du logos. Hachette: Paris, pp. 599–605. (Originally published in 1983.)

Thom, R. (1990b) Semiophysics. Addison-Wesley: Redwood City CA. (Originally published in French in 1988.)

Thomas, R. D. K./Reif, W. E. (1993) The skeleton space: A finite set of organic designs. Evolution 47: 341–360.

Tjian, R. (1995) Molecular machines that control genes. Scientific American 272(2): 54–61.

Travisano, M./Mongold, J. A./Bennett, A. F./Lenski, R. E. (1995) Experimental tests of the roles of adaptation, chance, and history in evolution. Science 267: 87–90.

Turvey, M. T./Shaw, R. E. (1999) Ecological foundations of cognition: I. Symmetry and specificity of animal-environment systems. Journal of Consciousness Studies 6: 95–110.

Van Cleve, J. (1990) Mind-dust or magic? Panpsychism versus emergence. In: Tomberlin, J. E. (ed) Philosophical perspectives, vol. 4. Ridgeview Publishing Co.: Atascadero CA, pp. 215–226.

van der Steen, W. J. (1994) New ways to look at fitness. History and Philosophy of the Life Sciences 16: 479–492.

van Gelder, T. (1995) What might cognition be, if not computation? Journal of Philosophy 91: 345–381.

van Gelder, T. (1998) The dynamical hypothesis in cognitive science. Behavioral and Brain Sciences 21: 615–665.

Vitiello, G. (1998) Structure and function: An open letter to Patricia Churchland. In: Hameroff, S. R./Kaszniak, A. W./Scott, A. C. (eds) Toward a science of consciousness II. Bradford Books/MIT Press: Cambridge MA, pp. 191–196.

Wachbroit, R. (1994) Normality as a biological concept. Philosophy of Science 61: 579–591.

Wächtershäuser, G. (1997) The origin of life and its methodological challenge. Journal of Theoretical Biology 187: 483–494.

Wainwright, P. C./Reilly, S. M. (eds) (1994) Ecological morphology. University of Chicago Press: Chicago.

Weber, B. H./Depew, D. J. (1996) Natural selection and self-organization: Dynamical models as clues to a new evolutionary synthesis. Biology and Philosophy 11: 33–65.

Webster, G./Goodwin, B. (1996) Form and transformation. Cambridge University Press: Cambridge.

Weibel, E. R. (2000) Symmorphosis. Harvard University Press: Cambridge MA.

Weibel, E. R./Taylor, C. R./Bolis, L. (eds) (1998) Principles of animal design. Cambridge University Press: Cambridge.

Welch, G. R. (1992) An analogical 'field' construct in cellular biophysics: History and present status. Progress in Biophysics and Molecular Biology 57: 71–128.

Weng, G./Bhalla, U. S./Iyengar, R. (1999) Complexity in biological signaling systems. Science 284: 92–96.

West, B. J. (1985) An essay on the importance of being nonlinear. Springer: Berlin.

West, B. J. (1990) The disproportionate response. In: Mickens, R. E. (ed) Mathematics and science. World Scientific: Singapore, pp. 258–290.

Wheeler, J. A. (1990) Information, physics, quantum: The search for links. In: Zurek, W. H. (ed) Complexity, entropy and the physics of information. Addison-Wesley: Redwood City CA, pp. 3–28.

Wheeler, M. (1997) Cognition's coming home: The reunion of life and mind. In: Husbands, P./Harvey, I. (eds) Fourth European Conference on Artificial Life. Bradford Books/MIT Press: Cambridge MA, pp. 10–19.

Wick, D. (1996) The infamous boundary. Copernicus/Springer: New York.

Wicken, J. S. (1987) Evolution, thermodynamics, and information. Oxford University Press: New York.

Williams, G. C. (1992) Natural selection. Oxford University Press: New York.

Williams, R. J. P./Fraústo da Silva, J. J. R. (1996) The natural selection of the chemical elements. Clarendon Press: Oxford.

Wilson, K. G. (1983) The renormalization group and critical phenomena. Reviews of Modern Physics 55: 583–600.

Wilson, M. (1993) Honorable intensions. In: Wagner, S. J./Warner, R. (eds) Naturalism. University of Notre Dame Press: Notre Dame IN, pp. 53–94.

Wimsatt, W. C. (1985) Some problems with the concept of 'feedback'. In: George, F./Johnson, L. (eds) Purposive behaviour and teleological explanations. Gordon and Breach: New York, pp. 165–181. (Originally published in 1971.)

Wimsatt, W. C. (1997) Aggregativity: Reductive heuristics for finding emergence. Philosophy of Science 64 (supp.): S372–S384.

Wolynes, P. G. (1991) Search and recognition: Spin glass engineering as an approach to protein structure prediction. In: Peliti, L. (ed) Biologically inspired physics. Plenum: New York, pp. 15–37.

Woodfield, A. (1976) Teleology. Cambridge University Press: Cambridge.

Wouters, A. (1995) Viability explanation. Biology and Philosophy 10: 435–457.

Wright, L. (1998) Functions. In: Allen, C./Bekoff, M./Lauder, G. (eds) Nature's purposes. Bradford Books/MIT Press: Cambridge MA, pp. 51–78. (Originally published in 1973.)

Yates, F. E. (1993) Self-organizing systems. In: Boyd, C. A. R./Noble, D. (eds) The logic of life. Oxford University Press: Oxford, pp. 189–218.

Yates, F. E. (1994) Order and complexity in dynamical systems: Homeodynamics as a generalized mechanics for biology. Mathematical and Computer Modelling 19: 49–74.

Yates, F. E. (1997) Microsemiosis. In: Posner, R./Robering, K./Sebeok, T. A. (eds) Semiotics, vol. 1. Walter de Gruyter: Berlin, pp. 457–464.

Yates, F. E./Kugler, P. N. (1984) Signs, singularities and significance: A physical model for semiotics. Semiotica 52: 49–77.

Yockey, H. P. (1992) Information theory and molecular biology. Cambridge University Press: Cambridge.

Zak, M./Zbilut, J. P./Meyers, R. E. (1997) From instability to intelligence. Springer: Berlin.

Winfried Behr

Homo Ridens

A Speculation on the Origin and Biological Function of Humor

Abstract

Laughing at (derision, ridicule), laughing with (genial laughter) and smiling are functionally related behavioral phenomena of partly differing evolutionary origin.
Derision is the earliest phenomenon. It started as a communication of censure and disapproval of aberrant behavior directed from the group to a group member (mobbing call). Exclusion of the censured member increased the efficiency of group efforts and thus the individual fitness of the remaining group members. Derision is aimed at the comical. Criteria of the comical are odd (stupid) behavior or appearance, bad luck, and imitation.
Genial laughter is associated with verbal wit. Wit creates comical effects by verbal means. Having only imaginary, not real victims it has largely lost the overt aggression of derision. Verbal wit requires understanding in order to be enjoyed. Understanding of verbal wit is a criterion for admission to groups of individuals of a similar sense of humor. It thus promotes intellectual homogeneity in groups. Individual fitness gains partly by higher efficiency of group activities and partly by assortative mating within intellectually more homogeneous groups.
Smiling, as a reaction to wit, evolved as a means of communication between potential sexual partners displaying intellectual capabilities. It is a means of furthering intellectually assortative mating.

Key words

Wit, wittiness, comical, laughter, intelligence, smiling, sexual dimorphism, assortative mating.

1. Introduction

The phenomenon of the complex of laughter and wittiness, as well as wit, as a determinant of laughter, is a problem that has been discussed by philosophers since antiquity, but never satisfactorily explained (RAPP 1949; MORREAL 1987). In fact, an explanation of the phenomenon has never really been attempted as its origin and existence is closely related to the origin and existence of man and this problem—as late as the nineteenth century—could only be answered metaphysically. Philosophers, therefore, have limited themselves to defining the phenomenon of wit and comedy, that is, to describing in general terms the situation which causes in humans this particularly pleasant emotion, which is accompanied by laughter. The better known of these definitions are the so-called traditional theories of humor and laughter. These theories all cover certain aspects of the phenomenon satisfactorily, but usually fail in others. This is discussed by RAPP (1949) who quotes ARISTOTLE as one of the first to write about humor and laughter: "The laugh provoking may be defined as that error or deformity which is not painful or destructive" and "most witty remarks have their source in metaphor, and in misleading your listener… The same result is produced by jests that hinge on a change in letter (i.e., puns); for they are deceptive…". RAPP critizises that "the two statements are quite different; and more than that, they take no cognizance of each other" and, furthermore, they do not account for ridicule (laughing at someone): "Any good definition which covers wit does not fairly cover ridicule; and vice versa".

The same difficulty arises with KANT's well-known statement that "laughter is an affection arising from a strained expectation being suddenly reduced to nothing" (the relief from restraint theory). It is indeed a definition which is not suitable for many forms of wit and in addition cannot be applied to ridicule. RAPP continues with SCHOPENHAUER's incongruity theory: "The cause of laughter in every case is simply the sudden perception of the incongruity between a concept and the real objects which have been thought through in some relation, and laughter itself is just the expression of this incongruity". RAPP confirms this definition to be partially successful but reminds us that incongruity is a very broad term. "and at most incongruities we do not laugh".

The proposed hypotheses cover the phenomenon partially, at best, and this is also true of Hobbes' description of laughter as "a sudden glory arising from some conception of some eminency in ourselves, by comparison with the infirmity of others, or with our own formerly" (superiority theory). (All quotes from RAPP 1949, pp81–82).

The deficiency of the traditional theories on humor and laughter is stressed by KULKA (1990). To quote from the abstract of his article: "The incongruity theory is today the most popular mainly because its rivals (that is: relief from restraint and superiority theory) are considered discredited. It is considered to be particularly well suited to account for humorous laughter and amusement occasioned by jokes. [The] article is not concerned with the question of whether the concept of incongruity adequately covers all cases of laughter, for it is argued that the incongruity theory is inadequate even for those cases for which it is thought to be especially well suited. It is shown that the incongruity theory does not (and indeed cannot) account either for the pleasurable effect of jokes, or for aesthetic pleasure".

Thus, not only is there no explanation for wit, the comical and laughter, there is not even a commonly accepted definition of laugh-provoking (that is, wit and comedy), although from introspection their emotional effect is very well known.

Traditional theories of humor seem to suffer from an absence of useful concepts from which one could not only form workable definitions of wit and comedy, but also come closer to an explanation of the problem. In his well-known publication of 1900, BERGSON defines the comical as "mechanical inelasticity", or as behavior that is "mechanical or unresponsive to the existing surroundings". The purpose of laughter is, according to him, "social control". While based on keen observation and clearly correct, these conclusions do not lead to an understanding of the phenomenon because they are not related to any appropriate concept. And to be appropriate the concept must be based in biology.

If agreed the origin and existence of man is best explained as a problem of biology, then it stands to reason the phenomenon of humor and laughter, so inseparably connected to human nature, must also have its explanation based in biology.

This was correctly understood by RAPP (1949) who, therefore, proposed a phylogenetic theory of wit and humor. He assumes that the various forms of humor must have a common root: ridicule and wit "must all be related. For laughter is one". And "…there must have been a time when puns did not exist, there must have been a time when genial humor did not exist, there must have been a time when savage derision did not exist". And his question is also the one discussed in the present article: "How did these develop, one out of the other; or one out of something else?" (RAPP 1949, p83).

We may add here that ridicule and wit are certainly related biologically and that laughter is indeed "one". However, the usual assumption that the determinants of laughter too must be reducible to a uniform and universal definition is based on intuition rather than logic.

According to RAPP (1949), the origin of laughter was in the shout of thriumph of the winner of a duell ("thrashing laughter"), a shout repeated by the onlookers of the fight who thus sided with the winner. He claims that this scenario accounts for and includes the traditional theories of laughter: relief from restraint and feeling of superiority. The third—incongruity—was later included when the contest of brawn gradually developed into a contest of brain, consisting in riddle solving. This short and therefore necessarily deficient summary can naturally not do justice to RAPP's theory, which has the merit of treating wit and laughter for the first time as a problem of biology. However, the hypothesis has a basic flaw: while he calls his theory phylogenetic, he disregards the principles of ethology which are indispensible for the plausibility of phylogenetic assumptions. Which is the precursor behavior which by modification could have become "thrashing laughter"? Which adaptiveness lies in the "stentorian announcement of victory"? What was the driving force behind the assumed long development from triumphant shout to genial laughter? But ethology was far from being common knowledge in 1949 when RAPP presented his theory and, therefore, these questions could not be addressed. This reminds us of the essential role played by concepts in biology (as in any science, but more so in biology).

WEISFELD published in 1993 a comprehensive review of the traditional theories of humor and laughter. He points out their inadequacies and proposes his own tentative explanation: "Humor evolved to induce the subject to seek out informative social stimulation and to reward others for providing such stimulation (by laughter)" (1993, p162). According to this hypothesis, laughter is a form of applause for the jester to make him understand that more stimulation—humor—is desired. It is obvious that this explanation cannot account for the laughter of ridicule which is clearly not an expression of appreciation.

While philosphers have concentrated on the intellectual quality of the determinant of laughter, namely comedy and wit, biologists have paid attention to the physical expression of laughter. They have traced the facial features of laughter to precursory behavior of certain mammals (VAN HOOFF 1972). Amazingly, they have almost completely neglected the experience of pleasure, the emotional reward connected to wit and comedy, even though the strength of this feeling is remarkable and clearly waiting for an ethological explanation.

2. The Definition of the Humorous

Philosophers have failed to agree on a universal definition of comedy and wit. The variability of causes of the humorous makes it apparently impossible to determine a uniform pattern therein. We may thus have to reconcile ourselves to the possibility that more than one definition exists; that there are several forms of the humorous, independent from each other and yet equally valid. In the framework of an ethological hypothesis a functional unity may appear making the searched for logical uniformity dispensible.

Comedy and wit are the determinants of a pleasurable emotion which,in turn, causes us to laugh or smile. Laughter is indeed "one". However the determinants of laughter are not. And not only are there various, independent categories into which occasions for laughter can be grouped, there seem to be, in addition, two distinct forms of the phenomenon. On the one hand, a more primitive, predominantly non-verbal form, better described as comical (or comedy); and, on the other, a verbal form in a stricter sense of wit. Wit is the verbal construction of comical effects. While it uses largely the same elements as the comical, it differs from the comical in certain aspects to be discussed below in more detail.

The comical that is visual- its most prevalent form—is readily understandable. One observes modes of action, behaviors, processes, shapes or images and is compelled to laugh if certain conditions are met—the most important is an association with humans. One laughs about human behavior, about processes that involve humans, and about forms and images which are human-based. There are some exceptions: animals, if their behavior is suggestive of human behavior; and objects or preposterous constructions, be they toys or other creative devices, if capable of animation and seen in action. A comical effect is impossible to achieve, however, with inanimate objects or plants. To quote BERGSON (1900; cit. from DIETZSCH 1993, p34), "There is nothing comical which is not truly human. A landscape may be beautiful, pleasant, imposing, boring or ugly but never comical. One laughs about an animal, but only because one has discovered in it a human feature or expression".

Human involvement is the necessary precondition of the comical, yet further criteria have to be met: behavior betraying stupidity; or observed behavior, images, figures or courses of events in conflict with our expectations and thus deviating from the norm. The comical is further found in events showing the bad luck or the embarrassment of others, and, finally, in imitation or aping. It should be noted that these criteria for the comical—stupidity, deviation from the norm, malice and imitation—have different structures, and, except for certain similarities between stupidity and deviation from the norm, nothing in common. Apparently an emotional reward accompanied by laughter has evolved as a uniform reaction to the clearly variant causes:

- stupidity
- deviation from established norms, grotesqueness
- malice: bad luck or embarassment of others
- imitation or aping.

The difference among the causes is so pronounced that one has to exclude the possibility of a uniform definition of the comical (or wit). On the other hand, there is remarkable uniformity of reaction—a never changing sense of the comical, accompanied by laughter—which provokes one's curiosity and suggests the possibility that the behavioral complex of the comical and resulting laughter must have a uniform biological origin.

While stupidity, malice and imitation are clearly defined, the remaining criterion—deviation from the norm—is somewhat more diffuse. It encompasses deviations from conventional images (real or imagined), and from reasonable expectations. Expressive words such as absurd, grotesque, eccentric, and preposterous are used for the characterization of such deviation.

It is conceivable that the determinant "breach of a norm" predated the determinant "stupidity" in the phylogeny of the behavioral trait, as the latter encompasses a value judgement of considerable degree of abstraction, which at some time should have been beyond the mental capabilities of our early forebears. Stupid behavior often appears as deviant from an established routine and, therefore, the breach of a norm may originally have served as a functional mark of stupidity before this abtract term existed in the minds of our forebears.

3. Criteria of the Comical, the Comical as Opposed to Wit

The visual predominates among comical motives. We observe and feel compelled to laugh. We observe unreasonable or irrational behavior and we laugh. We observe one or more persons caught in an embarrassing situation and we laugh. We see figures or pictures deviant from our conventional expectations and we laugh about their "grotesqueness"—a word coined at the time of the Renaissance when Roman paintings, discovered in an Italian cave ("grotta"), were considered deviant from the standards of good painting, and thus "grotesque". For the comical derived from imitation, the exact opposite is true, as the more imitation resembles known behavior, the more comical it is.

The prevalence of the visual in the comical is probably due to the fact that the phenomenon originated at a stage of human evolution during which language was still rudimentary. Yet, in spite of this prevalence, the visual is not essential for the comical as it also may have auditory causes. Thus, linguistic errors are in this sense merely comical, despite being verbal, and not witty. One listens and often feels compelled to laugh or smile at the unintentional linguistic creations of a foreign speaker. The reluctance to express oneself in an imperfectly mastered foreign language has its roots in the fear of appearing ridiculous. The comedy of the ridiculous lies here exclusively in a deviation from the linguistic norm firmly established in the minds of native speakers.

Also someone making a stupid remark will be considered a comical figure to be laughed at. (But someone telling a joke is witty and his audience will laugh with him).

The comical of imitation too can be visual as well as auditory. Not only is the mimicking of facial expressions and behavior comical, but the comical is also found in the imitation of manners of speech and accents.

There are people who laugh at individuals with physical peculiarities, because they are exceptionally tall, short, heavy, skinny or have other characteristics of appearance at odds with the norm. Laughing at persons handicapped in one way or another, or at foreigners with linguistic peculiarities is considered rude and pardonable only if done by children. "Laughing marks a fool" is the saying. Yet not only children and fools consider grotesque appearances to be comical. Everyone indulges in guilt-free laughter at the humor in the cartoons of grotesque exaggeration because no real person is hurt by such amusement.

Laughing about the predicament of others is also unacceptable unless the observed embarrassing situation is one of theatrics and not of reality. In a play one is no longer aware that laughing at such things as people caught in the struggle of having to swallow the spilled beans, pompous dignitaries bumping into each other after the offering of polite entreaties to move first, or speakers hit in the face by custard pie, etc., is really laughter about the misfortune of others. Laughing at the antics of a clown, may also be laughter about a person's bad luck, yet does not cause feelings of guilt as one knows the predicaments into which the clown perpetually falls are theatrical with no real consequences.

We defined the comical as a more primitive, predominantly visual or physical, and sometimes auditory form of humor. It is more primitive than verbal wit because it is probably phylogenetically older, and its understanding requires less intellect, and lastly because laughter about the comical is often of an ethically unacceptable nature. Nevertheless, we admire great clowns as great artists who cater to our doubtful needs for entertainment. The admiration is justified, as the art of a clown civilizes those instincts of ours which have basically an uncivil origin.

4. A Possible Mechanism for the Evolution of Laughter

A possible mechanism for the evolution of laughter as a reaction to the comical may be found in the probable scenario in which the above discussed criteria of the comical occurred along with a precursor of laughter.

And this may have been the scenario: a group observed the unreasonable, inappropriate behavior of a member and warned him of this by laughter. Laughing in such a situation means something like: Do not do it! Stop it! Leave it! Go away! In other words laughter at stupidity. If this inappropriate behavior resulted in the unfortunate consequences foreseen by the group, the laughter turned into that of gloating (malice)—Justly so! You deserved it! In a similar fashion imitators may have become objects of the laughter of derision. The group observes an individual in the action of mimicking the behavior of another without understanding the internal concept of what he is doing and punishes him with laughter.

This hypothesis assumes the origin of laughter is rather an "action whose function is to correct or to repel deviant or non-conforming individuals" (EIBL-EIBESFELDT 1986) than VAN HOOFF's (1972) relaxed open mouth display. The original form of laughter would thus have been mobbing calls.

There is, in fact, a form of laughter which still does not require humor or wit as its cause, and it expresses nothing but: Stop it! Leave me alone!. It is the laughter in reaction to physical molestation, i.e., tickling. A similar type of laughter—albeit produced by a respiratory technique different from ours—has also been observed in chimpanzees (PROVINE 1996). The circumstances show that laughter in chimpanzees is also a demand for cessation and being left alone. Actually, one can expect living beings, in general, to often have a need to express this. It is, therefore, not surprising that a corresponding behavioral expression evolved long before language. Mammals only remotely related to us show the so-called "hissing, bared-teeth display", which is considered by ethologists to be a forerunner of laughing (VAN HOOFF 1972). Such a display could only be directed towards conspecifics, as only they could be expected to understand. Thus, the precursor of laughing already excludes inanimate or vegetative objects as subjects of laughter.

The proposed scenario accounts for the precondition of human involvement and for the independent determinants of laughter (stupidity, malice and imitation), but only partly for the aspect of deviation from the norm. Aberrant behavior may be interpreted as stupid, yet there remains the laughter at deviant physical appearance. One is reluctant to accuse our forebears of the cruelty of laughing at individuals with physical handicaps and, yet, such practice must have existed as it exists even today, albeit considered inappropriate and a breach of ethical norms.

5. Individual Fitness Gains From Higher Group Efficiency

Laughter about the comical is associated with a strong emotional reward. One loves to laugh, one seeks opportunities to laugh and one is even willing to pay for being made to laugh. As we may be sure that nature—or rather evolution—does not do anything unreasonable, we have to accept that the complex of the comical and its emotional reward of laughter confers some survival value, that it is adaptive. If we consider, in addition, how strong the emotional reward of laughter is—only sexual activities are likely to be more rewarding --, we must assume the survival value or the fitness gain conferred by this complex to be very important. Thus, we must assume that it is adaptive to laugh at an outsider because his physical peculiarities or aberrant behavior are in conflict with group expectations; that it is adaptive to laugh at him because he relies on imitation when the demands of the situation are beyond his limited mental capabilities to react otherwise; and, finally, that it is adaptive to exhibit pleasure over the bad luck of an outsider caught in an unfortunate predicament. The laughed at, less capable member is excluded from the group and this exclusion could be the key to the adaptiveness of the process. The exclusion of a less capable member makes group endeavors apparently more efficient and this increase in efficiency is reflected in a gain of the individual fitness of the remaining group members (increased individual fitness, that is a higher number of surviving offspring). Only thus, by an increase of individual fitness, could the complex of the comical and laughter have become fixed in the human genome. The fitness gained on account of higher group efficiency must have been a very significant one as the exclusion of a group member is, per se, something detrimental to average individual fitness.

The increased fitness of the remaining group members was acquired at the expense of the excluded. At an early stage of human development survival outside the group was difficult, if not impossible. In addition, there were greater obstacles for the initiation of sexual contact. Thus a lower survival rate and lower mating success diminished the number of offspring of the laughed at outsider and thus his biological fitness. While at an earlier stage of evolution, the reduced chances for survival outside the group were probably more crucial, the difficulty of establishing sexual contacts became more important as a determinant in a later period when verbal wit had come into being. A long-term, unintended benefit of the painfull practice of derision must have been an increase in the over-all intelligence of the next generation, as the criteria of the comical give a certain reason to assume a below average intelligence was often the cause of exclusion from the group.

That not only laughing at aberrant behavior, but also laughing at grotesque appearance could have been adaptive, is, at first glance, not quite understandable. But it is possible the exclusion of individuals with physical disparities had individual or familial advantages less clearly understandable than punishment of aberrant behavior. It may be that uncommon physical characteristics were a sign of lower physical efficiency making their rejection as efficiency enhancing as the rejection of aberrant behavior. And it is also not impossible that uncommon physical characteristics (e.g., receding forehead) signified intellectual defects. This could have, for example, lowered the likelihood of interbreeding between anatomically modern humans (our forebears) and Neanderthals, which, indeed, did not take place, as was learned from the analysis of mitochondrial DNA from Neanderthal bones (KRINGS 1997), even though Neanderthals lived for a long time in close contact with anatomically modern humans and may have even taken over some cultural practices from them, such as tool fabrication and the use of red ochre (MELLARS 1998).

6. Emotional and Behavioral Peculiarities Become Understandable

The hypothesis for the evolution of the complex of the comical and laughter outlined above is supported by the fact that certain, otherwise unexplainable, phenomena fit into the proposed concept, above all, our nearly panicky fear of being laughed at, which becomes understandable through its assumed evolutionary history. To be laughed at is accompanied by the feared threat of exclusion with all the fateful consequences that it had in the early stages of human development. It diminished the chance to beget off-spring, and it continues to do so. This can explain why our fear of being laughed at is so strong, even disproportionally strong when one keeps in mind that there is no danger of any physical attack. Even smiling, normally an encouraging gesture, can be painful, if it can be interpreted as derision, as shown by the following examples.

When in Thomas MANN's "The Magic Mountain", Dr. Krokowski, a "very brilliant entity", "practicing soul analysis" (that is, a psychoanalyst) meets Hans Castorp and his cousin, he "greeted the new inmate with a jovial and robust heartiness, as who should say that with him all formality was superfluous, and only jocund mutual confidence in place" and shakes hands with "cordial smiles". The message is misunderstood and Hans Castorp is irritated as he sees "in the smiling … signs of benevolent mockery". Certainly a false interpretation and we may wonder why MANN, of all people, had a psychoanalyst commit such a blunder in non-verbal communication.

Then in STENDHAL's "Red and Black" when Monsieur de Rênal on his way to church one Sunday notices in the distance the sawmill owner, Sorel, smiling at him, the incidence creates a "fateful day". De Rênal now begins to suspect this is because he probably paid too high a price for a piece of land he purchased from Sorel. A social superior, de Rênal can afford the financial loss, but he cannot tolerate the suspicion of being laughed at.

7. Wit, the Creation of the Comical by Verbal Means

7.1 Wit, as opposed to the comical

For the sake of easier presentation of our argument, the discussion was limited so far to what we chose to call the comical (i.e., the humorous, that is visual, physical, directly understandable). This approach is also warranted by the assumption that the comical and derision predated the evolution of verbal wit and genial laughter to be discussed in the following.

The emotional reward connected to derision must have at a certain time initiated attempts to verbally create situations to laugh at without having to wait for chance events as occasion of ridicule. Provided that the comical and derision already existed, verbal wit was to come into existence as an almost inevitable evolutionary consequence of the progress of speech.

Verbal wit is a much more common phenomenon than the comical, that is visual or physical. *Homo ridens* is first of all *homo loquens*. The use of speech has, by its capability to create comical effects by verbal means, not only dramatically increased the scope of humor, but also its importance.

Wit is a verbal construction of the comical or to use the term of PROVINE (1993) a "structured attempt at humor". It uses more or less the same elements as the comical, but in contrast to the comical, poses certain challenges to comprehension which must be overcome by intellectual effort. While reaction to the comical rarely requires high intelligence, a joke has to be understood if it is to be enjoyed. This requirement of intellectual understanding is an added dimension practically absent from the comical. The simplest form of a joke verbally presents a behavior or opinion which is perceived as silly or absurd and reacted to with laughter. The build-up to a joke sets a scenario familiar to an audience, which can anticipate its consequences. Then the dimwitted figure is introduced, made to say or do something absurd or stupid showing he is—in contrast to the audience—not up to the situation, and not aware of what response could save him from unpleasant consequences. The audience must both understand the context and the appropriate response, and, thus, the absurdity of the response of the imaginary dull person—an intellectual feat that must be performed instantly and without effort in order to earn the emotional reward that comes with the understanding of wit. The object of derision is an imaginary victim, who cannot suffer from his role, cannot be excluded from the group and whose success of reproduction cannot be affected. Thus the absence of overt aggressiveness is—next to the requirement of understanding—the second aspect by which verbal wit differs from the comical. The joke or verbal wit in general, starts to assume other functions, such as strengthening group cohesion, an important, but certainly not its only function.

Jokes must be understood instantly and without effort in order to have the biological function pointed out below. This explains the "suddenness" often discussed as an essential element of humor (compare KANT's statement on laughter). But it is essential only for the functioning of a joke: the understanding must be sudden, followed by the evidence of this sudden understanding: laughter. Sudden events or surprising communications as such are, in most cases, not humorous and often rather unpleasant.

7.2 Forms of wit

Wit may assume various forms. The most common one is the humor of self-irony which consists of the speaker assuming the role of an object of laughter or derision.

The speaker presents himself as a dullard with inane opinions, as one ready to act the fool or play the clown, as a font of false analogies (such as puns), or of the gullible. The audience understands this as playful and not real, and is grateful to the speaker for the fun accorded by his wit, not to mention the sense of superiority granted by the feigned humbleness of the performer. This type of humor still has a real, not only imaginary, person to be laughed at—the performer himself—, but in no way does he have to fear exclusion from the group. The aggression of derision has disappeared, laughing at has become mere laughing with. The roles that are witty because they are pretended would be merely comical, if anything, were they real.

The former Canadian president, TRUDEAU, once heard the cries of a baby during a campaign speech. He interrupted his remarks and asked: "Has anyone seen my wife?" TRUDEAU, whose wife only shortly before had left him, insinuated with this question the naive possibility of his family's problems being the cause of the baby's crying. A leader making himself the object of irony is, however, quite the exception, as it is not a recommendable strategy for attaining and defending a leadership position. But the politician instinctively recognized that humor was the best reaction to this situation and that there would be less damage to his reputation if he beat the audience to connecting the baby's crying to his own unfortunate family ordeals. He demonstrated not only humor, but also ready wit; something which strengthens leadership positions. Great leaders are more inclined towards sarcasm. This is wit directed not at oneself, but rather the weaknesses of others. When CHURCHILL became First Lord of the Admiralty, he tended to disregard established customs. When asked by some admirals if he was going to question the traditions of the navy, he replied, "The traditions of the navy are mutiny, sodomy and rum", or according to another version, "sodomy, rum and the lash". His irreverent response, thereby, exposed the admirals to ridicule. Sarcasm has real, not only imaginary victims, yet its purpose is not exclusion from the group. The victims are either not members of the group, or, if they are, the purpose is struggle for dominance in, rather than exclusion from the group.

Sarcasm, in contrast to humor, is clearly aggressive. Occasionally, a humorous remark may seemingly have an aggressive aspect, as in the following scenario: an entertainer instigates laughter in his audience with questionable jokes, and then compliments them on their receptive reaction, which in turn evokes more laughter. One could assume aggression is involved here, since, after all, the entertainer accuses his audience of having bad taste. But, in truth, the wittiness of his remark lies not in the aggressive content but rather in its inappropriateness, stupidity and imprudence, as the entertainer is dependent on the goodwill of his audience. The audience, of course, understands this breach of etiquette or lack of good sense is just part of the desire to be entertaining, and goes along with it.

A breach of the norm can also be witty when stupidity is in no way involved. The Yiddish word "Chutzpa", defined as "gall, brazen nerve, effrontery" (ROSTEN 1968), is the calm rejection of commonly accepted moral norms, an imposing of one's own questionable position on others. For example, a beggar invited into the kitchen by a housewife helps himself to cake from the table. “There is also bread”, says the housewife. “Yes”, says the beggar, “but cake tastes better”. “But cake is more expensive”, replies the housewife. “And it’s worth it”, says the beggar.

The beggar cannot be accused of stupidity. His remarks just represent the impudent dismissal of common etiquette. While this may not be amusing to the housewife of this tale, it would be for the teller and his audience. It is wit, a verbal construction of the comical, and its humor is found in the breaching of the norm.

Not only defiance of standard behavior, but the calm rejection of customary views, conventional opinions, polite phrases can provoke laughter, as the following anecdote about the writer Nikolai GOGOL illustrates. Having moved to St. Petersburg to further his career, he called on PUSHKIN his first morning there. PUSHKIN’s servant who answered the door explained his master was still asleep. “He must have worked late last night”, said GOGOL. “Your Honour means he was playing cards”, replied the servant. It is difficult to decide which of the proposed elements of the comical is being used here. Is it that GOGOL’s polite remark is to be considered stupid? Or is the rejection of politeness by the servant stupid? The example shows to what degree the function of verbal wit has outgrown the simplistic roots of the comical to be laughed at. Essential for the witty effect is the rapid understanding of an indirect message, in this case the interpretation of GOGOL’s politeness as obsequiousness. Whether this obsequiousness is really a stupidity is of little relevance and so is whether the servant consciously wanted to give the visitor a lesson.

A grotesque image can also be verbally brought to mind by a joke. For example: Why are there always ditches along roads in certain flat coastal regions? So the inhabitants can move their arms while they walk. The humor of this joke, if there is any, lies in the imagining of grotesque human figures. The joke has something embarrassing and ethically unacceptable about it, as is often the case in physical comedy. The laugh provoking that lies in the mental generation of a grotesque image has also to some degree varied its original form. While in most cases the image conjured up verbally is one of grotesqueness it is not necessarily so in all cases. Often the image created mentally is not grotesque any more, but just unexpected or merely uncommon. The emphasis of the process has shifted away from the quality of the image to the mental creation of the image. The more complicated the process of creation, the more significant the intellectual achievement, the higher the amusement.

The comical found in stupidity or deviation from the norm is extensively used, and with great variety in verbal jokes; but apparently not the comical derived from malice and imitation. The humor in an embarrassing situation can only be experienced by seeing it and not just by hearing a description. Whether the humorous aspect of malice can never be produced verbally, is, however, up to the judgement of the reader, which he may exercise on the following little story: “Our coworker did not come to work for a week because of psychological problems. Her wedding had been announced and, with the assistance of the bridegroom, wedding gown and white shoes had been purchased. But then a few days before the event, the bridegroom panicked and made his escape”. Does anyone laugh? (The story is authentic. It was noticed that people who did not personally know the unfortunate young lady found something funny in it, while those who knew her did not.) It is, however, apparently impossible to duplicate the comical effect of mimicry indirectly in a verbal joke. Thus, verbal wit does not use the element of imitation and only rarely the element of malice. But by varying the remaining two elements —stupidity and deviation from the norm—it extends its scope tremendously and often is of great subtlety. This makes the definition of wit such a difficult task. What was originally an encouragement to chastise stupid behavior has developed into emotional reward for understanding consequences of stupid behavior and, finally, reward for mere understanding. The emphasis has shifted: Ridicule aims at the outsider who is to be punished for stupidity; in contrast the purpose of verbal wit is to reward insiders for understanding. Enjoying a completely different and surprising view of a situation or problem (as in an unexpected retort) has so little in common with the derision of stupidity that it amazes one that the two events should both provoke the same reaction of laughter. It should be noted, however, that an uncommon, unconventional approach to a problem is very often a means to a surprisingly clever solution of the problem. If wit is some kind of a natural intelligence test as argued below, evolution had a good reason to include the understanding of a surprisingly original view among the determinants of wittiness.

This is true in particular of a humorous situation that is so far removed from the originally laughed at stupidity that it almost warrants a separate category of the laugh-provoking. It is the amusement experienced when a person unintentionally betrays his secret thoughts. The person doing so unintentionally may be considered stupid (and someone doing so intentionally, a good entertainer) but the emphasis has totally shifted from the intelligence of the person betraying himself to the intelligence of the person understanding the true meaning of what is said, although not explicitly. This stresses the importance—the adaptiveness—of the understanding of hidden or indirect messages in social interactions. We describe the unintentional slip of tongue as wittiness, although it does, in most cases, not meet the conditions of our definition of verbal wit, that is, indirect creation of a comical situation: slips of tongue are rarely elements of jokes, but are rather experienced and enjoyed directly whenever they occur. And likewise they do not meet the conditions of our definition of the comical because a certain degree of intelligence is indispensible for the understanding of the true meaning of a slip of tongue.

The betrayal of secret thoughts is as an element of a joke not common, yet it does occur, as in the following: a boy and a girl meet on a trail in the fields and walk together. The boy leads a goat on a rope and carries a stick, a live chicken and a kettle. When they come to the entrance of a gorge the girl says, "Here I cannot follow you any longer because in the gorge you will try to embrace and kiss me". Says the boy, "How should I do this the way I am loaded?" The girl replies, "You will ask me to hold the goat, you will push the stick into the ground and you will put the chicken under the kettle". "Allah blesseth your wisdom", says the boy and they walk on (it is the Arabs to whom we owe this charming little story).

7.3 Understanding of wit is an intellectual accomplishment

Understanding is essential for the enjoyment of wit. The degree of complexity of any witticism varies greatly with the consequence that not everyone understands every witticism. Next to simple jokes there are witticisms whose understanding is a considerable intellectual accomplishment. Individually such an achievement is not noticed as such, as the precondition for enjoying wit is that it be understood instantly and effortlessly. That this feat of immediate understanding is essential for the enjoyment of wit is demonstrated by the fact that if it is followed by an explanation its effect is blunted to the point of irritation. Such an explanation either insults the intelligence of the audience or implies the attempt at wit was too feeble to stand on its own. In either case, its appreciation by the audience is undermined and the jokester or entertainer is rejected.

The phenomenon of involuntary humor also shows the essentiality of understanding wit. While the speaker is not aware or does not understand the ramifications of what he is saying, the audience does and laughs at him. An example is found in the following text quoted from a company publication attempting to sell a miracle cure to young mothers of a country with a high birth rate: "In our neighbour country XXX the problem was solved by administering XYZ tablets to nursing mothers. Tests have shown that the tablets enhance milk secretion. The results also showed that the consumption increased the breast volume considerably. The rate of increase was equal in both breasts."

The author of the text discusses earnestly the results of some supportive investigation of questionable scientific merit, but readers aware of the availability of scientists today willing to confirm the results of any study, if paid enough, know the absurdity of this. He assumes all measurements of the investigation were—if at all—executed with an unscientific bias to obtain the desired result. In this context the claim of scientific thoroughness ("both breasts measured, both of equal size") is a complete absurdity, as is the thought of physical asymmetry being an unwanted result of the miracle cure.

The more subtle the witticism the more significant the intellectual achievement of understanding, as in the following examples:

In the former Soviet block a joke was told of a Russian who during STALIN's rule was punished by ten years of forced labour camp in Siberia. He served his term, found work afterwards, succeeded in a career and became a factory director ten years later. Then wishing to see his old mother again he invites her to join him. Arriving by Trans-Siberian rail, she immediately recognizes her son awaiting her on the platform and embraces him. Astonished the man says, "Mother, how could you recognize me right off after twenty years?" "Your jacket", she says touching his lapel, "Your jacket".

The wit of this little story may be obvious, but it gains a degree of sophistication by assuming a background knowledge, not explicitly mentioned, of the absurd contrast of the promises of future happiness, tiringly and tirelessly made by the communists and the grotesque reality as experienced by the victim: Ten years work camp (for a trivial cause is plainly understood), ten years professional success, and twenty years the same jacket.

When the author of the present article not long ago (1997) congratulated a St. Petersburg cab driver on the prospects of the new Russia, the outraged man replied, "For eighty years this talk about the great future!"

CICERO relates a joke which has not lost its witty effect after 2000 years: In it CATO pretended to be astonished that a haruspex did not laugh upon seeing another haruspex (*…quod not ridet haruspex haruspicem cum videt*). The intellectual feat here lies in understanding the complexity of a joke in a joke.

7.4 Jokes featuring animals

Jokes having animals as objects deserve special attention. Humans are deeply aware of the difference between human and animal behavior. Therefore, it is absurd and in a given situation comical or witty if human behavior is attributed to animals or animal behavior to humans. Examples show that not keeping this difference in mind is the source of the humor: "A dog is more likely to accumulate a stock of sausages than a politician to resist the temptation to spend a budget surplus." Here the wittiness lies in the absurdity of a dog practicing human foresight combined with the scorn inherent in the allegation of a politician having a mentality not so different from that of a dog.

The aspect of scorn is always present when humans are compared to animals, such as the arrogant Roman nobleman, Publius Nasica, asked a farmer with whom he had just shaken hands, why his hands were so hard—did he use them for walking?

A South African Bushman observes a hatching ostrich hen keeping, for unknown reasons, one egg separate from the clutch. He calls the hen forgetful as she needs the odd egg to remind her of what she is doing.

This is proof that the structure of wit is the same in all human societies and that cultural differences have an influence only in so far as they provide certain behavioral norms and experiences—the breach of which supplies the material of wit. To endow an ostrich hen with human forgetfulness and the capacity for an ingenious remedy would be thought absurd by any society and, hence, suitable for a witticism anywhere.

7.5 The comical in drawings

The comical in cartoons also deserves separate consideration, as various categories of the comical and wit overlap in this case.

As cartoons are a visual medium, one should expect the comical to predominate; that is, the drawing should show exaggeration of figures and shapes. This is, indeed, the essence of a cartoon's humor. In reality, however, the genuinely comical is not common in cartoons, as it is found only in the drawings of rare great cartoon artists whose work is of such high caliber that it cannot be described verbally. The humor in cartoons lies not only in the exaggeration, but also in the affected naivité of the cartoon artist. The cartoonist plays the role of an entertainer at whom we may laugh.

There are cartoons that present the comical in a purely visual, non-verbal manner, as in the drawing, for example, of a man who has tied one end of his hammock on a lone tree and, shading his eyes with his hand, scans the horizon in a futile search for another tree. This is a non-verbal joke in the form of a drawing, which gains zest by the drawing. Verbally presented, it hardly provokes a smile.

It is remarkable that cartoons are often enjoyed in spite of the very poor quality of their humorous essence. Clearly, cartoons are a particularly efficient media for the creation of a comical effect. This may reflect the phylogenetic history of the comical—its most original form was visual.

The captioned cartoon is transitional to verbal wit. In some, the cartoon drawing may be just an accessory to the verbal joke in the captions. An example would be a cartoon published in the liberal newspaper, "Jerusalem Post", when Prime Minister Begin was in office. In it were the outlines of two people, possibly journalists of the JP, and the following captioned conversation:

"Mr. Begin says the JP is partial."

"He also blames the JP for supporting our enemies."

"In addition, he says the JP undermines national morale."

"But still, he did not go all the way. He did not say the JP is controlled by Jews."

8. The Biological Function of Wit

8.1 The understanding of wit as a means of qualification for admission to groups

Scorn and sarcasm exist as exceptions within witty communication; much more common is unaggressive, harmless, entertaining humor. Laughing is not laughing at any more; no one is the laughed at outsider to be excluded from the group. Therefore, the question is, what is the function of ever-present wit?

A careful consideration makes one suspect wit and humor are by no means completely benign. Wit can be a means of exclusion of an individual, as anyone will quickly become an outsider if he does not understand and share the group's sense of humor. A minimum level of intelligence is indispensible for the understanding of group humor by its members—a feat which is sometimes impossible for those of lesser intelligence. This points to what the adaptive function of wit may be: while practicing humor the group selects group members with regard to their intelligence. Only those are accepted who meet average group intelligence, as they have the ability to share the group's sense of humor. This group reaction has the same effect as the practice of derision—of laughing at someone. The less gifted is identified by his lack of or inappropriate response to jokes or witticisms (he does not understand them), is excluded from the group and his opportunities for reproduction diminished. This should have—as in case of derision—the consequence of higher group efficiency and, thus, the higher individual fitness of the remaining group members.

8.2 Laughing as group phenomenon

This hypothesis is supported by certain well-known facts, the first of which is that laughing is contagious. We are predisposed by selection to laugh with a group even if we find the humor less than compelling and would not have laughed in different circumstances (cf. BERGSON 1900). It is apparently advisable to laugh along with one's group in order to avoid the possibility of exclusion. Secondly, the above proposed evolutionary scenario explains the fact that laughter is a reflex, and that its suppression is difficult if we feel something is truly humorous. Only as a reflex can laughter become a group phenomenon, and it is a group phenomenon most of all. The reflexiveness of laughter must also have an adaptive function and that function seems to be to create a situation in which a less capable individual betrays—by not laughing—his defectiveness and loses his chance for admission into the group. This explains why we have such a finely-developed sense of hearing enabling us to recognize hollow or mimicked laughter.

A certain analogy can be drawn between the impossibility of credibly producing false laughter and the phylogenetically older humor derived from imitation. Through the process of selection we have acquired the ability to identify, for the purpose of group expulsion, fakers who pretend to understand a joke or witticism with forced laughter, and, likewise, the ability to recognize and eventually expell those who imitate but do not understand the concepts that guide group behavior. The original function of imitation is so remote that imitation is practiced today exclusively for entertainment.

However, imitation is the most precarious form of clowning in the striving for group esteem, as the subconscious memory that imitation was once a sign of inferior intellect, is still very much alive.

We may be reminded here of BERGSON's definition of the comical as "mechanical, inelasticity" of behavior, a behavior which we interpreted as "imitation". As a definition, BERGSON's phrasing is perfect as it has a high degree of abstraction. In contrast, "imitation" is already an interpretation of "mechanical and inelastic" behaviour and, thus, the use of the term imitation allows a connection to an ethological hypothesis, something which is impossible with the term inelasticity. The same may be said of "unresponsiveness towards one's surroundings" and the interpretation of such behavior as grotesque or stupid.

8.3 The role of laughing in the selection of sexual partners

The inclination to form groups is remarkably strong with the young among whom laughter is of particular importance. They are very prone to laughter with a much lower threshold for doing so than adults who consider such youthful jocularity as pure silliness, but tolerate it because their own youthful behavior is often remembered with fondness. It is further remarkable that groups formed in youth usually disintegrate later on, allowing one to conclude that such groups, with their special sense of humor, fulfill a certain objective and then become obsolete. It is not farfetched to assume that the objective, so important in the youthful phase of life, is a sorting-out process for sexual partnerships. The group functions as the market place where potential sexual partners meet, evaluate each other and pair.

Since wit and laughter play such a notably important role in groups of youths, one must assume this behavioral pattern has an important function, and this function still lies in the exclusion of the less intelligent. At first sight, it is difficult to understand how this function is beneficial to the fitness of the group members. The higher efficiency of group cooperation cannot be drawn into consideration because youth group activities are notable for not being connected to joint work efforts, but instead, seemingly, to shared amusement (and so actually to sexual partner selection). How can the exclusion of potential partners result in an increase of individual fitness? The loss of potential sexual partners is for average individual fitness something like a zero-sum game, and, actually, even has a negative effect. The exclusion of a female is a disadvantage for the males that is not completely offset by the advantage of the remaining females enjoying less competition. The same is true—*mutatis mutandis*—when a male is excluded. The exclusion of any individual means, on average, a loss of individual fitness for the remaining group members. It cannot, therefore, form the basis of a mechanism of selection. That the offspring of the group should be more intelligent as a consequence of such exclusions is not a valid argument. Fitness is an individual characteristic. It manifests itself in the number of offspring begot by individuals. That a behavior could be selected for that would increase the fitness of future generations, while diminishing the fitness of the present is not a plausible evolutionary mechanism.

8.4 Wittiness and laughter promote assortative mating

The exclusion of a male or female from the group has by no means the same results as implied above, as the consequences of sex differences also have to be considered. The exclusion of a male can almost be completely compensated for by the forming of multiple partnerships of the remaining males. Multiple partnerships by the females do not, on the other hand, result in a higher birth rate which would balance the exclusion of members of their sex. There is a higher average individual fitness decrease resulting from the exclusion of a female than with the exclusion of a male. This is a consequence of the different contributions of the sexes to procreation leading to dimorphisms in physical appearance and behavior of the sexes. It should also be responsible for the the sexual dimorphism in respect to laughter discussed in paragraph 8.5.

Since the exclusion of a group member of either sex leads to a decrease of average fitness in the remaining group members—albeit of unequal effect depending on the sex of the excluded member --, one has to search for a mechanism that compensates for this decrease in order to account for the obvious importance of humor in youth groups discussed above. Wherein could the fitness gain lie which motivates an individual (regardless of sex) to accept only those (regardless of their sex) into the group who share the group's sense of humor? As a similar sense of humor is equivalent to being of similar intellect, one may guess sexual partnerships between those of similar intelligence are more successful than between those of differing intelligence—more successful in that such couples will beget on average a higher number of children for whom they will be able to care until they are sexually mature in turn. This is a reasonable assumption as families are economic units with a chance for greater success if their partners understand each other and can discuss matters on equal intellectual terms. Groups of youth with an ongoing exchange of wit and laughter, with all its consequences, provide their members an increase of individual fitness by reason of the fact that all sexual partnerships formed within the groups are more efficient. The higher efficiency is due to a group's higher intellectual homogeneity, which, in turn, diminishes the likelihood of errors in partner selection.

This view is supported by the evidence of intelligence indeed being an important parameter in partner selection. What is assumed on the basis of mere intuition is confirmed by psychometrists: "Sharing genetic wealth is … even rarer than sharing capital assets" (JENCKS 1972, p74).

The function of wit in partner selection is, like verbal wit itself, apparently a later development, partly related but not identical to the older phenomenon of exclusion from the group by derision. It is notable that non-admission to a youth group is a far less painful experience than being laughed at. Derision evolved when social organization was fairly primitive. There was no society but the group. Being excluded from the group left one no alternative and was a matter of life and death—hence the intense feeling of pain when being laughed at. More advanced societies, however, are not only larger, but also stratified. There are various groups that can be joined. One strives to be accepted in the group considered appropriate, but rejection is not a catastrophe as alternatives are still available.

CAVALLI-SFORZA and coauthors (1994, p65) discuss the assumption (since proven, see KRINGS 1997) that the ancestors of modern man and the Neanderthals did not interbreed. They conjecture Neanderthals had a lower linguistic capability than anatomically modern humans and assume that this worked as a cultural barrier against interbreeding. As a linguistic basis is indispensible for wittiness, an assumed absence of this skill (but not necessarily of all ability for comical humor, especially) in Neanderthals may have contributed to this barrier. Laughter and smiling as part of courtship behavior (as discussed below) should be seen in this context. Courtship behavior is an efficient isolating mechanism between populations (MAYR 1997, p185).

8.5 Sexual dimorphism in laughter

GRAMMER/EIBL-EIBESFELDT (1990) have evaluated experimental dyad meetings of young adults of both sexes previously not known to each other. Only insignificant behavioral differences between the sexes were observed in same sex meetings: Man with man laughed as frequently as woman with woman and the frequency of synchronized laughter (laughing together) was also not different. But differences appeared in mixed sex meetings: Laughing frequencies were markedly depressed in both sexes, but more so in males. And so was synchronized laughter. In addition, it was found that the intensity of female laughter (vocalized vs. non-vocalized) was a measure for the female's (self-reported) interest in further contacts with the male. On the other hand vocalized laughter of a female roused the interest of a male while non-vocalized laughter suppressed it. Synchronized laughter indicated mutual interest. The observations were exclusively made on dyads of same or mixed sex, thus not on larger groups. Yet, we dare to conjecture that the data demonstrate the working of the group as the marketplace where potential sexual partners meet and evaluate each other. Within a group, interaction can gradually shift to a personal communication which is instinctively understood by the involved individuals. Vocalized laughter is reflexive and indicates the male attempt at humor was a success. The laughing woman signalizes interest and rouses at the same time the interest of the man. The fact that they share the sense of humor is a sign that they are probably also intellectually a good match. Non-vocalized laughter is a more controlled laughter, it may be understood as a sign of mere politeness, if not as a sign of mockery. The two involved individuals are not attracted to each other, probably for a good intellectual reason.

PROVINE (1993) describes and interprets in a publication 1200 "laugh episodes" observed and recorded at random mostly on a university campus (mostly dyads, but also larger groups). While GRAMMER/EIBL-EIBESFELDT (1990) do not discuss the causes of laughter and which sex took the initiative for witty talk, PROVINE confirms men are more often engaged in witty talk and women laugh in general more often than men. In addition he found that an audience of either sex is more inclined to laughter if the speaker is male. Similar observations have been made by various other authors (CHAPMAN et al. 1980;, ZIV 1984; CASTELL/GOLDSTEIN 1977; MCGHEE 1979).

It amazes that in PROVINE's publication only 10% to 20% of the laugh episodes were estimated by the observers to be humorous. The laugh-provoking in the episodes may have included "a multitude of nonverbal and postural cues" (that is, what we called comical). Also the age of the sample must be considered (an average of appr. 24 years). This is an age of a notoriously low laughing threshold and of particularly low standards of humor. Nevertheless, laughing as a display without connection to the humorous certainly occurs. This is explicitly stressed by the author: "Laughter deserves more attention than that accorded it as a behavioral curiosity related to humor".

The assumption that wit and laughter favor intellectually assortative mating is not an explanation for the sexual dimorphism observed. Why do males take the intitiative to amuse while women wait to be amused? Why is it not the way around? The answer is almost banal and lies in the different contributions of the sexes to procreation. The male expenses of time and metabolic energy in procreation are much lower than the expenses carried by women. Showing the initiative in courtship (and other fields) is more rewarding for men than it is for women because initiatives may lead to multiple partnerships, something which increases the fitness (that is, the number of offspring) of men, but not of women.

The role of laughter as a courtship ritual may explain sufficiently why women are more ready to laugh and men more likely to try to amuse. Yet, this behavioral tendency may have been reinforced by the interpretation of laughter by the group. Not to laugh—that is, not to share the group's sense of humor—increases for both sexes the chances for an expulsion from the group. But there is a good reason not to expell women. The exclusion of a less capable man increases the group efficiency and consequently the average individual fitness of remaining group members but the exclusion of a woman is first of all detrimental for the average fitness. The ability of women to not only bear children but to nourish them for nine months before being born and for several years after is such an important contribution to average fitness of the group that in comparison the intellectual capabilities of women have less weight. For a good reason does the intelligence test of laughter work less rigidly with women.

It may be necessary here to stress that this does of course not mean that women are less intelligent than men. What it means is that in human evolutionary history intellectual capabilities were less important for the fitness of a female than of a male.

It is, anyhow, well known that selection affects the male more than the female sex. In populations under normal selection pressure, to remain without offspring is more likely for males than it is for females (HESCHL 1998, p220).

8.6 The humorous entertainer as a result of selection

The role as courtship ritual should not overshadow the fact that wit and laughter is still most of all a group phenomenon with sexual ramifications and not primarily a courtship ritual. As such it is probably a later and dependent development. This view is supported by observations that laughter is more frequent in same sex rather than opposite sex dyads as otherwise it would not or would occur less frequently in groups. However, the opposite is true: it is more frequent in same sex dyads (GRAMMER/EIBL-EIBESFELDT 1990). It is more plausible to assume wit and laughter evolved first in groups and was adapted to a courtship ritual later, rather than the way around.

We must assume, therefore, that the individual characteristic of humor, that is the capability of being a good entertainer, evolved in groups, too.

The importance of humor and wit as conveyers of certain important advantages is what initiates the need for humorous entertainment. A good entertainer is needed by a group (of youth) almost as a basis of their existence. Youth groups often form around humorous individuals. Without persons thus talented the group would disperse.

An analogy exists between humorous entertainers and musicians (and the artistically talented in general), the fitness of whom consists mainly of their ability to initiate the formation of groups—something all individuals need if they want to earn the benefits of higher group efficiencies. “The adaptive value of cohesion in human societies can hardly be overestimated” (DOBZHANSKY 1962, pp214f).

It is understandable that men compete for such group functions of a certain prominence as these accord opportunities of multiple partnerships with the consequence of added fitness. And it is also understandable that women do not take part in this competition as their fitness could not expect to benefit.

This may explain why comics, composers, musicians and artists in general are much more likely to be male rather than female.

While the different contribution of women to procreation suppresses or does, at least, not encourage initiative in group activities it does, in contrast, encourage other abilities. A careful selection of sexual partners is more important for women than for men and this makes women keener observers with a more reliable judgement. “What do women say about him?” was the question of Francis I of France before forming his opinion of a man.

In various publications submissiveness is considered to be one of the messages of laughter (e.g., GRAMMER/EIBL-EIBESFELDT 1990) and, therefore, dominance a trait of the humorist. “Professional comics seem to be highly dominant” (MCGHEE 1986). Thus, the clown as great dictator. This view is probably a fallacy as they are not uncommon in the discussion of wit and laughter; it may belong to the same category as “suddenness”, “relief from restraint”, “incongruity” etc. What is correct in some cases is not necessarily a general characteristic of the trait. When the boss tells a joke and laughs, his coworkers are well-advised to laugh with him. This is common knowledge but the interpretation as submissiveness is not necessarily correct and the extention of this feature to other occasions of laughter is certainly not warranted. Not to laugh at a joke signals to the jokester that his sense of humor is deficient and that he is the object of pity rather than respect. Politeness and prudence advises one not to confront a boss with such a message. One may call this submissiveness, but it will be difficult if not impossible to find such submissiveness in the numerous other occasions of laughter. The same is true for the laughter of a woman in reaction to the joke of a man, since it signals encouragement, not submissiveness.

It may be appropriate at this point to stress that almost no characteristics exist which would have a genetical basis exclusively male or female. The sex determining chromosome Y is too small to hold much information that would go beyond mere sex determination. All individuals of a population—male or female— share one gene pool, that is, they share all genes at random. What differs is the extent of expression of some genes. And we may speculate that here also lies an explanation for the smallness of the Y chromosome. Genetic information specifically connected to sex must have been detrimental to fitness and, therefore, evolution eliminated from or never allowed general information on the Y chromosome (which is true at least for all mammals).

9. Smiling in Response to Wittiness

The earlier assumption that smiling is an expression of happiness was rejected by KRAUT/JOHNSTON (1979) on the basis of statistical oberservations in favor of the interpretation of smiling as a signal in social interaction. (The assumption of smiling as expression of happiness was favored by DARWIN himself. Thus, sometimes even good HOMER dozes.) The function of smiling as a social interaction is nowadays generally accepted and the assumption of a phylogenetic relationship to laughter (the “diminuitive theory”) is generally rejected (EIBL-EIBESFELDT 1978, p249).

Smiling can also be observed among chimpanzees where its easily identified “silent bared teeth display” expresses a friendly, appeasing, submissive disposition: “What I am doing is not meant to be threatening” (compare also GOLDENTHAL et al. 1981). Various forms or functions of smiling in humans are nowadays assumed: that of submission to counteract aggression, as in primates; that which strengthens or creates special bonds between mother and newly born child; and the flirtatious smile of courtship (EIBL-EIBESFELDT 1978).

It amazes that smiling in response to wittiness is hardly mentioned in the literature.

How could smiling have adopted the additional function of indicating the appreciation of wit and humor? Sexual selection should have contributed significantly to the evolution of this variant expression. As with laughing, smiling is a form of communication. Yet smiling is the more intimate and subtle form as it is usually between two people with mutual understanding of the humor in what is occurring and thus able to demonstrate to each other an equal intellect. A situation exists in which it is critical for the two people involved to have a sense of each other’s intellectual capability, and that is the selection of a sexual partner. With intelligence being, in the words of one expert “the most appreciated human characteristic”, and an important aspect of the biological fitness to be passed on to one’s offspring, it is, therefore, a determining trait in one’s choice of a sexual partner. Smiling, as a test of intelligence, reduces the likelihood of one forming a sexual partnership with an individual of inferior intellect.

It is easy to understand how smiling could have evolved as a reaction to humor. Originally a generalized expression of non-aggression, smiling became (possibly later) a facial expression of the bonding between mother and child and finally a part of flirtation in the courtship process.

Reacting to humor with smiling may have begun as a chance variant of the already existing mode of expression. The new type of smiling could not serve as a means of communicating intellectual capability as long as the frequency of the controling gene remained low and, therefore, the message not being understood. What accelerated the establishment of smiling in response to humor as a general human trait may, however, have been the possibility that it could be misunderstood. By being taken as a message of encouragement in flirtation, the “humor smile” may have been slightly adaptive from the very beginning, especially if encountered in a female. With increasing frequency of the gene, the message was more often understood correctly to the extent spreading of the gen would have become a self-accelerating process.

The assumed establishment of the trait profited, thus, from the circumstancial and temporal closeness of flirtation and the all-deciding sexual act, which, after all, is the consummation of biological fitness. In this sense, all traits have a connection, and sometimes a close connection, to the sexual act, a fact which may have fostered a sinister understanding of sexuality in certain psychological theories, which continues to amaze and amuse those brought up in DARWINIAN thinking.

Smiling and laughing are related only due to the fact they are both a reaction to the display of wit. In our closest animal relative, the chimpanzee, smiling and laughing are completely independent phenomena—the common bond of humor being missing.

10. The Social Significance of Humor

Intelligence is of eminent social importance, a fact rarely spoken of but universally known. Because of the relationship between intelligence and humor (in its widest sense), humor is of considerable social importance. It is, therefore, understandable that it functions as a criterion of differentiation—that it is specific to socioeconomic groups.

In “Romeo and Juliet”, Lady Capulet discusses Juliet’s age with her nurse when considering her daughter’s marriage. The conversation reminds the nurse of an incident in Juliet’s childhood which she remembers as being extremely funny. When little Juliet fell and burst out in tears, the nurse’s husband—”God be with his soul!”—made the humorous remark, “Thou wilt fall backward when thou hast more wit, wilt thou not, Jule?” to which the crying child replied, “Ay”. In spite of the attempt of the annoyed Lady Capulet and Juliet to silence her, the nurse repeats the story saying she would not forget it "should she live a thousand years".

It is remarkable that it was the nurse's husband who made the witticism. The nurse shared with her husband the same sense of humor and, therefore, the same intellect. We may ponder the test of smiling that once brought the two together?

Author's address

Winfried Behr, Friedrich-Breuer-Straße 86
D-53225 Bonn, Germany.
Email: behrbonn@t-online.de

Acknowledgments

I wish to thank Elke BERGER of Berlin for her continuous interest in an esoteric topic, and Raymond TUBBS of Hartford, Connecticut, for editing the text. He imparted brevity and eliminated unintentional humor derived from the linguistic creations of a non-native English speaker.

References

Bergson, H. (1900) Le Rire: essai sur la significance du comique. Félix Alcan: Paris. Quoted from Dietzsch (1993) Luzifer lacht. Philosophische Betrachtungen von Nietzsche bis Tabori. Reclam Verlag: Leipzig

Castell, P. J./Goldstein, J. H. (1977) Social occasions for joking: A cross cultural study. In: Chapman, A. J./Foot, H. C. (eds) It's a Funny Thing, Humor. Pergamon Press: Oxford, pp. 193–197.

Cavalli-Sforza, L. L./Menozzi, P./Piazza, A. (1994) The History and Geography of Human Genes. Princeton University Press: Princeton.

Chapman A. J./Smith, J./Foot, H. C. (1980) Humor, Laughter and Social Interaction. In: McGhee, P. E./Chapman, A. J. (eds) Children's Humor. Wiley: New York, pp. 141–179.

Dietzsch, S. (1993) Luzifer lacht. Philosophische Betrachtungen von Nietzsche bis Tabori. Reclam Verlag: Leipzig.

Dobzhansky, T. (1962) Mankind Evolving. Yale University Press: New Haven.

Eibl-Eibesfeld, I. (1986) Die Biologie des menschlichen Verhaltens. Grundriß der Humanethologie, Piper: München, Zürich, pp. 302–332. Quoted in Grammer, K./Eibl-Eibesfeldt, I. (1990).

Eibl-Eibesfeld, I. (1978) Grundlagen der vergleichenden Verhaltensforschung. Piper: München, Zürich.

Eibl-Eibesfeld, I. (1989) Human Ethology. Aldine de Gruyter: New York.

Goldenthal, P./Johnston, R. E./Kraut, R. E. (1981) Smiling, Appeasement, and the Silent Bared-Teeth Display. Ethology and Sociobiology 2: 127–133.

Grammer, K./Eibl-Eibesfeldt, I. (1990) The ritualization of of laughter. In: Koch, W. A. (ed) Natürlichkeit der Sprache und der Kultur. Bochumer Beiträge zur Semiotik, Brockmeyer: Bochum, pp. 192–214.

Heschl, A. (1998) Das intelligente Genom. Springer Verlag: Berlin, Heidelberg.

Hooff, J. A. R. A. M. Van (1972) A comparative approach to the phylogeny of laughter and smiling. In: Nonverbal Communication. (Hinde, R.A., ed.). Cambridge University Press, Cambridge, pp. 209–241. See also Eibl-Eibesfeldt (1978).

Jencks, C. (1992) Inequality. Basic Books: New York, London.

Kraut, R. E./Johnston, R. E. (1979) Social and Emotional Messages of Smiling: An Ethological Approach. Journal of Personalyity and Social Psychology 37(9): 1539–1553.

Krings, M./Stone, A./Schmitz, R. W./Krainitzki, H./Stoneking, M./Pääbo, S. (1997) Neandertal DNA Sequences and the Origin of Modern Humans. Cell 90: 19–30.

Kulka, T. (1990) The Incongruity of Incongruity Theories of Humor (in Hebrew). Iyyun 39 (2): 223–235. (The edited abstract of this article is quoted).

Mayr, E. (1997) This is biology: the science of the living world. The Belknap Press of Harvard University Press: Cambridge MA.

McGhee, P. E. (1979) Humor: Its Origins and Development. Freeman: San Francisco.

McGhee, P. E. (1986) Humor across the life span: sources of developmental change and individual differences. In: Nahemow, L./McCluskey-Fawcett, K. A./McGhee, P. E. (eds). Humor and Aging. New York: Academic Press, pp. 27–51. Quoted from Weisfeld (1993).

Mellars, P. (1998) The fate of the Neanderthals. Nature 395: 539.

Morreal, J. (Ed.) (1987) The philosophy of laughter and humor. SUNY Press: Albany.

Provine, R. R. (1993) Laughter Punctuates Speech: Linguistic, Social and Gender Contexts of Laughter. Ethology 95: 291–298.

Provine, R. R. (1996) Laughter. American Scientist 84: 38–45.

Rapp, A. (1947) Towards and Eclectic and Multilayeral Theory of Laughter and Humor. The Journal of General Psychology 36: 207–219.

Rapp, A. (1949) A Phylogenetic Theory of Wit and Humor. The Journal of Social Psychology 30: 81–96.

Rosten, L. (1968) The Joys of Yiddish. McGraw-Hill Book Company: New York.

Weisfeld, G. E. (1993) The Adaptive Value of Humor and Laughter. Ethology and Sociobiology 14: 141–169.

Ziv, A. (1984) Personality and Sense of Humor. Springer Verlag: New York.

Ladislav Kováč

Fundamental Principles of Cognitive Biology

Abstract

Cognitive biology aims at a synthesis of data of various scientific disciplines within a single frame of conceiving life as epistemic unfolding of the universe (the epistemic principle). In accord with evolutionary epistemology, it considers biological evolution as a progressing process of accumulation of knowledge. The knowledge is embodied in constructions of organisms, and the structural complexity of those constructions which carry embodied knowledge corresponds to their epistemic complexity. In contrast to evolutionary epistemology, cognitive biology is based on the assumption that the molecular level is fundamental for cognition and adheres to a principle of minimal complexity, which stipulates that the most efficient way to study any trait of life is by studying it at the simplest level at which it occurs. Several principles of cognitive biology are similar to extremum principles of physics and may represent tight junctions between ontic and epistemic realms. A principle of minimisation of suffering is derived from the evidence of human conscious emotional experience. It has a bearing upon such notions as onticity, fitness, altruism, science, future of mankind. This principle, just as all the other principles of cognitive biology, is descriptive, not normative.

Key words

Algorithmic complexity, cognitive biology, complementarity, GÖDEL, epistemic complexity, epistemic principle, extremum principles, JAYNES, logical parity, minimisation of suffering, molecular recognition, ratchet.

A modern molecular biologist might paraphrase the poet Pope by saying, The proper study of mankind is the bacterium (KOSHLAND 1977).

Introduction

The expansion of molecular biology, which has begun in the early 1960s, is both exciting and appalling. It is providing ever deeper insight into the mechanisms of functioning of the living cell. This enables, in turn, ever more powerful interventions into those mechanisms, with consequences which may be already transcending human capacity to foresee, evaluate and control. The results of manipulations with the nucleus of the cell may turn out to be more far-reaching than have been the results of manipulations with the nucleus of the atom.

Disquieting is the progressing instrumentalisation of molecular biology. More and more science is turning into technoscience, or even a sheer engineering. Priorities of research are being reoriented toward technology. In a highly competitive field of research, with complicated methods, instrumental knowledge has a priority. The situation in molecular biology has its parallels in other areas of contemporary scientific endeavour. Our instrumental knowledge, our ability to manipulate things and events, singled out and treated out of their context, is great, our understanding the underlying complexity, which results from their interconnections and of which they are parts, is poor and lagging behind. In addition, by assembling the disentangled things and events into new combinations we are creating new systems of unprecedented complexity.

The discrepancy in how little we know relative to how much we can do may be unique in the history of terrestrial life (but not, as it will be reasoned later, in the history of life in the universe). Doing and knowing have been so far a unit, two sides of the same evolutionary coin. Applying to non-human organisms they may be considered as two names for the same target of evolutionary selection. The precarious disparity in human knowing and doing, in cognition and behaviour, need be urgently diminished or abolished and this calls for setting priorities of scientific research toward profound comprehension of human cognition and behaviour. Konrad LORENZ, who had been incessantly stressing this need (LORENZ 1983a), would be pleased to witness the recent upsurge of cognitive sciences.

In general, however, cognitive sciences aim at revealing *mechanisms* of cognitive processes. A danger of instrumentalisation and also of misuse of knowledge is considerable (CHANGEUX, J. P., cited in BUTLER 1998). In a somewhat vague contradistinction of cognitive sciences is epistemology, which aims at understanding the very *nature* of knowledge: what is it, how is it possible, how it evolved, what is its role in the universe. Epistemology has been traditionally a major discipline of philosophy. However, recent attempts at its "naturalisation", mainly under the name of genetic epistemology (PIAGET 1967) or evolutionary epistemology (RADNITZKY/BARTLEY 1987, RIEDL/DELPOS 1996, as well as the references therein), may be conceived of as a part of a process of detachment from philosophy of its hitherto genuine disciplines.

Philosophy does not produce knowledge, although it has been intended to do so; it has been and continues to be a source of emotionally and socially effective beliefs. Only some of them may, by chance, prove to map some features of the world into human-made reality. This statement does not underrate the importance of philosophy. As human longing for meaning and universal understanding springs up deep from the emotional ground, philosophy will continue to thrive (the backlash of postmodernism and the fashion of pseudoscience notwithstanding) and epistemology will continue to be part of philosophy. It is conceivable that epistemology as part of science may evolve not to become a natural science, but rather a formal science like mathematics, cybernetics, statistical physics, game theory.

Can molecular biology constitute a part, or even foundation, of "naturalised" epistemology? Instead of serving just as a powerful instrumentarium in revealing the mechanisms of brain functioning, may it open a novel path toward understanding *the essence* of cognitive phenomena, including emotion and even consciousness? However subtle, or even futile, the distinction may appear at first sight, it may, in fact, have a profound meaning. It would promote some views and concepts of molecular biology, successful so far mainly in instrumental science, into the realm of conceptual science. A "bottom-up" approach to epistemological problems, that encompasses molecular biology, has been called cognitive biology (KOVÁČ 1986a). Owing to its ample use of concepts and reasoning of thermodynamics, it may be considered as an outgrowth of bioenergetics (KOVÁČ 1986b; 1987). Some pioneering ideas have been formulated by GOODWIN (1976). The main credit should be given to Hans KUHN. For him, life from its very beginning, starting from self-copying nucleic acids, was an unceasing process of accumulation of knowledge (KUHN 1972; 1988).

This paper is a short outline of essential principles which, altogether, provide a rationale for cognitive biology. Their list is not exhaustive. Although some of them may appear to have a normative character, it will be argued that they all are descriptive. Their more formal and detailed description will be subject of subsequent publications.

Elaboration of Conceptions

1. The principle of logical parity

Regarding the interest of biologists in the most various aspects of human cognition it is rather surprising that little attention has been given to the biology of logic. Even in studies of evolutionary epistemologists such considerations are rare. POPPER insisted in his early writings upon a strict separation of logic and psychology in the analysis of scientific discovery (POPPER 1957). According to VOLLMER (1987, p180), the foundation of logic is one of the most difficult problems and is hitherto unresolved. It can be inferred that for LORENZ human logic was a species-specific outcome of the human evolutionary trajectory, yet correctly reflecting relevant features of the world (LORENZ 1973). On the other hand, for PIAGET, logic is not innate to any human individual in the sense that it exists at any age. Logico–mathematical structures are extracted in the course of ontogenesis from operations on the surroundings; accordingly, there are a sort of abstract manipulations of the adult human subject with the objects in his/her environment (PIAGET 1967). Probably the most extensive analysis of the evolutionary nature of logic has been provided by RIEDL (1979; 1992). According to him the evolution of logic is closely linked to the evolution of language and logic is also determined by a language. European logic with its subject/predicate structure is conditioned by the structure of European languages and may differ from that of Chinese.

Against the view that logic is "human-specific" or even "culture-specific" and that a number of internally consistent but mutually excluding logics may be feasible in the world, it may be argued that computers, non-living machines, do obey the rules of formal logic. A rather naive rejoinder may point out that it need be so because the computers are the invention of a single culture, the same that has "invented" formal logic. Less naive may appear another rejoinder stating that the operation of computers must be isomorphous with the operation of the human mind, not culture-specific but necessarily species-specific just as the mind itself is, since the computers are just "*exomental instruments*" (in analogy with mechanical tools and machines that have been called "exosomatic instruments" by LOTKA and the term popularised by GEORGESCU-ROEGEN 1971, p307) of the human species.

It will be shown later by evolutionary reasoning that the computer metaphor of the human mind is misplaced and misleading, and accordingly the two arguments with the computers lose substance. It seems feasible that, in the course of evolution, motor behaviour of living organisms, mechanical work on the surroundings, moving and rearranging objects in space and time, has been continuously becoming less overt, has been internalised, more and more reduced to pretended actions, transformed into internal virtual manipulation with the objects. The final achievement of this ever-growing abstraction may be human thinking—an abstract motor behaviour. VOLLMER (1987, p104), referring to LORENZ (1943), considered "das Hantieren im Vorstellungsraum" (handling in the imagination space) as the initial form of thinking.

An object cannot simultaneously be and not be at the same place; two different objects cannot at the same time occupy an identical region of space, etc. But this is precisely what is being said, in abstract terms, by logic. All rules of formal logic are nothing but an abstract translation of the physical necessity. By implication, logical operations of our mind or of our computers are mere tautological *translations*, applied in order to make the facts of the world more comprehensible to us. Due to these operations we are able to organise our sensations and conceptions to *construct* reality. Reality is, indeed, species-specific and also culture-specific, and to a considerable extent even individual-specific, but there is no species-specific or culture-specific logic: Logic and the world are the same thing, split into two by the construction and limitations of our mind. This separation, illusory duality and, at the same time, complementarity of the world and the mind, is being named the principle of logical parity.

The logic as a phenomenon resulting from the world/mind splitting is a ground on which the logic as a formal science, with all its branches (such as relational, modal, temporal, deontic logic), has been built up. Only the latter is, as science in its entirety, our construction, part of human-specific reality. It should be properly called *the science of logic*. It helps us to create and categorise concepts in such a way that they facilitate our understanding of the world (the conceptual aspect) and our manipulation with the world (the instrumental aspect). Cognitive biology insists upon a clear definition of concepts, chosen according to a criterion of their maximal conceptual and instrumental usefulness, stipulating that all concepts belong to the reality and none of them to the world.

It would be too pretentious to label the principle of logical parity by another name: the principle of general complementarity. Considering other limitations of the human mind, BOHR introduced a principle of complementarity to account for particle–wave duality of quantum physics. We live in a world of "medium" dimensions, a macroworld. (This is the term standardly used in physics. It corresponds to the term "mesocosm" proposed by VOLLMER 1987). Our mind can perceive phenomena of this macroworld and to conceptualise them appropriately, but has not been constructed to conceive of the microworld, the world described by the mathematical formalism of quantum physics. To give this formalism a "human-tailored" explication we have to take recourse to our percepts and concepts fitting the macroworld. This is why we conceive of an electron as a particle in explaining the results of one experimental arrangement and as a wave in interpreting the data from another experimental arrangement.

According to BOHR (1964) we have to apply the same principle of complementarity in explaining brain–mind dualism. Following his reasoning we may call the pertinent world the "psychoworld". It seems that a similar principle may be needed for explaining phenomena of the "megaworld", the world of galaxies, the subject of cosmological inquiry. Our world is not of "medium dimensions" in space only, but also in complexity. Complementarity of the "psychoworld" may be mainly conditioned by its "high-dimensional" complexity. This is why we may need the principle of complementarity in our analysis of still another world, the "socioworld": a world created by cultural evolution, which, by its complexity, may escape straight-forward human understanding and need several complementary explications, each of them internally consistent and all of them mutually incompatible. The duality world/logic may underlie all these other complementarities.

The principle of logical parity has a bearing upon the concept of rationality. It will be shown later that it may be useful to distinguish conceptual rationality and instrumental rationality. Each of the two represents not a state, but a process. In both of them the aim is to connect two distinct points by a trajectory.

Ideal rationality corresponds to the shortest trajectory which, in a simple case, would be a straight line. The complete absence of rationality is represented by random walk. The "real" rationality, including the "bounded" rationality of humans (SIMON 1983), is situated between the two extremes, and is conditioned by the amount of available knowledge. It is immediately apparent that the concept of rationality is on a par with the concept of extremum principles in physics, which seem to be the basis of all fundamental laws of nature (FEYNMAN et al. 1966b). There must be a profound meaning in this parity that has not been sufficiently appreciated yet.

If physical necessity can always be translated into logical necessity, the inverse translation is also a possibility (except that logic deals with all feasible worlds, not just the one created by the contingencies of evolution of our universe). What is logical, is always possible, either virtually, in principle, or as a fact of nature. DARWINIAN interpretation of evolution has been generally considered to be a theory, a useful explanation of our observations of living nature. It is not a theory. It is a logical necessity and, by translation, a fact of any universe containing self-replicating entities. In any environment, containing restricted amount of resources, exponential growth, competition and selection of self-replicating entities are ensuing automatically. Biological evolution is but a particular case of the fact of evolution (EIGEN/WINKLER 1975). The replicator equation (SCHUSTER/SIGMUND 1983) is one of the most fundamental equations of the universe.

But how to prove that our universe, with its logic of self-replicating entities, is not virtual? This will be a matter of the last principle of those enumerated here.

2. The principle of double epistemic closure

There is a universal characteristic of any living system to sense relevant features of its surroundings and to react appropriately upon them in order to preserve its own permanence, its *onticity*. Indeed, the relevance of the features is determined by their value, positive or negative, for the maintenance of onticity of the particular living system; those which are neutral have no relevance and are not being sensed. The human species may had not differed for a long time in this characteristic from all other species, until the exuberant redundancy of the human brain has enabled cultural evolution. As cultural evolution has been progressing, myths, a species-specific adaptive arrangement assuring onticity by suppression of cognitive chaos and cementing group cohesion, have been transmuting into philosophy and later into experimental science. Once established, philosophy and science, by their autonomous dynamics largely independent of human intention, have made neutral features, of no biological relevance, subjects of human sensing and appreciation, and inciters of new forms of behaviour. Self-consciousness, itself possibly a product of brain redundancy, has been allotted with an additional new function: reflection of (and on) the world.

It appears, in a simplifying course-grained view, that for the majority of early philosophers of the birth-place of philosophy, the ancient Greece, human reflection had no limits, however restricted and superficial may have been immediate perception and cognition. The *logos* of Nature was considered to be isomorphous with the *logos* of human mind and thus fully accessible to human comprehension. Gradually, hesitations and doubts were accumulating and they found their culmination in KANT's transcendental philosophy. KANT's views on limits imposed on human cognition have been biologically reinterpreted by LORENZ (1983b). This reinterpretation, anticipated by a number of LORENZ's predecessors (listed in CAMPBELL 1974) may be designated as *a foundation stone* of evolutionary epistemology. Pre-KANTIAN epistemological scepticism, for instance four kinds of "idols" of Francis BACON, may be easily reinterpreted and "naturalised" in a similar way. Deficiencies of the human mind have been extensively analysed by all evolutionary epistemologists. RIEDL did in several books and succinctly presented in a paper (RIEDL 1995).

The limits of the human mind, its possibilities and constraints, imposed by contingencies of evolution of the species, seem to be insurmountable. Due to them, our reality, a model of the world, is species-specific. (Implying that other species construct their own species-specific reality.) They confined us to the world of medium dimensions and low complexity. The worlds outside are separated by barriers which may be called *KANT's barriers*. When we attempt to cross the barriers, complementarity may be the only means of how to interpret the phenomena encountered there.

The species-specific delimitation of the cognitive capacities may be called the first epistemic closure.

There is another closure, more general and more fundamental. It was anticipated by a number of philosophers, starting from a minority of philosophical "dissidents" in the antiquity (PROTAGORAS, TIMON, ZENO), through HUME up to KANT. It has first received a rigorous formulation in 1931 by GÖDEL in his incompleteness theorem. GÖDEL's theorem concerns formal systems of mathematics and restrictions imposed upon them (for a simple exposition see NAGEL/NEWMAN 1958, HOFSTADTER 1980, PENROSE 1994). Whatever set of consistent rules one adopts for manipulating mathematical symbols in a system of axioms, there must always be some statement, framed in the language of these symbols, whose truth or falsity cannot be decided using those axioms and rules. In addition, in a closed system of finite axioms and rules, there is also no way of telling whether or not the starting assumptions are logically consistent or not. If one tries to solve the problem by adding a new rule or a new axiom, one just creates new undecidable statements. To understand the system of mathematics fully one must go outside mathematics.

As HOFSTADTER pointed out, it can have a suggestive value to translate GÖDEL's theorem into other domains, provided one specifies in advance that the translations are metaphorical and not intended to be taken literally (HOSTADTER 1980, p696). Mathematicians have often shown irritation or scorn upon witnessing mathematical outsiders backing their scepticism in various non-mathematical domains by referring to GÖDEL. However, as shown by CHAITIN (1990), GÖDEL's theorem can be seen not as an isolated paradox but a natural consequence of constraints imposed on cognition capacity by information theory. In the theory of algorithmic complexity (KOLGOMOROV 1965, CHAITIN 1975), complexity of a series of digits is equal to the size in bits of the minimal program of the series. A system may be represented by a very long series of digits, but its complexity is low if the minimal program that will yield the given series consists just of a small number of digits. The information of the series is present in a "compressed" form in the minimal program. If the minimal program is approximately of the same length as is the series, the complexity of the series is maximal, and its information cannot be compressed, because the series of digits is random. Since the minimal program cannot be compressed, it is always random. Within a system of axioms and rules of inference, it can be proved that a series is non-random—by finding a program which is shorter in digits—but there is no way how to prove that it is random and cannot be compressed any more. A computer program running in order to find out this proof will never halt. In a formal system of certain complexity no series of digits can be proved to be random (that is, of maximal algorithmic complexity), unless the complexity of the series is less that than of the system itself.

This shows immediately that GÖDEL's theorem is related to TURING's Halting Problem (see HOFSTADTER 1980, p425), TARSKI's theory of truth (TARSKI 1949) and also POPPER's falsification principle (POPPER 1957). Indeed, all these formulations imply the same statement about a system of a specific complexity (which should be valid not only for algorithmic complexity but also for other conceptions of complexity): It is impossible—and the impossibility is grounded on logic, and hence, by translation, on the properties of the material world—to achieve a complete knowledge of a system of a specific complexity with the means available exclusively within the system; to achieve it one needs also means from another, larger system of greater complexity, a *metasystem*. The size of complexity of a formal system determines the size and the limits of its *epistemic capacity*.

PENROSE's attempts to show that the conscious activity of the human brain transcends beyond computation (PENROSE 1989; 1994) can be understood along this line. Human mind may be more complex than is complexity of formal systems of mathematics. If these systems, and logic underlying them, are translatable into properties of the physical world, the transcendence of the formal systems would imply that some novel, hitherto unknown, physical principles may govern the non-computational faculty of the conscious brain. A deep relationship between GÖDEL's Incompleteness Theorem and what he called a paradox of the human brain was already taken up by KUHLENBECK (1982). He pointed out that our world of consciousness is a phenomenon of the brain, but our brain is also a phenomenon of the brain. Hence, a closure. To solve the paradox one would need another hypothetical brain that would be placed not in a domain of consciousness but in an extramental public space-time system. GERLACH (1988), who brought out this close relationship between GÖDEL and KUHLENBECK, has proposed to acknowledge the merit of KUHLENBECK by renaming the brain paradox and call it "KUHLENBECK's paradox".

HOFSTADTER has also analysed this problem. He maintained that there may be no fundamental, i.e., GÖDELIAN, reason that would bar the understanding of the human mind. It "may be completely clear to more intelligent beings" (HOFSTADTER 1980, p707). Unfortunately, biological evolution on earth has not supplied such beings yet.

Despite this evolutionary hint, the principle of the second epistemic closure is fundamentally different from that of the first closure: The statement that a system cannot be fully understood with the means of the system itself applies to any formal system, whatever is its complexity. It holds for any living systems, independently of the evolutionary contingencies which had shaped its cognitive abilities and, eventually, its species-specific reality. It is a principle of logic, not of evolution. In KANT's epistemology, both the first epistemic closure and the second epistemic closure are implicit.

Instead of evoking some divine "metabrain" to achieve a most comprehensive understanding of the essence of human cognition, one can undertake an opposite approach: to reach first a full comprehension of cognition of living beings simpler than human. Since the complexity of simpler organisms is surely lower than is the complexity of human—and this may apply also to epistemic "formal systems" of these simpler organisms—the human cognitive system may be conceived of as a "metasystem" with respect to the cognitive systems of simpler species. An extrapolation of this knowledge toward human cognition may then be a way of how to circumvent the "*KUHLENBECK's barrier*". There is out of these considerations that stems the next principle.

3. The principle of minimal complexity

There are fundamental and unresolved issues associated with the definition of complexity. Indeed, this single word is commonly used to describe quite different characteristics of quite different systems. To avoid confusion, it will be attempted in this paper to use the word "complexity" with an adjective in all cases in which unambiguous, and possibly mathematical, definitions are available, e.g., algorithmic complexity, sequence complexity, thermodynamic complexity, epistemic complexity. When used without adjective, it will mean an intuitive, even if rather vague, notion of something that is not simple, but complicated, compounded, organised.

Only a tiny part of what we know today of human heredity has been obtained in studies on human subjects. The major part has been a result of extrapolations from studies on simpler organisms. From pea of the founding farther of genetics, MENDEL, *downwards* to the fruit fly of MORGAN, still lower to yeast of EPHRUSSI and LINDEGREN and bacteria of LEDERBERG and MONOD, down to the ground of the phage of BENZER, DELBRÜCK and LVOFF, such has been a victorious *upward* path of genetics and molecular biology.

Max DELBRÜCK, who had been a successful physicist, before he turned to biology, has particularly marked this trajectory. His ambition was to describe biological systems with a similar precision as had been applied to physical systems. It was obvious to him that, to achieve this goal, most simple biological systems should be chosen. This is why he studied phototaxis of simple fungi *(Phycomyces)* as a model of organisms' reacting at the external stimuli. Historical was his decision in 1937 to take up the study of bacteriophage: it may be seen as a start of molecular biology (STENT 1963). Bacteriophage can be considered to be the simplest living "thing". The study of bacteriophage enabled the elaboration of the concept of the gene as a unit of heredity, an insight into the internal structure of the gene by intragenic mutations, the elucidation of the nature of genetic recombination. All this had been accomplished before nucleic acids were discovered as material carriers of genes and before biochemistry assumed the dominating position in molecular biology.

DELBRÜCK's success in genetics, made possible by employing the simplest system exhibiting heredity, has inspired his pupils and collaborators to use a similar approach in another discipline of science, in neurobiology. *Aplysia*, an animal with just a few hundreds of neurones was studied and later another simple organism, *Caenorhabditis elegans*. The sequence of all genes of *C. elegans* has now been reported (*C. elegans* sequencing consortium 1998) and already the first comparative analysis of genes coding for neuronal functions has revealed a striking homology with many of highly conserved neuronal genes of mammals and of human disease genes (BARGMANN 1998).

This experience, well-proven also in other branches of science, substantiates the formulation of a principle of both heuristic and conceptual significance: The most efficient way to study a concrete biological phenomenon is by studying it on the simplest organism in which this phenomenon can be found—here it is experimentally best accessible and, because of its evolutionary simplicity, theoretically the most comprehensible.

This is the principle of minimal complexity. To acknowledge DELBRÜCK's merit, it may just well be called DELBRÜCK's principle.

The principle would be easy to apply, almost automatically, if it were easy to detect the lowest evolutionary level at which the concrete phenomenon occurs. To find the lowest level is, however, the most difficult problem. The problem is linked to the question of homology and analogy. Two phenomena may be similar in appearance and nevertheless of entirely different evolutionary origin. Even if identical in form and function, their conceptual unification will only be a metaphor. A metaphor may be most misleading and may incite to naive interpretations. This is often the case when phenomena of a lower level are being explained by phenomena of a higher level, as is the standard case of antropomorphisms.

The question of homologies is connected with another difficulty. What should be the minimal resemblance and the maximal evolutionary distance to make it meaningful to speak of a homology? (WRAY/ABOUHEIF 1998). Comparative morphology has often faced this ambiguity. It is encountered in an elementary form in molecular biology when comparing sequence homology of nucleic acids or proteins. In a sense, all our genes originate from a single, evolutionary oldest, gene and relative to it all our genes are homologous. This, however, is a trivial statement of no explanatory value. A recent conference has amply shown how elusive the concept of homology is (TAUTZ 1998).

The essence of the problem becomes clear when we search for homology at levels higher than the molecular one. To what extend is social behaviour of mammals homologous to that of insects or even bacteria? Is human consciousness homologous to consciousness of other animals? If we assign to human 100% of consciousness, has the chimpanzee 80%, the mouse 2%, the fruit fly 0.01 and yeast 0.000001% of consciousness? The question concerns the very essence of cognitive sciences: is cognition a characteristic of all living forms, does life equal to cognition, or, in an opposite view, is cognition exclusively a human faculty? It is amazing that both extreme views coexist within contemporary cognitive sciences and have their respective supporters and militants. Even in evolutionary epistemology the views diverge (HESCHL 1998, VOLLMER 1985, p294). All aspects of this crucial question are too extensive to be dealt with here and will be a subject of a separate publication.

The core of the problem is obviously the phenomenon known under various names: emergence, fulguration, discontinuity, qualitative transition. To take up the example of consciousness, mentioned above, if one did not admit that there must be a threshold for the evolutionary origin of consciousness, under which there is no consciousness at all, one would obviously end in panpsychism granting consciousness even to elementary particles. Some qualitative transitions, such as phase transitions in thermodynamics or percolation in statistical physics, have their rigorous mathematical underpinning. On the other hand, some qualitative changes have been described by science as only apparent, such as visual perception, with qualitative discontinuities in colour corresponding to continuous changes in wavelengths of the electromagnetic radiation. It is in this area that the major challenge for epistemology exists: some "phase transitions" do occur in the world and are being translated into "phase transitions" in species-specific reality; some others concern phenomena of the world with no repercussion in reality; still others—the vast majority of them may be of this kind—concern exclusively the reality endowing it with qualitatively distinct phenomena which have no parallels in the world (and this is made often still more complicated due to the nature of our concepts: all concepts are mutually demarcated, enforce discontinuities and parcel out, sometimes quite arbitrarily, the reality). To discern between these three different categories of qualitative transitions is a matter of empirical research and, hence, will always be provisional. This is why a choice of a subject of research, which would exhibit minimal complexity of the phenomenon of interest, must be mainly tentative and its adequacy remains to be proved subsequently by results of the very research.

At any case, the principle presupposes that there are levels of complexity in the living world and that, in the course of biological evolution, there has being a continuous growth of complexity.

4. The epistemic principle

Ever since DARWIN there have been incessant discussions in biology as to whether biological evolution is progressive and has a direction. If there is progress in biological evolution, one can speak of simpler and more complicated, lower and higher organisms, one can attempt to find and define evolutionary tendencies or even formulate some laws of evolution. If there is no progress, such terms have no sense and may be refuted as antropomorphisms.

The substance of the argumentation of those who do not admit any progress in evolution is the statement that in a specific environment individual organisms—or, adopting a "gene eyes' view", their genes—compete with other organisms for a single utility: DARWINIAN fitness. Fitness is being achieved by various means. In the same environment different organisms optimise their fitness by adopting different strategies and many strategies may be equally successful. The environment for a single organism are not only physical conditions but also all the other organisms. Fitness of a single organism is therefore a highly dynamic function in which are variables the fitness of all other organisms. Man may appear to be an organism with a high fitness, with the present population number of 6 billions. However imposing the figure may be, it is negligible when compared with the number, or even with a cell mass, of bacteria. Some bacterial species have existed on earth since several billion of years, not much changing, and it seems probable that they will continue to exist after the species *Homo sapiens* will no longer be here. If they have survived for such a long time and also face a bright future, there is no ground for seeing a progress in evolution and no reason to label humans as higher, and bacteria as lower organisms.

The controversy about the progress in evolution may stem from a misunderstanding which is due to the ambiguous connotation of the word "progress". In European culture, at least since the Age of Enlightenment, progress has been seen as something to be wished, something valuable, "good". It has been considered as inevitable in cultural evolution: anything more progressive has had a better chance to push through, it has have, we may say, higher DARWINIAN fitness when compared with something conservative and retarding.

Omitting the normative connotation, the word "progress" is left as a neutral, valueless expression of an evolutionary tendency. The tendency is undeniable: the tendency toward appearance of ever more complex organisms. The very fact that bacteria, with simple cell organisation and simple behaviour, are evolutionary old and man, a being with the brain as an organ with the highest structural complexity as has ever appeared in evolution, is a proof. The fact is not changed by the possibility that the fitness of bacteria may be higher than is the fitness of man. It is also possible that some bacteria are evolutionary younger than man and, at the same time, of low complexity. Such a possibility does again not argue against the universal evolutionary tendency, which may be pictured by a metaphor of a complicated *maze*: life incessantly, at all levels, by millions of species, is "testing" all the possibilities of how to advance ahead. The vast majority of the species perish or end in deadlocks where they survive with no possibility to advance. To advance—where, toward what goal?

Progress in evolution has its thermodynamic reason. Dynamics of the world is irreversible, directed by the second law of thermodynamics. Without that law, the world would have a NEWTONIAN character: as an ideal pendulum in an ideal void it would persist in a monotonous, eternal movement, symmetric in time. In such a case, neither the evolution of life would have an arrow of time: DARWINIAN variations would resemble endless musical variations on a single theme. Natural selection would play reversibly with such variations.

As pointed out recently by FONTANA et al. (1997, p210), resounding the recurrent focal question of contemporary biology, "selection has no generative power; it merely dispenses with the "unfit", thus identifying the *kinetic* aspect of an evolutionary process. The principle problem in evolution is one of *construction:* to understand how the organisations upon which the process of natural selection is based arise, and to understand how mutation can give rise to organisational, that is, phenotypic novelty. " In principle, the answer was provided already more than three decades ago by non-equilibrium thermodynamics: systems far from thermodynamic equilibrium spontaneously evolve to ordered states, called dissipative structures, as a result of fluctuations (PRIGOGINE 1967). This, of course, is not a full answer: *constructions*, present in living systems, are mainly conservative structures, not dissipative, and the task remains to elucidate how the constructions, more and more complex, arise and how they are maintained. It is here that the value of the principle of minimal complexity comes to the fore.

It has been often argued that the origin of life must have been a highly improbable event, having taken place perhaps once in the history of the universe. A support for such a claim has been found, quite erroneously, in the second law of thermodynamics. It has been overlooked that the second law operates in the world governed by fundamental forces. In the field of forces elementary particles, atoms, molecules associate, forming larger units and the dissipation of potential energy has thus a creative role in generating structures. It seems now to be virtually sure that in any part of the universe, where the thermodynamic conditions of temperature and pressure are similar to those that had been some four billion years on our Earth, life must arise as a physical necessity.

With the advent of self-copying nucleic acids another new quality has appeared in the universe: *molecular recognition*. There is no recognition between two atoms of hydrogen and oxygen which, under specific conditions, associate to form a molecule of water. The association is a physical necessity, it is a *nomic* process. The association of nucleotidetriphosphates with the macromolecular single-stranded template of self-replicating nucleic acid and the resulting WATSON-CRICK pairing of the corresponding bases is another kind of process. In is an *ex post* necessity, a pseudo-nomic process, behind which is hidden historical contingency: a selection of a macromolecule with a certain specific sequence of units (which, incidentally, may have been degenerate) from an ensemble of similar macromolecules: a molecule endowed with a specific function, and, hence, exhibiting teleonomy. In contradistinction to simple atomic or molecular associations, molecular recognition is a *teleonomic* process, a result of evolutionary contingency and selection.

A molecule with self-copying ability must have a certain minimal complexity; in this case, the sequence complexity. It is this complexity that enables the molecule to fulfil a function, to do a specific work on its environment, the result of which is the dynamic maintenance of the molecule's *onticity*, the molecule's survival. To be so, the molecule must have a certain minimal knowledge of the relevant features of the environment, there must be a correspondence, however coarse-grained and abstract, between these features of the environment and the structure of the molecule. In general, at all levels of life, not just at the level of nucleic acid molecules, a complexity, which serves a specific function, and only that, corresponds to an *embodied knowledge*, translated into the constructions of a system. The environment is a rich set of potential niches: each niche is a problem to be solved, to survive in the niche means to solve the problem, and the solution is the embodied knowledge, an algorithm of how to act in order to survive.

Hence, life from its very beginning is a cognitive system: the self-copying molecule, pursuing its onticity in the world, accordingly, the simplest teleonomic system, is already a *subject* facing the world as an object. At all levels, from the simplest to the most complex, the overall construction of the subject, the embodiment of the achieved knowledge, represents its *epistemic complexity*. It is the epistemic complexity which continually increases in biological evolution, and also in cultural evolution, and gives the evolution its direction.

What forced the original self-replicating molecules to increase in complexity in the course of evolution, to associate with one another and with other molecular species, to produce higher levels of complexity, and, hence, other levels of embodied knowledge?

In SPIEGELMAN's experiments with self-copying of nucleic acid of the phage Q_β in an artificial system, evolution did not progress toward increasing complexity of nucleic acids, but toward its diminishing—the highest DARWINIAN fitness had the molecules which replicates at the highest rate and these molecules became shorter than the original one, with a lower algorithmic complexity (SPIEGELMAN 1971). This should occur in a simple and closed environment. Evolutionary reactors operates under steady state, with constant influx of substrates and ouflux of products, but also under simple and constant conditions. Under such conditions there is a selection for simple, rapidly replicating molecules (KÜPPERS 1979).

Different is the situation in the "evolutionary reactor" of the world. Replicating systems are present in a "vessel" of unlimited size, in an environment which is complex and steadily changing. A tiny change is enough to bring a rapidly replicating simple system into the environment which will its replication slow down or which will it destroy. Systems which are *accidentally* more complex and which would be, in an unchanging environment, eliminated, may gain advantage just because in their complexity a potential for "survival" under the changed conditions may have resided. It may be said that their greater complexity represents a more complex formal system and thus a larger epistemic capacity of the replicating molecule as a *subject*.

The higher is the rate of replication of a complex system, the higher is the consumption of resources, the higher energy dissipation in the evolutionary reactor, the larger a distance from thermodynamic equilibrium. Increasing the distance from thermodynamic equilibrium continues also when conservative structures, constructions, begin to be built up, embodying ever greater evolutionary knowledge. For maintaining them, energy dissipation is no longer required: there are kinetic barriers which keep their thermodynamic distance and retard their transition into equilibrium.

This enables biological evolution to be a continual growth of knowledge: creation of subjects with ever greater embodied knowledge, ever less probable, placed ever farther from thermodynamic equilibrium. Biological evolution is *inventive*, and this is *the* reason why it is *progressing*.

It should be made clear that algorithmic complexity of the sequence of units in unidimensional space may have approximately corresponded to complexity of the first replicating nucleic acids. When nucleic acids have subsequently produced a more complicated auxiliary devices, membranes, tissues, individuals, societies in order to make their onticity much more robust, complexity of the systems has assumed a form which cannot be expressed in such a simple manner. Complexity of nucleic acids themselves of the genome exceeds their sequential complexity: it is a complexity consisting in appropriate timing of gene transcription, implicating not only three dimensions of the space, but also the forth dimension of time (JACOB 1981, p89). It will be shown in another publication that this complexity is related to thermodynamic depth (LLOYD/PAGELS 1988). And that, in order to express it, object and subject should be considered as a unit, in inseparable interaction, as pointed out in a different context by GRASSBERGER (1989) and GELL-MANN (1994).

There is an intriguing relationship between epistemic complexity, fitness and truth. This will be analysed in another paper. An individual subject with high epistemic complexity exhibits robustness with respect to fluctuations in the environment. But also a large set of simple agents, such as a species existing in many identical copies, widely dispersed or closely collaborating, can be robust and survive under various attacks from the side of the environment, and even compete out complex subjects with much greater embodied knowledge, if the latter are sparse comparing to the former. It is this intricacy which serves as an argument to those which oppose the idea of progress in evolution. The intricate relationship is also obvious in cultural evolution. The saying that "a majority is always right" is far less trivial than it may seem. The relationship makes the epistemic "maze" more complicated but does not violate the universal evolutionary tendency.

Embodied knowledge enables teleonomic systems to proceed toward goals (underlain by the ultimate goal of onticity) by minimising the length of the trajectory. A virtually random walk at the beginning is being more and more biased in the course of evolution. In this sense, any teleonomic system is always rational, with the degree of rationality being determined by the difference between ideal trajectory, a straight-line (corresponding to the ideal rationality) and the biased zigzag trajectory allowed by the limited amount of the embodied knowledge.

The biological rationality has an essential shortcoming: it is always the rationality of RUSSELL's hen (a recurrent theme of RIEDL's epistemological analysis, e.g., RIEDL 1994). A hen, fed by a farmer, anticipates the future as a continuation of this benefice with no idea that this is just preparing her for a pan. Even though anticipating, she does not see the future, the anticipation is essentially an extrapolation from the past experience. The only exception has appeared at the level of rationality of the individual human person: even though with great difficulties, man can have a restricted foresight and make prognoses. But this may well turn to be just a tiny flash in the history of life on earth. Cultural evolution, with autonomous dynamics of memes, may have a rationality superior to that of an individual human person, but it appears to be again no more than the rationality of RUSSELL's hen.

It is so even in the case of science, a triumph of human rationality. As stated by POPPER, a scientific theory "can only prove its "fitness" to survive those tests which it *did* survive; just as in the case of an organism, "fitness", unfortunately, only means actual survival, and past performance in no way assure future success" (POPPER 1976, p103).

There is no reason of why self-copying molecules should not arise anywhere in the universe. On the basis of some reasonable assumptions from contemporary science, it can be easily calculated from DRAKE's equation (DRAKE 1990) that life must be a general phenomenon of the universe. The origin of life seems to be a nomic process, and only then teleonomic processes set in. As implied by the anthropic principle (GOTT 1993), our earth, and our species, do not occupy any unusual, or exceptional, position in space and time. The universe as a whole is *epistemically unfolding* by creating localised foci at which processes of knowledge accumulation are running ahead. The maze metaphor, used to explain progression of life on earth but also the failure of the vast majority of actors to succeed and their inevitable extinction, should apply to the entire universe.

The tendency toward the epistemic unfolding of the universe is named the epistemic principle.

Our earth, our species, we are actors in the unfolding. We shall keep in mind: we occupy no privileged position in the universe. We have no reason to suppose that we have been elected. We have many reasons to assume that this has not been the case.

5. The principle of ratchetting

As has been pointed out, dissipation of energy in evolution enables not just the maintenance of dissipative structures but also the formation of conservative structures, constructions. Constructions are systems far from thermodynamic equilibrium, separated from it mainly by kinetic barriers. Thermodynamics of constructions has not been worked out. Growth of knowledge in evolution means the accumulation of ever more complicated constructions.

Dynamics of the living systems, at all levels of hierarchy, consists in uni-directional ratchetting. The idea of a ratchet has been introduced into science by FEYNMAN, who used it to illustrate some implications of the second law of thermodynamics, in particular, that useful work cannot be extracted from equilibrium fluctuations (FEYNMAN et al. 1966a). A simple mechanical ratchet consists of *a wheel with assymetrically skewed teeth and a spring-loaded pawl*, which allows it to spin in one direction only and prevents backward motion. Molecular, evolutionary, developmental, cognitive, social ratchets are all based on this simple principle. At the molecular level, constructions allow life to use molecular ratchets: arrangements allowing to bias the BROWNIAN motion of particles in an anisotropic medium without thermal gradients, a net force, or a macroscopic electric field (MAGNASCO 1993, PESKIN et al. 1993, ASTUMIAN 1997). Random thermal motion of particles is rectified to serve a function, to do a useful work on the environment. A relation to the concept of rationality, as has been outlined above, is immediately apparent: random walk means “no rationality”, a straight line means an “ideal rationality” (but also causal connection, *nomicity*) and anything in between means a “bounded rationality”. Molecular ratchets are devices exhibiting *molecular rationality* and may be considered as the prototype of systems with “bounded rationality”.

Molecular ratchets have been proposed mainly to account for working of molecular motors, such as muscle proteins or ATP synthase. In essence, however, even simpler proteins may function as ratchets. This may apply to the basic cognitive devices, molecular receptors. BROWNIAN motion of a molecule of ligand is biased by the electric field of a receptor, ligand is being bound to receptor and a part of binding energy, instead of being dissipated straight away, is used to accomplish a molecular work by receptor: transfer of signal across the protein molecule. Ligand binding is coupled, *conjugated,* with signal transduction. This is an elementary form, at a molecular level, of a universal phenomenon of conjugation. Forms, numbers, and levels of conjugations have been increasing in evolution. Described in these terms, life in its entirety is but a huge system of countless conjugations through which the flow of energy, starting from the radiation energy of the sun, is canalised, forced to jump by steps, through a few selected degrees of freedom, down to the inevitable sink of thermal energy, instead of being disssipated downright. In this optics, even the most complex human activities, including science and art, are just evolutionary inserts into the flow and the eventual dissipation of the solar energy.

The incessant inventing of constructions is itself a process of ratchetting. Evolution as a whole is a ratchet. MULLER’s ratchet has well been known in the evolutionary biology: accumulation of deleterious mutations, resulting in an increase of the mutational load and an inexorable, ratchet-like, loss of the least mutated class (MAYNARD SMITH 1989, p241). MULLER’s ratchet is a virtual one, it may never operate, or perhaps, just exceptionally (ANDERSON/ HUGHES 1996): sex may have been an invention of how to prevent its operation. The preventive effect of sex may be amplified by outbreeding ratchets, which encompass various mechanisms to prevent inbreeding, including for instance *incest tabu* in humans.

An insight into the most elementary evolutionary ratchets has been provided by computer modelling of RNA evolution (FONTANA/SCHUSTER 1998). The probability of transition between two different RNA molecules, which differ from one another by a single mutation, is not symmetric: the destruction of a structural element through a single point mutation is easier than its creation.

The most effective evolutionary ratchet is made possible by WEISMANN’s barrier, separating genotype from phenotype. Thanks to WEISSMANN’s ratchet, the entire battlefield of an individual organism’s “Dasein”, with all its failures, disappointments, degeneration, senile resignation, is separated from the playground of the *evolutionary dicing,* which takes place at the genome level and, the case of parasitism disregarding, pushes the genome unidirectionally toward greater complexity. It is at this elementary, molecular level that is rooted the universal CAMPBELL’s “variation and selective retention” phenomenon (CAMPBELL 1974).

It has been pointed out that evolution pulls organisms ever farther from thermodynamic equilibrium. It has been aptly put by PRIGOGINE and his collaborators at a number of occasions (e.g., NICOLIS/ PRIGOGINE 1987) that matter far from equilibrium is creative. The larger dissipation of energy, the more powerful is self-organisation and the more order is being created in an irreversible manner. To appreciate the pioneering studies of PRIGOGINE and his insight, we may call this working of evolution PRIGOGINE’s ratchet.

Ratchetting plays a major role in development. By a simple analogy, the excellent discovery of George BEADLE from the early days of biochemical genetics, indicating that there is, generally, one gene coding for one enzyme—the "one gene-one enzyme" hypothesis—might be transferred to morphogenesis by inferring "one gene-one morphological trait". Notwithstanding the probability, that no scientist has ever attempted to make such a sweeping analogy, the assumption has long been a scapegoat of all brands of anti-NEODARWINISTS and biological structuralists. In order to specify the activity of each gene, in each cell of a multicellular organism, at each state of development, the program controlling this process should be enormous and, instead of being characterised by a great, but manageable, complexity the organism would need be endowed with some miraculous "supercomplexity". Perhaps even the ingenious metaphor of a cake recipe (DAWKINS 1987, p294), less demanding of program space, may still be too pretentious. Instead, a kind of somatic DARWINIAN mechanisms may operate (EDELMAN 1987, KUPIEC 1997, BRITTEN 1998), constrained to such an extent that it must end in a quasi-determined final state, but without requiring a too excessive program and with no need of a central controlling authority, a "genome brain" (or, for that matter, a "genome central committee"). At each step of morphogenesis, starting from two cells of the earliest embryo, various combinations of merely local associations, essentially stochastic or only slightly biased, may be occurring, but only the appropriate ones are developmentally retained by stage-specific developmental ratchets, which, by clicking round one notch, open a stage for new, higher-level associations. A complex global order is spontaneously emerging from exclusively local interactions of simple units.

This seems to be the universal principle of ordering in evolution, development, cognition and megasocieties's structuring. Again, the principle of minimal complexity suggests that molecular biology can make a major contribution towards its full elucidation. Nowadays, it may be receiving the strongest support from studies in artificial intelligence (MAES 1997): A complex behaviour emerges from the interactions of autonomous simple agents, situated in the environment, each of them assigned to fulfil a set of simple goals. There is no general planner, no internal structure corresponding to "the plan" of the system, no central representation shared by the agents.

Another principle is implicit in the principle of ratchetting. It may be called the principle of *hierarchical continuity of design* and loosely linked to JACOB's principle of tinkering (JACOB 1977). Evolutionary dicing combined with evolutionary ratchetting makes of evolution a process in which any new move must necessarily build upon the previous ones. Devices and constructions invented in the past are being remodelled into new ones and also serve to support the latter. Even if no longer functional, they are rarely thrown away but rather kept in store or disassembled and their parts used as in other combinations as modules for other purposes. This is also one of the reasons of increasing redundancy in evolution at all levels of biological organisation, from the redundancy of genes up to the majestic redundancy of human neurones, and, for that matter, of human culture. Increasing redundancy in evolution has been changing life *from game to play*. Adaptive behaviour has been complemented with *expressive behaviour*. Upon the onset of cultural evolution, the two forms of behaviour have been extended by an additional, *meme-enforced behaviour*.

Shaping and constraining evolution, the continuity of design is of fundamental importance in development. Much more strictly than in evolution, in development is any new move conditioned by all the previous moves. Because of continuous branching of developmental moves, previous moves recede deeper and deeper into the hierarchy. Since development is genetically controlled, the results of very early moves remain almost immutable. Any modification of the early genes of the embryonic stage, no matter how advantageous in itself, would presuppose a simultaneous accommodation of all the genes involved in later stages of ontogenesis—the probability of such a co-ordinated modification is virtually zero. There is a progressive and irreversible *encapsulation* of all previous achievements. This process has been named *generative entrenchment* by WIMSATT (1986). He has illustrated its virtual irreversibility at a model of a developmental lock: a digital cylindrical lock consisting of wheels, each with a number of possible positions, in which the correct position of a wheel is dependent on the actual position of a preceding wheel. The resetting of a wheel placed early in the process would incur the necessity to readjust simultaneously the correct positions of all subsequent wheels, but not in the other way round. Indeed, the digital developmental lock may be interpreted as representing a sort of developmental ratchet. A case in point demonstrating ratchetting, with both continuity of design and tinkering, is the gene *Pax6* (DESPLAN 1997). A gene, the initial function of which was to regulate photoreceptor differentiation in a primitive "eye" formed only of photoreceptors, has been promoted to the contemporary position of a master regulator in eye formation in flies, mice, and humans. It controlls the genes that were added later in evolution as the eye was becoming more complicated.

The same general principle underlies human cognitive ontogenesis. With *imprinting* at the bottom, through the very first filling in of the genetically determined abstract, but nevertheless specific, mind's "letterboxes" with concrete concepts, ideas and habits (*incipation),* through contingencies of reinforcement up to the conscious reflection. And, at all levels, variations, local interactions, selective retention by the imposed cognitive ratchets, new levels built up upon the unmoveable deeper levels, resulting in a unified, coherent and dynamic structure. Generally, the earlier in individual life has a cognitive module been assembled, the more resistant is it to any subsequent modification. It has been argued that HUSSERL's notion of *Lebenswelt*, taken over by existentialists, unique to every human being and apparently inaccessible to others, corresponds to the reality constructed in this idiosyncratic way in ontogeny of every individual (KOVÁČ 1992). By extension, specific human cultures arise and evolve on the basis of the same principle (KOVÁČ 1999).

Ratchetting in evolution, in development, and in cognition fulfils the same essential function: it allows step-wise accumulation and meaningful application of knowledge and prevents its futile diminution or degradation by running the process backwards. Ratchets operate at many hierarchical levels, from molecules up to megasocieties. The concept of *granulation,* analogous to the concept of graining of statistical physics, is instrumental in analysing these hierarchies. It will be dealt with in a separate paper.

6. The principle of minimal prejudice

"Von Anfang an muß das Leben ausgestattet gewesen sein mit allgemeinem Wissen, dem Wissen, das wir gewöhnlich Wissen von Naturgesetzen nennen. Selbstverständlich nicht Wissen in dem Sinne vom bewussten Wissen." (POPPER 1987, p32.) This is an unfortunate formulation. It may have been one of the reasons why POPPER supposed that the origin of life must have been an "unbelievably improbable" event. Life is *constrained* by *all* laws of nature, but this does not imply that, from its very beginning, it should know them. POPPER himself, like all evolutionary epistemologists, asserted that knowledge of an environment means *adaptation* to it ("…die Anpassung des Lebens an seine Umgebung ist eine Art von Erkenntnis"). This clearly implies that adaptation to a simple environment is equal to a simple knowledge and that, in the course of evolution, the continuous increase in complexity of niches which life can occupy means a continuous growth of knowledge—hence, the continuous increase in epistemic complexity.

Limited knowledge was a major cause of the extinction of species. The laws of nature have enforced themselves mercilessly when life did not know them. Ambiguous with regard to extinction has been another characteristic of living beings: inflexibility of beliefs, fanaticism. Organisms are *fanaticists*. Simple organisms with no capacity to learn, and even self-copying nucleic acid molecules, are absolute fanaticists. Organisms do not invent and maintain hypotheses, they abound in *beliefs,* and only some of these beliefs represent, in a specific environment, pieces of knowledge. Which implies that only some of the constructions of organisms are embodied knowledge, the others are but *embodied beliefs.* Once adapted to its environment, a simple organism remains totally inflexible. If we take a mutation in a bacterium as a new belief about the environment, we can say that the mutant would sacrifice its life to prove its fidelity to that belief. In organisms with learning capacity the situation is not as different as we may assume: as already mentioned, the principle of continuity of design takes care for maintaining and preserving in the course of individual life those beliefs and that behaviour that had been acquired early in ontogeny. From the point of view of the common gene pool, this has been nevertheless largely an adaptive arrangement: it is one of the main source for generating and maintaining polymorphism, for survival of a species in fluctuating environment and, by increasing variance, for enlarging its evolutionary potential. Under a specific fluctuation, some fanaticists would perish while others gain in fitness; upon a swing of the environmental pendulum into the opposite direction, the chances would turn round; but species, consisting of pure liberals, of pure sceptics, or of pure opportunists would not be robust enough to survive the fluctuations and would get extinct.

The human species has been no exception. Man, like all other animals, is not a POPPERIAN rationalist eager to expose his/her explanations of the world to testing and ready to replace them by new ones. Human beings are *mythophils:* they firmly stick to their beliefs, often ready to die for them just as sturdily as are ready the bacteria. The environment, in which our main mental dispositions have been shaped by selection, did not favour fitness of individuals who were irresolute, hesitating, tolerant, amazed at the complexity of the world and susceptible to cognitive chaos and existential anxiety. There must have been also a strong selection pressure for group conformity. Myths, unmoveable, indisputable, all-encompassing and omniscient explications of the world, have been most efficient group "glues", making of a group a powerful unit and exacerbating intergroup competition. They continue to fulfil this function in their contemporary form of ideologies. The "deficiencies of human reason" (RIEDL 1995), generating cognitive illusions and extremely biased prejudices, function as an excellent nutrient medium for exuberant growth and spreading of the memes constituting collective myths and ideologies. A particularly important cognitive illusion for reinforcing group cohesion is individual's *self-deception*: SOCRATES' maxim "Know yourself" must be much more difficult to achieve than anybody of us in our self-deception would admit.

It has been said repeatedly that many features of the physical and mental outfit of humans, selected for life in small nonanonymous groups of hunters and gatherers in the savannah, may no longer be adaptive in the socioworld created by cultural evolution. Experimental science, a unique invention of European culture, may provide a partial corrective. Not so much by conscious activities of scientists—an individual scientist may be no less a mythophil than is a layperson, he/she firmly sticks to his/her beliefs—but by its manner of how the world is being transformed into reality: experimental results allow no biases in rationality, they must conform to the laws of nature, underlain by the extremum principles. They impose upon the work of scientists a principle, which itself is an extremum principle, and which runs counter human "natural" mental disposition: the principle of minimal prejudice. It may be called JAYNES' principle, according to a physicist who first gave it a precise formulation. It has been anticipated by many philosophers and scientists by such ideas as OCCAM's razor (RUSSELL 1961, p462), economy of thought (MACH 1923), parsimony (SOBER 1992).

JAYNES has given the principle a mathematical formulation. According to him, if one has an incomplete knowledge of the subject, the minimally prejudiced assignment of probabilities is that which maximises SHANNON's entropy, subject to the given information (JAYNES 1957). The corollary of his argument has been the demonstration that the laws of thermodynamics can be derived as consequences of the principle. The thermodynamic entropy of CLAUSIUS becomes a special case of SHANNON's entropy if one asks the right question. It may not be too exaggerated to expect that foundations of some other disciplines of science may also be derived from JAYNES' principle.

There is an obvious link between the principle of minimal prejudice and the notion of rationality sketched above. And yet, there have often been reflections on rationality in which the principle of minimal prejudice has been violated. Rationality has been almost exclusively considered to be a matter of reasoning, of mental calculation, of conscious appreciation of profits and losses. A "*wisdom of the body*", achieved by evolutionary selection, has been ignored, a possibility of rational action has been denied other animals, rationality has been ascribed to the individual human person and nothing has been known of the superior rationality of adaptive dynamic systems. And, above all, rationality of emotions and their decisive role in meaningful behaviour has not been recognised. The belief in the power of the individual human reason has been a strongly biased prejudice, particularly in European culture. A prejudice that has no evolutionary justification.

The very principle of the minimal prejudice substantiates the next principle of cognitive biology.

7. The principle of minimisation of suffering

If bacteria had consciousness and were capable of self-reflection, their world view would be definitely "bacteriocentric". The same would hold for rats, except that the latters would construct reality that would be "ratocentric". The formers and the latters would be proud of their evolutionary prosperity and would scoff at man who has been led, by redundancy of the human brain and associated cultural evolution, toward such absurd evolutionary oddities as the atomic bomb or the mass television entertainment. We, human beings, can ask such theoretical questions as to what it is like to be a bat (NAGEL 1974), but will never penetrate into the bat-specific reality. We are confined to our human-specific reality and, by all our evolutionary and developmental ratchets, forced to be anthropocentric.

A bacterial philosopher, sentenced to life imprisonment in his/her species-specific formal system, would ask precisely the same most general questions as has been asking for two and a half thousand years the human philosopher: (1) Does the world exist at all; is it not but my illusion? (2) If something exists, why it does exist, why there is something rather than nothing? Just as his/her human colleague, the bacterial philosopher will find no answer. In order to find them he/she should jump out of his/her formal system—and it is impossible.

Evidence, the conclusive one, that the world exists, do we, humans, get in a different way. We get it through our *conscious experience of emotion*. By consciously feeling joy and pain.

Emotions have evolved as an efficient adaptive arrangement to secure onticity of living beings. Many attempts have recently been made to trace their evolutionary origin (PLUTCHIK 1991, DAMASIO 1994, WIMMER 1995), but, just as in the case of cognition, opinions diverge, with two extremes: one, considering emotions as a privilege of humans, and another, ascribing emotions even to non-living thermostats. The principle of minimal complexity should be useful in this analysis. It is tempting to search for emotions in any teleonomic system with built-in devices for evaluation of external stimuli (KOVÁČ 1982).

A peculiarity, and possibly the uniqueness, of the species *Homo sapiens* is the coincidence of emotions and self-consciousness. It is this parallelism, or rather inseparability, of these two evolutionary achievements that renders possible double human transcendence—jumping out of the formal system of our reasoning, and surmounting the fundamental imperative of life: struggle for individual onticity, and nothing but individual onticity, at any price! As will be argued later, both self-consciousness and the capacity for deep emotions, indeed, hyperemotionality, may be end products of *a singular run-away process* in human evolution, so that they widely exceed adaptive qualification. Emotions are a most powerful motor of expressive behaviour.

In spite of successful attempts at explaining human altruism by the elegant proofs that account for altruism in all biological species, including plants and micro-organisms, we know from introspection that human altruism has a specific, most efficient source: *empathy*. Empathy does not only conduct our behaviour toward another human being; it also provides the most pervasive evidence of his/her existence (BUBER 1923). It is much stronger that would be Bayesian reasoning which is also used to justify the anthropic principle: none of us occupies any special position in the universe. If we combine the former and the latter arguments, a conclusion, voiced already almost three decades ago by Linus PAULING (1970), is inevitable:

"The evidence of my senses tells me that I am a man, like other men. When I cut myself I am hurt, I suffer, I cry out. I see that when some other person cuts himself he cries out. I conclude from his behaviour that he is suffering in the same way that I was. None of my observations leads me to believe that there is something special about me that sets me apart from other human beings, in any fundamental way; instead, I am led to believe that I am a man, like other men. I want to be free of suffering to the greatest extent possible. I want to live a happy and useful life, a satisfying life. I want other people to help me to be happy, to help to keep my suffering to a minimum. It is accordingly my duty to help them to be happy, to strive to prevent suffering to other people. By this argument I am led to a fundamental ethical principle: the decisions among alternative courses of action should be made in such ways as to minimise the predicted amounts of human suffering [...]

I have contended that the principle of the minimisation of human suffering is a scientific principle, with a logical, scientific basis. I do not disagree with Professor Jacques MONOD, who said that ethics must be based on axioms, just as geometry is based on axioms. Professor WADDINGTON then pointed out that, although different geometries may be developed on the basis of different axioms, all people agree that in the practical world we can accept Euclidean geometry and its axioms. I feel that, although we have theoretical freedom allowing various ethical systems to be formulated, the choice of a reasonable and practical ethical system is highly restricted by our knowledge about the nature of the physical and biological world, and that the only acceptable ethical systems are those that are essentially equivalent to that based upon the principle of the minimisation of human suffering."

Any ethical norm can be maintained in a population by two different manners. First, it may be increasing fitness of the individuals who observe the norm. In the simplest case this would be due to reciprocal altruism—a sort of calculation, which, indeed, may be implicit in such universal moral commandments, as is "do not do to your neighbour what you don't wish he/she would do to you". Or it may be a strongly virulent meme, which spreads in the population even though it may reduce individual fitness of the infected human carriers. The principle of the minimisation of suffering, based on empathy, combines both biological and cultural contributions. Empathy as a biological feature would be one of those abstract and specific mind's "letterboxes", mentioned above, that can be filled in either by compassion, charity and self-sacrifice, or by envy, vengeance or malicious cruelty. Neither of the two opposite kinds of behaviour has ever been observed in non-human animals, which proves that they are made possible by a faculty exclusive to humans. It depends mainly on the composition of the meme pool of the specific cultural environments which of the two kinds of behaviour will predominate.

Nature is indifferent to suffering. A part of human suffering is a consequence of deliberate human action, but a larger one is unintended, caused by ignorance and by human incapacity in the face of blind forces of nature and society. Ignorance, impotency, inferiority, fear have constituted the nutrient medium for memes of envy and cruelty, their opposites have favoured countenance, compassion, devotion, social playfulness. For centuries, science has been explained and justified as an activity aiming at reducing the ignorance and the incapacity. All the principles of cognitive biology substantiate the concept of science as an organised reduction of ignorance. At the same time, the string of the presented arguments associates science more directly with the reduction of suffering. The "search for truth" has been often presented as an internal norm of science. It is not: science with lies is simply no science. In the same vein, the principle of minimisation of suffering gives science an additional dimension. Not as a norm: the more "genuine" science is, the closer it is to this extremum principle. While PAULING's reasoning ended with a normative proposal of a basis for an ethical system, this statement is purely descriptive. The origin of science, and its subsequent evolution as an institution, have been inherently linked with the reduction of human worries: pain, distress, labour, misery, anxiety. Science has become the main instrument in human efforts to minimise pain and to maximise pleasure. Cognitive biology just explains why it is so.

This is not to say that a research in which suffering, unintended or intended, is incurred, is no science. It is a science with a large proportion of ignorance. As life on earth, as life in the universe, science itself progresses forward in a maze: there is a major evolutionary tendency, but there are also many false paths and deadlocks. The success is not prescribed.

Discussion

Is there any need for the new term of cognitive biology? Is cognitive biology part of evolutionary epistemology or evolutionary epistemology itself under a new guise? The opposite may hold: cognitive biology is a larger set, with evolutionary epistemology as a subset. Cognitive biology has grown out of molecular biology, with an assumption that the elucidation of molecular recognition, of processing of molecular signals, of the organisation of gene networks, of protein computation may provide a clue for understanding higher cognitive processes. At the same time, its close association with physics and chemistry may help to end the perennial controversies and confusions concerning the relations between information and physical entropy and to enable a more precise and formalised description of knowledge, epistemic complexity and rationality.

Some of the most prominent physicists have anticipated the fundamental importance of extremum principles of physics for our comprehension of nature (PLANCK 1958, FEYNMAN et al. 1966b, LANDAU/LIFSHITZ 1969). In the sense of logical parity, the same principles may somehow govern the working of mind. The fact that some of the principles of cognitive biology may resemble, or be related, to the extremum principles of physics, may not be a fortuitous coincidence. Duality of ontology and epistemology may turn out to be a major cognitive illusion.

It has been pointed out that Thomas KUHN's theory of scientific revolutions has not had a good influence on cognitive science. Many cognitive scientists present their theories as new KUHNIAN paradigms by discrediting others (TAATGEN 1999). Cognitive biology is no paradigm shift. It builds upon the traditional views of a number of scientific disciplines, its only virtue being an attempt at synthesis. Even if not clear at first sight, it should be emphasised that it is mainly out of the deep sources of physics and molecular biology that the idea of the pivotal role of human conscious emotional experience emerges. Implying consequences which, however, may be a revolutionary flash in biological evolution: it is no longer onticity of selfish genes, permanence of self-replicating entities, but a suffering of the individual conscious person that has become—even if possibly just for a short period, and perhaps not for the first time in the history of the universe—the plot of the world drama.

It should be made clear that none of the principles of cognitive biology, not even the principle of the minimisation of suffering, are normative. They are descriptive statements derived from its axioms. It would be premature to try to axiomatise cognitive biology. Some basic postulates, idealised as are the postulates of Euclidean geometry, may provide sufficient proof of the descriptive character of the principles. Existence, onticity, of human suffering is a distinct initial postulate. The second postulate is empathy as another emotional quality. Minimisation of one's own suffering is, in the ideal case of unlimited empathy, inseparable from the minimisation of suffering of one's neighbour. Science as a specific human invention is instrumental in this action. Neither individual and inclusive fitness, nor reciprocal altruism take part in the reasoning. To claim that this minimisation principle is normative would be equivalent to the claim that extremum principles of physics, such as the principle of minimal action, are normative.

This, of course, does not imply that scientists are driven by compassion and a conscious effort to help or be useful to humankind. The tendency of science to reduce human suffering comes out of its inherent dynamics, in which public acceptance and support of science, as well as of technoscience, play a major role. The trend toward a full elimination of suffering, a triune result of run-away processes of human hyperemotionality, consciousness and meme dynamics (which includes the emergence of science), transcends human biology, including human DARWINIAN fitness: suffering has been an adaptive device in striving for existence and zero suffering (combined with maximum of artificial pleasure which may soon be provided by techniques of virtual reality) may well reduce the striving to zero. But need survival continue to be the supreme value of that kind of conscious life that has evolved on Earth?

The latter point is crucial for a propre understanding of the scope of cognitive biology. It would be misleading to conceive it just as an attempt at laying down a molecular foundation of cognition. The principle of ratchetting (in addition to the epistemic principle) explains how the unidirectional operation of evolutionary and developmental ratchets generates everincreasing complexity, culminating in human conscious emotionality and in science. Cognitive biology encompasses the analysis of human transcendence, tracing it back and down to its biological, and molecular, roots.

In addition to its status as a science, cognitive biology is also a conceptual program, and as a program it has normative features. The program has its substantiation in the conviction that may have been first voiced by Claude LÉVY-STRAUSS: the 21st century will be the century of science on man—or will not be. The main statement of the program reads: The aim of science *should be* the minimisation of human suffering *plus* the optimisation of human material and spiritual comfort. (This statement, in contrast to the previous description of the tendency, is normative.) Cognitive biology as a program holds to the tenet that the appropriate way to achieve this, at the present state of our knowledge, is the study of any specific trait of human nature by using organisms, or even purely molecular systems, which are of minimal complexity and still exhibit that trait. The introductory motto of this paper may be slightly modified (and expressed with a slightly normative flavour): The most efficient approach to mankind, to human suffering and happiness, may still be for some time the study of the bacterium.

Acknowledgements

I thank my colleagues at the Konrad-LORENZ-Institut für Evolutions- und Kognitionsforschung for an excellent atmosphere in which this work could have been accomplished and for critical reading of the manuscript. This work was partially supported by grants from the Howard HUGHES Medical Institute (No. 75195-547301), and the Austrian Bundesministerium für Wissenschaft, Forschung und Kunst (No. GZ 45. 425/1-VIII/B/3a/97).

Author's address

Konrad Lorenz Istitut für Evolutions- und Kognitionsforschung, Adolf-Lorenz-Gasse 2, A-3422 Altenberg/Donau; and Centre for Cognitive Biology, Faculty of Science, Mlynská dolina CH-1, 842 15 Bratislava, Slovakia. Email: kovacl@fns. uniba. sk

References

Anderson, D. I./Hughes, D. (1996) Muller's ratchet decreases fitness of a DNA-based microbe. Proceedings of the National Academy of Sciences of the US 93: 906–907.

Astumian, R. D. (1997) Thermodynamics and kinetics of a Brownian motor. Science 276: 917–922.

Bargmann, C. I. (1998) Neurobiology of the Caenorhabditis elegans genome. Science 282: 2028–2033.

Bohr, N. (1964) Atomphysik und menschliche Erkenntnis, vol I. Vieweg: Braunschweig.

Britten, R. J. (1998) Underlying assumptions of developmental models. Proceedings of the National Academy of Sciences of the US 95: 9372–9377.

Buber, M. (1923) Ich und Du. Insel Verlag: Leipzig.

Butler, D. (1998) Advances in neuroscience "may threaten human rights". Nature 391: 316.

Campbell, D. T. (1974) Evolutionary epistemology. In: Schilpp, P. A. (ed) The philosophy of Karl Popper. Open Court Publishing Co: La Salle, pp. 413–463.

C. Elegans Sequencing Consortium (1998) Genome sequence of the nematode C. elegans. A platform for investigating biology. Science 282: 2012–2018.

Chaitin, G. J. (1975) Randomness and mathematical proof. Scientific Americain 232 (5): 47–52.

Chaitin, G. J. (1990) Algorithmic information theory. Cam-

bridge University Press: Cambridge.
Dawkins, R. (1987) The blind watchmaker. Norton: New York.
Desplan, C. (1997) Eye development: governed by a dictator or a junta? Cell 91: 861–864.
Drake, F. D. (1990) The Drake equation: A reprisal. Byron Press.
Edelman, G. M. (1987) Neural darwinism. Basic Books: New York.
Eigen, M./Winkler, R. (1975) Das Spiel. Piper: München, Zürich.
Feynman, R. P./Leighton, R. B./Sands, M. (1966a) The Feynman lectures in physics. Vol. II. Addison-Wesley: Reading.
Feynman, R. P./Leighton, R. B./Sands, M. (1966b) The Feynman lectures in physics. Vol. III. Addison-Wesley: Reading.
Fontana, W./Wagner, G./Buss, L. W. (1997) Beyond digital naturalism. In: Langton, C. G. (ed) Artificial Life. MIT Press: Cambridge, MA, pp211–227.
Fontana, W./Schuster, P. (1998) Continuity in evolution: On the nature of transitions. Science 280: 1451–1455.
Gell-Mann, M. (1994) The quark and the jaguar. Freeman: New York.
Georgescu-Roegen, H. (1971) The entropy law and the economic process. Harvard University Press: Cambridge.
Gerlach, J. (1988) Gödel-Theorem and Kuhlenbeck-Paradox. Naturwissenschaften 75: 393–398.
Goodwin, B. (1976) Analytical physiology of cells and developing organisms. Academic Press: London.
Gott III, J. R. (1993) Implications of the Copernican principle for our future prospects. Nature 363: 315–319.
Grassberger, P. (1989) Problems in quantifying self-generated complexity. Helvetica Physica Acta 62: 489–511.
Heschl, A. (1998) Das intelligente Genom. Springer: Berlin
Hofstadter, D. A. (1980) Gödel, Escher, Bach: An eternal golden braid. Penguin Books: Harmondsworth.
Husserl, E. (1936) Die Krisis in der europäischen Wissenschaften und die transzedentale Phänomenologie. Philosophia, Band 1. First published as a book in 1954 by Nijhoff: Hague.
Jacob, F. (1977) Evolution and tinkering. Science 196: 1161–1167.
Jacob, F. (1981) Le jeu des possibles. Fayard: Paris.
Jaynes, E. T. (1957) Information theory and statistical mechanics. Physical Review 106: 620–630.
Kolgomorov, A. N. (1965) Three approaches to the quantitative definition of information. Problemy peredachi informatsii 1(1): 3–11.
Koshland Jr., D. E. (1977) A response regulator model in a simple sensory system. Science 196: 1055–1063.
Kováč, L. (1982) Evolution of affective evaluation of external stimuli. In: Novák, V. J. A./Mlíkovská, J. (eds) Evolution and environment. Czechoslovak Academy of Sciences: Prague, pp867–874.
Kováč, L. (1986a) Úvod do kognitívnej biológie. (In Slovak. With an abstract in English: Introduction to cognitive biology.) Biol. listy 51: 172–190.
Kováč, L. (1986b) The future of bioenergetics. EBEC Reports 4: 26–27.
Kováč, L. (1987) Overview: Bioenergetics between chemistry, genetics and physics. Current Topics in Bioenergetics 15: 331–372.
Kováč, L. (1992) The natural world ("Lebenswelt") as a biological problem. Vesmír 71: 505–506.
Kováč, L. (1999) European culture in the global conflict of cultures: a view of a biologist. In: Fukač, J. (ed) The crossroads of European culture. Vutium Press: Brno.
Kuhlenbeck, H. (1982) The human brain and its universe. Karger: Basel.
Kuhn, H. (1972) Selbsorganisation molekularer Systeme und die Evolution des genetischen Apparats. Angewandte Chemie 18: 838–862.
Kuhn, H. (1988) Origin of life and physics: Diversified macrostructure—Inducement to form information-carrying and knowledge-accumulating systems. IBM Journal of Research and Development 32(1): 37–46.
Kupiec, J. J. (1997) A Darwinian theory for the origin of cellular differentiation. Molecular and General Genetics 255: 201–208.
Küppers, B. (1979) Towards an experimental analysis of molecular self-organization and precellular Darwinian evolution. Naturwissenschaften 66: 228–243.
Landau, L. D./Lifshitz, J. M. (1969) Kratkiy kurs teoreticheskoy fiziki. Vol. 1. Nauka: Moscow.
Lloyd, S./Pagels, H. (1988) Complexity as thermodynamic depth. Annals of Physics 188: 186–213.
Lorenz, K. (1943) Die angeborenen Formen möglicher Erfahrung. Z. Tierpsychologie 5: 235–409.
Lorenz, K. (1973) Die Rückseite des Spiegels. Piper: München.
Lorenz, K. (1983a) Der Abbau des Menschlichen. Piper: München.
Lorenz, K. (1983b) Kants Lehre vom Apriorischen im Lichte gegenwertiger Biologie. In: Lorenz, K./Wuketits, F. M. (eds) Die Evolution des Denkens. Piper: München, Zürich, pp95–124. First published in 1941 in Blätter für deutsche Philosophie 15: 94–125.
Mach, E. (1923) Die ökonomische Natur der physikalischen Forschung. In: Populärwissenschaftliche Vorlesungen. Barth: Leipzig, pp217–244.
Maes, P. (1997) Modelling adaptive autonomous agents. In: Langton, C. G. (ed) Artificial life. MIT Press: Cambridge MA.
Magnasco, M. O. (1993) Forced thermal ratchets. Physical Review Letter 10: 1477–1481.
Maynard Smith, J. (1989) Evolutionary genetics. Oxford University Press: Oxford.
Nagel, E./Newman, J. R. (1958) Gödel's proof. Routledge and Kegan Paul: London.
Nagel, T. (1974) What it is like to be a bat. The Philosophical Review 83: 435–451.
Nicolis, G./Prigogine, I. (1987) Die Erforschung des Komplexen. Piper: München.
Pauling, L. (1970) Scientists in Politics. In: Tiselius, A./Nilsson, S. (eds) The place of values in a world of facts. Almquist and Wiksell: Stockholm, Willey, New York, London, Sidney, pp97–05.
Penrose, R. (1989) The Emperor's new mind. Oxford University Press: Oxford.
Penrose, R. (1994) Shadows of the mind. Oxford University Press: New York.
Peskin, C. P./Garrett M. O./Oster G. F. (1993) Cellular motions and thermal fluctuations: The Brownian ratchet. Biophysical Journal 65: 316–324.
Piaget, J. (1967) Biologie et connaissance. Gallimard: Paris
Planck, M. (1958) Das Prinzip der kleinsten Wirkung. In: Vorträge und Reden. Vieweg: Braunschweig, pp91–101.
Plutchik, R. (1991) Emotions and evolution. International Review of Studies on Emotion 1: 37–58.
Popper, K. (1957) The logic of scientific discovery. Hutchinson: London. First published in Vienna in 1934 under the title Logik der Forschung.
Popper, K. (1987) Die erkenntnistheoretische Position der evolutionären Erkenntnistheorie. In: Riedl, R./Wuketits, F.

W. (eds) Die Evolutionäre Erkenntnistheorie. Parey: Berlin, Hamburg, pp29–37.

Prigogine, I. (1967) Introduction to thermodynamics of irreversible processes. Wiley: New York.

Radnitzky G./Bartley III, W. W. (eds) (1987) Evolutionary epistemology, rationality, and the sociology of knowledge. Open Court: La Salle.

Riedl, R. (1979) Biologie der Erkenntnis. Parey: Berlin.

Riedl, R. (1992) Wahrheit und Wahrscheinlichkeit. Parey: Berlin.

Riedl, R. (1994) Darwin, Zeus und Russells Huhn. Kremayr & Scheriau: Wien

Riedl, R. (1995) Deficiencies of adaptation in human reason; a constructivist extension of evolutionary epistemology. Evol. Cogn. 1: 27–37.

Riedl, R./Delpos, M. (eds) (1996) Die Evolutionäre Erkenntnistheorie im Spiegel der Wissenschaften. WUV Universitätsverlag: Wien.

Russell, B. (1961) A history of western philosophy. Unwin Paperbacks: London.

Schuster, P./Sigmund K. (1983) Replicator dynamics. Journal of Theoretical Biology 100: 533–538.

Simon, H. A. (1983) Reason in human affairs. Stanford University Press: Stanford CA.

Sober, E. (1992) Parsimony. In: Keller, E. F./Lloyd, E. A. (eds) Keywords in evolutionary biology. Harvard University Press: Cambridge MA, pp249–254.

Spiegelman, S. (1971) An approach to the experimental analysis of precellular evolution. Quart. Rev. Biophysics 4: 213.

Stent, G. S. (1963) Molecular biology of bacterial viruses. Freeman: San Francisco, London.

Taatgen, N. (1999) The atomic components of thought. Trends in Cognitive Science 3: 82.

Tarski, A. (1949) The semantic conception of truth and the foundations of semantics. In: Feigl, H./Sellars, W. (eds) Readings in philosophical analysis. Appleton-Century-Crofts: New York, pp52–84.

Tautz, D. (1998) Debatable homologies. Nature 395: 17–18.

Vollmer, G. (1983) Mesokosmos und objektive Erkenntnis – Über Probleme, die von der evolutionären Erkenntnistheorie gelöst werden. In: Lorenz, K./Wuketits, F. M. (eds) Die Evolution des Denkens. Piper: München, Zürich, pp29–91.

Vollmer, G. (1985) Was können wir wissen? Band 1. Die Natur der Erkenntnis. Hirzel: Stuttgart.

Vollmer, G. (1987) Evolutionäre Erkenntnistheorie. Hirzel: Stuttgart. (First published in 1975.)

Wimmer, M. (1995) Evolutionary roots of emotions. Evolution and Cognition 1: 38–50.

Wimsatt, W. C. (1986) Developmental constraints, generative entrenchment, and the innate-acquired distinction. In: Bechtel, W. (ed) Integrating scientific disciplines. Nijhoff: Dordrecht.

Wray, G. A./Abouheif, E. (1998) When is homology not homology? Current Opinion in Genetic Development 8: 675–680.

Werner J. Patzelt

Institutions as Knowledge-Gaining Systems

What Can Social Scientists Learn from Evolutionary Epistemology?[1]

Abstract

What can institutional research and the comparative analysis of political systems learn from evolutionary epistemology (EE)? A five-step argument will illustrate it. First of all, it is necessary to describe the basic ideas of EE. Then I will show how we should follow the ideas of EE in conceiving social and political systems as knowledge-gaining systems. Subsequently, I will explain how we can understand 'institutionality' as a result of learning and 'historicity' as a learning process. Finally, I will put everything in concrete terms by using an outline of a comparative analysis of political systems in evolutionary–morphological perspective as an example.

Key words

Evolution, evolutionary epistemology, sociology, ethnomethodology, history, institution(s), comparative systems research, political science, learning, knowledge, reality construction.

Central Ideas of Evolutionary Epistemology

One of the founders of evolutionary epistemology has been the ethologist Konrad LORENZ (VOLLMER 1983; WUKETITS 1990; RIEDL/WUKETITS 1987; RIEDL/DELOPOS 1996). His earliest relevant publication in this field is an article originally published in 1941 about "KANTs Lehre vom Apriorischen im Lichte gegenwärtiger Biologie" ("KANT's doctrine of the a priori in the light of contemporary biology"). His most famous publication in this context "Die Rückseite des Spiegels. Versuch einer Naturgeschichte menschlichen Erkennens" (LORENZ 1973, engl. version "Behind the Mirror" 1977). The Viennese zoologist and marine biologist Rupert RIEDL has elaborated and developed further this empirical epistemology based on the theory of evolution. In 1980, he published "Biologie der Erkenntnis", subtitled "Die stammesgeschichtlichen Grundlagen der Vernunft" (RIEDL 1980, engl. version "Biology of Knowledge. The Evolutionary Basis of Reason" 1984; cf. also RIEDL 1975, 1985, 1986).

The leading idea of all this work is to conceive organisms as systems that are gaining knowledge under the pressure of evolutionary adaptation. It is thus proceeding from the hardly disputable conception that all life is based upon exchange processes between an organism and its environment, the latter being called "reality existing 'outside'". Those features of an organism which are important for such exchange processes, have to correspond to the characteristics of the organism's environment. Only then, the exchange processes can be carried out stably. The proponents of EE call this 'Passung'. Obviously, only those organisms, whose genotype and phenotype contained such 'knowledge' about the relevant properties of their environment, were fit to survive. Konrad LORENZ called the process of getting such knowledge *information*, since it is a matter of 'in-forming' (the Latin term for 'engraving upon') the relevant and important features of its environment to an organism. Being 'in-formed' by mutation and selection this way, the design of an organism reflects those environmental features that are most important to it. In this sense, the tail fin of a fish is a reflection of the water in which he swims; the organ of equilibrium is reflecting gravity; and the eye corresponds to the light radiated by the sun.

Organisms reached a new evolutionary level, when two things happened. First, relevant environmental characteristics were no longer 'reflected' in an organism's physical structure alone, but in their central nervous system as well. Second, the development of the central nervous system opened up three additional opportunities:

■ Information about characteristics of the environment can be stored as a kind of *memory*.
■ The memory can distinguish environmental conditions in a way that the organism reacts selectively with—genetically fixed—adequate behavioral repertoires.
■ The memory stores information about environmental conditions, and an organism can subsequently *recall* this information voluntarily or even consciously. Likewise, the organism can *imagine* patterns of behavior before it activates them.

This last step is decisive for human cognitive capabilities. The reason is that, at this point, the reality existing 'outside' can be reflected—even consciously and intentionally—'inside' an organism as well. The storing of knowledge in libraries and the simulating, experimenting, or playful–creative dealing with such knowledge, is just a cultural continuation of this biological chain of evolution. In this context, the term "world picture apparatus" ("Weltbild-Apparat") becomes important. It was introduced by Konrad LORENZ and re-defined by Rupert RIEDL, who later also used the term 'ratiomorphic apparatus' when referring to human beings. Both terms conclusively characterize the interaction of the following processes: Information about the environment reaches the organism via its *sense organs*; once the information is inside the organism, the *central nervous system* transforms it—if necessary—into behavior and into consciousness; and sense organs and the central nervous system *together* make sure that the living system reflects the reality existing 'outside'.

Of course, 'world picture apparatuses' have not evolved 'per se'. They evolved only in relation to the *specific* exchange processes of a particular species. As a consequence, the reality existing 'outside' is not depicted 'per se' or 'entirely', but only in certain spectra and partially, and—after all—only good enough to secure the survival of the particular species. Therefore, 'world picture apparatuses' *will* provide true representations of the reality existing 'outside'; but these representations will always be *selective*, since they are affected by *perspective particularities* according to the organism's need of information. For example, it is for good reasons that the hunter in the wood is 'seen' differently by a tick than by his dog or friends.

The construction plans of both species-specific 'world picture apparatuses' and of the creatures themselves (to which the 'world picture apparatuses' deliver correct, selective and perspective reflections of the reality existing 'outside'), are fixed genetically. Each organism develops in the course of its own genesis—the ontogeny—the 'world picture apparatus' which evolved phylogenetically. As a consequence, any individual's 'world picture apparatus' highly determines the way of perception and cognition open to this individual. We can even say that it gradually 'teaches' an organism (1) which elements of the reality existing 'outside' are available to it, and (2) which environmental characteristics are important for it. Thus, prior to any *individual* experience, there is an overall *phylogenetic* experience. This phylogenetic experience leads to adaptation of the species to its environment and, thus, increases the chances of survival of any individual organism. The individual creature therefore never starts 'at zero' with the cognitive task to secure its survival. Quite on the contrary, the 'world picture apparatus' provides an individual organism with an 'inborn teacher' who *always* indicates to an organism *beforehand* what to notice and how to interpret and treat it. Hence any individual *a priori* is a phylogenetic *a posteriori*. This exactly was the topic of the KANT essay by Konrad LORENZ.

However, these 'inborn teachers', or 'world picture apparatuses', only deliver correct and reliable information in relation to those exchange processes between an organism and its environment which continue to proceed as they did at the time when the correspondence between the organism and its environment came into being and was genetically fixed. The 'inborn teachers' do *not* provide the organism with any experiences about anything that developed at a later period within the species' environment. Everything of this kind is *not* 'covered' by a genetically secured correspondence. Therefore, individuals and entire species can fail to master exchange problems with their environment. In this case, they will be extinguished by environmental changes in connection with adaptation processes that are simply too slow.

But human beings succeeded in conquering new living spaces and in creating novel environments in quite short periods of time. The reason is that they became, thanks to their superior central nervous system, 'universal' to the extent that they developed civilizational techniques allowing them to engage in, or even to provoke, such 'gaps of phylogenetic experience'. They survived wherever and as long as they subsequently managed to bridge these gaps on an ontogenetic and individual base *for all practical purposes*. However, by generating a cultural environment that changes at much faster rates than the natural environment to which we are adapted, very specific problems of our species arise. The famous title

"Urmensch und Spätkultur" (engl. version "Primitive Man and Late Culture") by Arnold GEHLEN is indeed appropriate for an analysis of these problems (GEHLEN 1986, 1993).

Social and Political Systems as Knowledge–Gaining Systems

The structure of the intersection between the evolutionary–epistemological approach to phylogenetic learning and the social–scientific approach to institutional learning can easily be detected. We merely need to call 'organisms' living systems or even social systems; we simply speak of 'information structure', 'system-intern information management' and 'ideology' instead of 'world picture apparatus'; and the concept of 'inborn teachers' will be applied to socialization agencies of all kinds. By accomplishing this conceptual and theoretical task of translation, we will quickly leave behind the realm of EE proper and enter the field of 'normal' sociology. There, we immediately recognize that social and political systems are (potentially) knowledge-gaining systems as well. In addition, we recognize that they gain knowledge in exactly the way described by EE for (higher) biological systems.

Economic enterprises, for instance, obviously are systems that either succeed with, or just fail in, the attempt to gain necessary knowledge. Via the supply of goods and services on the market as well as by profits—realized through the satisfaction of demand—they are coupled back with their environment. Unless prevented by subsidies, they will escape collapse only if they *correctly* recognize the demand structures in their environment and react *quickly* to the fluctuations of the market or to the competitive efforts of other suppliers. Apparently, we are dealing with the same mechanism that we encountered when we were looking at evolutionary selection pressure. Just like biological systems, a social system incorporates knowledge about characteristics of its environment by fashioning its own (information-processing *and* proactive) structures in a way that they can secure its maintenance in this given environment. Thus, it is a knowledge-gaining system which can become, in a process of evolution, an ancestor of further systems.

Of course, not only economic enterprises depend on a sufficient gain of knowledge about their environment in order to survive.[2] The same is true for the political system and its sub-systems. Just a few examples will make this plausible. A system of government is characterized by a particular division of labor between head of state, executive branch of government, and parliament. Such a system may stand different crises, develop further, or break down in critical situations—like that of the Weimar Republic or of the IV[th] French Republic. And political parties gain voters due to a political program fitting the electorate's wishes, or they use electoral defeat for a restructuring of their personnel and politics, or they will just become marginalized.

On balance, *political* systems possess at least *two* environments with which they have to deal and which they must 'in-form' themselves in the sense of the EE. One of them is the governed society itself. The other one is this society's environment, which spans from the society's ecological base to the surrounding system of other societies and their states. As a matter of fact, the political system must adapt to *both* environments if it is to remain stable. From them, it must correctly gather, and process, the information necessary for its steering tasks. At the same time, the political system must be capable of self-reproduction by political socialization and recruitment. And it must succeed in winning generally binding recognition for its established rules and actual decisions. Throughout this process, political regulations and decisions must both secure their own success and avoid to endanger the correspondence between system and environment. Of course, there is no guarantee for that to happen. There is no guarantee either that good adaptation of a political system to its societal environment will go along with good adaptation to that society's natural or trans-national environment. Much rather, history gives numerous examples of configurations of social and political systems that indeed fitted each other, but no longer fitted their suddenly or gradually changed international or natural environment. They therefore either vanished or experienced periods of substantial crises.

In order to get a handle on these phenomena, it is useful to treat political structures and the political systems built from them as knowledge-gaining systems. If the necessity arises, they may know how to reconstruct and reform their framework of institutions and organizations, of offices and agencies, of norms and practices. Thus, they remain functioning in spite of changes within and outside society. The conversion of existing structures may serve this purpose as well as modifications in the functions performed by the very same structures. And the disassembly of traditional structures can serve this goal as well as building upon them by their functional or structural differentiation. Since the High Middle

Ages, the history of the British constitution has demonstrated how such a learning process of a political system can take place without major crises—with only few exceptions, such as the War of the Roses, the Civil War, and the Glorious Revolution. On the other hand, the constitutional history of France since the Great Revolution has taught, how a society can experiment by trial and error with quickly changing political structures, until it finally finds a political order which is (at least for a longer period) adequate for both the ruled society and its problems. Such an order was established in the systems of the III[rd] and V[th] republic.

These examples also demonstrate that the effects of a change of the system's environment and the system's performance of adaptation are closely intertwined. Of course, the degree of such interweaving is much higher in political systems than at the level of 'world picture apparatuses', of organisms, or of simpler social systems like middle-class enterprises. In the realm of political systems, the complex network of their sub-systems' feedback loops time and again leads to chaotic system behavior. Therefore, it is much more difficult to see and demonstrate in a political system (rather than in a biological or simple social system) the functioning of mechanisms which are basically the same like in organisms and economic enterprises. Additionally, there are two reasons why political systems can, and often do, protect themselves against a restructuring gain of knowledge, at least for some time. As a consequence, political systems are knowledge-gaining systems only in principle and for a certain period of time, but not in every single case.

The *first* reason has to do with the state's capacity to make choices simply because it has the power to do so. The central point is its *de facto* monopoly of (legitimate) coercive power. Because of that, governments are capable of monopolizing material resources to a great extent. Raising taxes and duties, governmental control of economy, or full socialization of means of production are means towards this end. The employment of substantial resources appropriated in this way will then, at least for some time, work as a buffer between the political system and its (double) environment.[3] The exploitation of man and nature along with a corresponding allocation of resources, later to be missed in other areas, can maintain expensive and nonproductive policies for decades. Other examples include state-controlled maintenance of steel-plants in regions that are poor in resources and energy, or state-controlled subsidies that make disappear the need for sectional or structural changes within an economy. Or governments may suppress opposition of a poorly governed population with a great repressive apparatus; or powerful political parties try to eliminate upcoming competitors through the establishment of cartels and the monopolization of financial and communication means. The following formula summarizes such occurrences: Whoever possesses power may mobilize resources, and resources may make free from incentives, or even pressure, for adaptation. Therefore, Karl W. DEUTSCH was absolutely correct when he defined power as the 'capacity of not having to learn' (DEUTSCH 1973). Since the political system is that social subsystem which usually administers the monopoly of (legitimate) coercive power, it has *typically* the privilege of *not* having to adapt to its environment for certain periods. On the contrary, political systems usually make attempts to adapt rather their environments to their own characteristics, goals and ideologies. Totalitarian dictatorship is the most remarkable type of such voluntaristic reality construction (PATZELT 1998).

However, the attempts of political systems to adapt rather their environment to themselves than trying adaptation the other way round, will sooner or later expose these systems to the danger of destruction or breakdown. The reason is that usually major discrepancies between system and environment arise. This is clearly shown by the breakdown of state socialism and by the dead-end road in which industrial society found itself by the 70ies for ecological reasons. So it is only *for some time* that the monopolization of power and the buffering employment of resources exempt a political, economic or social system from immediate reality control and need for adaptation. During those times, the fundamental mechanism forcing social structures either to gain knowledge or to be extinguished, is merely overlain, but not put out of force. This is why 'regulating catastrophes' destroy systems unwilling to adapt time and again, and lead to *forced* learning by its surviving parts. Such 'regulating catastrophes' include the breakdown of political systems, the removal of a government through a revolution, or the devastating electoral defeat of a political party.

The *second* feature of political systems leading to fundamental learning problems is the role played by political traditions, convictions, and ideologies in the planning, execution, and legitimization of political action. Within a political system, all this certainly produces identity and continuity. It also lends necessary stability to political thinking and action. Especially the processes of institutionalization, and institutions as their results, rely on traditions and ideologies. However, the outcome is not only stability. The result is also reduced search for alternatives, hesitant use of creative coincidences, and declining responsiveness to environmental impacts. Europe's communist states gave a discouraging example for this incapability of political systems to engage in apparent necessary structural adaptation to environmental challenges because of ideological fixation. Their finally forced learning could no longer prevent the breakdown of the system, most dramatically in the cases of the Soviet Union and the GDR. But even open societies[4] (POPPER 1947) with functioning pluralism, such as the Western democracies, have great problems to eliminate flaws of adaptation to their changing environment through structural changes. Because of deep-rooted traditions, convictions and ideologies, such problems may occur, for instance, in the ecological sector or in the distributive policies of a social welfare state that overstrains its economic basis. When both ideological conservatism and buffering employment of monopolized resources come together, it seems to be really inevitable for political systems to become unable to learn and to risk 'regulating catastrophes' after a certain period.

These 'regulating catastrophes' often take shape as system- or regime breakdown. History is full of such examples. They show that most political systems could not reproduce themselves after a certain period and, therefore, break down. They perished or were displaced like a no longer adapted species, unless there was an 'ecological niche'[5] allowing survival without adaptation to general environmental changes. MARX' terminology of a 'dialectic of basis and superstructure' described this phenomenon already more than a century ago, although with far less theoretical range than EE. Only few political systems, for example the British, succeeded in changing fundamentally without system-destroying 'regulating catastrophes'.

Considering such evolutionary processes, be they successful or not, makes two kinds of political structures especially interesting. On the one hand, these include institutions that outlived in single systems in spite of all system breakdowns and of multiple changes of their surrounding systems. On the other hand, those political structures and institutions are of utmost importance, that spread over to other political systems and fulfilled their (new) functions successfully in even widely differing structural contexts. Such institutions seem to be *inherently capable of learning*.

Institutionality as a Result of Learning

Our analysis has ascended from the micro-analytical 'lowlands' of the relationships between an organism and its environment to the macro-analytical 'plateau' of comparing the adaptation of whole political systems to their surrounding systems. In the course of this analysis, and always from the perspective of EE, we have traversed the 'layers' of social reality (RIEDL 1985, pp66–80; PATZELT 1993a pp40–43) from humankind's genetically fixed ethogram to culture-specific stocks of knowledge, and from roles and role structures, from organizations and institutions, to whole political systems, international regimes, and trans-national structures. But so far we have passed over the important phenomena of the intermediate or 'meso'–level. We have not yet discussed the processes of institutionalization, nor institutionality as a specific aggregate state of social reality, nor institutions themselves.

Institutionality and institutions, however, are particularly interesting in the perspective of EE. The reason is that they represent the counterpart of those morphological structures of fauna and flora whose similarities and relations have drawn the attention of evolutionary research all along. Evolutionists compare and investigate the structure of skeletal parts or inner organs; they study their development and transformation. They are also interested in the functions fulfilled by the skeletal parts and inner organs for the overall system of an organism. Quite close to such questions, institutional research deals with the following topics:

- In view of which environmental conditions do actions solidify into roles and role structures?
- Under which environmental conditions, and using which cultural techniques, do role structures transform themselves into institutions?
- What connections are there between types of environmental conditions and types of institutionalization processes on the one hand, and institutions on the other hand?
- How do types of institutions become components of self-evident everyday culture to such an extent that their constructive features serve as a basis of further institutionalization processes, and become a crystallization nucleus of morphologically related institutions?
- For which types of institutions can we detect morphological relationships? And why, in single cases, is there exactly *this* filiation of a type of institutions?

■ Which types of institutions are definitively results of learning by social systems under the selective pressure of evolutionary adaptation? *What exactly* did those systems learn? And what does, hence, the morphologically anchored information, i.e., the 'intelligence coagulated in institutions', really 'tell'?

Quite obviously, these questions are in line with the common questions of evolutionary research and EE as discussed above. One must not take analogies as equivalences; but it would be just as wrong *not* to use their heuristic potential. Since evolutionary research and EE have acquired and applied relevant insights, procedures and heuristic algorithms quite successfully—especially in the field of biological evolution—, the comparative investigation of so similar processes and results of social institutionalization would disregard this work only to our own disadvantage. So we should make use of analytical spill-over. Because EE has shown that morphological structures are *results* of learning processes, we should also investigate *institutions* as results of—more or less successful—learning processes in the course of social evolution. By the same token, we should analyze institutionalization processes as learning processes of social structures.

Adequate analytical tools for such an enterprise—other than the ones provided by EE—are made available by research on the construction of social reality (cf. BERGER/LUCKMANN 1971; HOLZNER 1972; GIDDENS 1984; PATZELT 1987). The guiding idea of such research is to proceed from the everyday knowledge of actors and then to examine how such everyday knowledge gives shape to actions in everyday life. These actions, in turn, create or maintain social roles and institutions. As long as this process goes on without (major) disturbances, social reality is (re-)constructed. And as long as this is case, social reality is a self-evident feature of everyday life and a matter of course for all everyday life activities. Seen in this way, everything influencing everyday life knowledge and everyday actions in daily life is an act of reality construction.

Interactionist theories of social action describe in detail how such actions of reality construction take place. First of all, complexity–reducing 'schematisms of interaction' (LUHMANN 1979) and patterns of habitualized actions arise. From this, roles and sets of roles (i.e., organizations) develop (BERGER/LUCKMANN 1971). Sets of roles may become institutions by symbolic self-representation, i.e., if the principles giving shape to those actions are symbolized as irrefutable interpretive schemes for both the actors and their audience.[6] Around roles and sets of roles, and within organizations and institutions, additional and more complex processes of habitualization, role formation and institutionalization will take place. In this way, a 'process of reflexivity' is going on,[7] a set of feedback loops in which man-made structures react on their very constructors. So the results of reality construction, both by their very existence and by their interpretation, become a framework of further reality construction. In this framework of meaning and interpretation, of expectation and actions (as it is maintained by these reflexive processes of reality construction), culture-specific conceptions of 'normality', 'rationality', 'competence', 'expectability' etc. are available. As members of a specific culture, we rely on these conceptions as matters of self-evidence, thereby reproducing role structures, organizations and institutions, and using these very results of reality construction as resources of further reality construction at the same time (Hans BUCHHEIM accordingly called institutions "standardized social dispositions").

What was called 'background fulfillment' (in German 'Hintergrunderfüllung') and described as an existence–relieving basic process of institution maintenance by the social anthropologist Arnold GEHLEN, is, in a very similar formulation, termed 'routinely non-discrediting of background expectancies' within the ethnomethodological theory of reality construction (GARFINKEL 1967/1984; PATZELT 1987). In both theoretical approaches, reflexive processes of interpretation and interaction are analyzed as they create, reproduce, and maintain (or undermine, transform or destroy) those 'matters of course' that give shape to human actions. These matters of course include cultural coding systems and stocks of knowledge as well as schematisms of interaction and roles, organizations and institutions. Ethnomethodology calls all of this an 'occasioned corpus', i.e., a stock of elements out of which the 'reality existing outside' is made, and classifiable within Karl POPPER's ontology along his concept of 'three worlds' (POPPER 1984). Of course, this 'corpus' requires continuous 're-occasioning' in order to remain available as something taken for granted by every competent 'cultural colleague'. But as long as this stock of elements is available, its very presence limits—and thus gives shape—to such behavior that we consider as 'normal', 'acceptable', 'expectable', etc. Fulfilling this molding function, the 'occasioned corpus' becomes itself a central resource of further reality construction (see PATZELT 1987). Part of it are the creation and maintenance of institutions.

But which types of institutions may allow to overcome—within what environmental conditions—which kinds of problems of human acting in a way that secures a given society's existence?[8] Only comparative analyses–especially including historical and inter-cultural comparisons—will provide reliable answers to this question. But even at their starting point it is obvious what we have to consider *at least*. These issues include the basic human competencies of action and interpretation, as set up in the human ethogram; culture-specific repertoires of acting and interpreting; (collective) stocks of knowledge and interpretative schemes as well as their social, regional, temporal and other theoretically relevant distributions; cultural-specific background expectancies of normality and rationality (PATZELT 1987); efforts to transmit such repertoires of acting and interpreting or stocks of knowledge in the chain of generations; and the roles, role structures, organizations, and institutions under specific analysis themselves.

Since social and political reality is—of course selectively and in biased perspectives—present in the everyday practical knowledge of humans as their constructors, we can speak—with Peter BERGER—of 'society in man' (BERGER 1963). Inversely, 'man in society' acts and reproduces, or changes and destroys, his own social and political reality. He does so by using those 'matters of course' (in the shape of which social reality is present) as resources for practical actions.[9] The results of these processes, such as confirmed, shaken, or dwindling matters of course, then in turn become a resource of further reality construction, thereby acting reflexively upon everything that (re-)creates them. In this context, the following insight is central. On the one hand, humans are the subjects of reality construction, since they create, maintain, change, or destroy 'their' institutions and social environment. On the other hand, humans are turned into 'subjects' in any social and cultural sense only by their surrounding social environment and by the institutions available to them (cf. Technische Universität Dresden 1996, p15). Personal subject formation—a contingent process—thus occurs by dealing with, and being shaped by, existing institutions. Although these institutions are contingent structures themselves, they are much less contingent than those individuals on which they have effects.

This double insight creates another interface of institutional research with EE. It can easily be seen that the relationship between humans and their reality–constructing stocks of knowledge on the one hand, and those institutions created and maintained by them on the other hand, is fully analogous to that nexus which exists on the level of an organism between its genotype and its phenotype. The phenotype relates the organism to its environment and represents the morphological result of learning by phylogenetic interaction between the species and its surrounding reality. The genotype, however, contains that selected material which allows the reproduction of the phenotype during ontogeny. It is handed down as long as the phenotype reproduced succeeds to survive in its environment or ecological niche. So there are some obvious parallels between morphological phenotypes and institutions, between genotypes and reality–constructing stocks of knowledge. These parallels have lead some theorists, interested in Greek-inspired plays on words, to call those stocks of knowledge—quite analogous to *genes*—simply '*mnems*' or '*memes*' (from the Greek expression *mneme*, i.e., recollection, memory). In this sense, memes are elements that become imprinting factors of culture-specific reality construction respectively institutionalization. They do so by way of combined action of cultural and individual memory. And whereas biochemical codes transmit genetic information, mnemic information is transmitted by verbal or symbolic codes. Both transmitters of 'in-formation'—this term being understood in the sense of EE—act as functional equivalents in their respective fields of application. We can draw this parallel even much further. The phenotype provides each species with stability in its own environmental contacts. But what does provide human actions time and again with such great stability (in the sense of uniformity, expectability, and regularity), given the availability of so many freely eligible alternatives for any single act? According to the usual answer, there are "mechanisms and structures that stabilize and maintain arrangements of order, by means of their symbolic self-representation", and they "create probability, repeatability, and expectability of social events" (cf. Technische Universität Dresden 1996, p8). Obviously, the symbolic codes that enable such 'symbolic self-representation' (and make it reality–constructing) resemble even at first sight those memes just mentioned.

'Institutions' is the common term for all these mechanisms and structures that are stabilizing themselves (and their surrounding social order) by symbolic self-representation. Gerhard GÖHLER, for instance, formulated the following definition: "Institutions are … social setups that give a certain solidity, continuance, and repeatability to the behavior and acting of individuals, that is to say: they constitute, and are, a restriction of options." (GÖHLER 1988, p16). Morphological structures do exactly the same for the biomass of an organism. And we do not leave the intellectual world of EE either if we formulate with Arnold GEHLEN: Institutions serve basically the 'relief of human existence', since man is a universal and risked being that succeeds only by use of institutions to reduce the openness of his acting and interpreting horizons to such an extent that he can cope with the remaining complexity. Therefore, we should attribute "institutionality to man as a necessity for the securing of his existence", (see Technische Universität Dresden 1996, p13) i.e., recognize his ability of social–structural learning as a precondition of his phylogenetic reproduction.

Obviously, we can look at all of this equally in the perspective of institution theory, from the standpoint of sociological theories of reality construction, or from the angle of evolutionary research and EE. Our analysis will always come to the same conclusion: Order always proves to be something that only comes into existence in the course of reality–constructing processes, since all structures of order—among them their reflexively stabilizing organizations and institutions—are nothing else but (more or less fragile) products of processes. This is why Harold GARFINKEL, the founder of ethnomethodology, correctly called social and political reality on its micro-level an "ongoing accomplishment of the concerted activities of daily life" (GARFINKEL 1984, VII). The basis of this accomplishment are the resources of reality construction mentioned before: the human ethogram; culture-specific stocks of knowledge; techniques of accounting, interpretive procedures and scenic practices competently mastered by 'cultural colleagues'; ethno-specific schematisms of interaction; and roles, sets of roles, and the organizations or institutions created and maintained by all of that.

Indeed even institutions are something 'firm' and 'solid' only as long as the processes reproducing and maintaining them take place without major disturbances. Earlier institution research tended to neglect this fact. It is, however, very well reflected in the following formulation: "What appears to us as an 'institution', is always a respective, preliminary result of processes". The 'point' of institutionalization processes, and basically of *all* reality–constructing processes for that matter, merely is that institutions, by means of self-symbolization, "conceal their own process character. They design images of unshattered stability and relate to constructions of an unbroken history or to foundation myths and traditions, that just do not reflect dynamics, but use categories of 'continuity'. Institutional analysis, therefore, has to understand the static of institutional configurations as a result of self-stabilization" (see Technische Universität Dresden 1996, pp18f). Ethnomethodological analysis of reality construction does nothing else when it turns the—at least in the 'natural attitude of everyday life'—"'uninteresting' essential reflexivity" of reality–constructing actions (as Harold GARFINKEL put it) into to its central topic.[10]

Additionally incorporating EE into such analytical frameworks and explanatory theories, will widen the scope of the analysis of institutions by the investigation of reflexive institutionalization processes and by the study of institutional learning. This, in return, will supply the key for an understanding why some arrangements of social order are more stable than others, and why certain forms of social and political reality may be stabilized and reproduced more easily than others within exactly the same environmental conditions.

Historicity as a learning process

Institutions give shape and continuity to social action. Possibilities and forms of acting 'coagulate' wherever institutionality comes into being. Whenever this happens, society adopts a 'more strengthened', sometimes even a widely consolidated 'aggregate state'. Its stability–producing institutions obviously restrict degrees of freedom for future development; they curtail certain possibilities of action and thus promote others; they initiate path-dependent processes; and they obstruct or, inversely, open up horizons of planning and acting. At least, they influence cost/benefit calculations or reflections about the relative soundness of alternative acting options. Just by doing so, institutions *bundle up* actions and sharpen their directions and forms, while otherwise much acting would remain amorphous. In this way, i*nstitutions give shape to unfolding history*. But since they are *products* of processes as well, they remain subject to history and to change themselves. This can be formulated as follows: Institutionality is an "aggregate state of social acting and communication structures between change and continuance", i.e., between "intended stability and actual change" (Technische Universität Dresden 1996, p8).

Again we move on a ground fully familiar to evolutionary research and EE. After all, we are dealing with questions of how structures form and establish themselves within environments that are partly set and partly changing; how these structures gain an existence of their own and thus may have retroactive effects upon their environment; and how it is possible to maintain a 'dynamic correspondence' (or perhaps even a mutual stabilization) of structure and environment, given both continuous exchange processes between structure and environment, and the need for periodical reproduction of *both* structure and environment. Institutional analysis and EE even share the question of how structures may fade away after a period of time—either as a disappearing species, or as an institution lost in history. In both branches of such *historical* analysis, the study of the history of a morphological structure reveals a 'protocol of systemic learning', with learning being understood as a process of 'in-formation' of relevant environmental features into a structure whose 'obstinate' acting subsequently has retroactive effects upon the environment having formed it. EE is indeed nothing else but a theory of both such historical learning processes and of the development of structures that result from them. For EE, it does not matter if it is about 'regular' morphological structures of an organism, or about the special case of the 'world-picture apparatus' or 'ratiomorphic apparatus' of a higher developed being.

Proceeding from this insight, one can—by way of heuristic formation of analogies—introduce some further innovations into the analysis of the nexus between institutions and history, or between institutionality and historicity. Morphological evolutionary research very carefully distinguishes long-term stable structures from short-range differentiations resting on them, or from either adaptive specialization or individual deviation. From such a perspective it becomes clear at first sight that we have to combine institutional analysis with a concept of 'layers of reality' and of specific 'time structures' typical of each of them. Within the science of history, such a concept of time structures was introduced by Fernand BRAUDEL (1976).

According to him, history unfolds, on the one hand, as a short-term *histoire événementielle*, a history of daily events. This history of events and facts, as well as all the individual interpretations and considerations of the actors creating this history, is essentially formed by institutions, be they social, economic or political institutions. Some institutions, such as parliaments facing elections every few years, have rather short endogenous temporal rhythms. Thus, they lead to a short-winded *histoire événementielle* themselves. Other institutions, like property systems as *institutions–choses* or governmental systems as *institutions–personnes*, function on much longer-term rhythms or even remain stable for a medium period of time. Their own rhythms, or the upcoming, acting, and disappearing of such institutions themselves, subsequently lead to *conjonctures*, a term with which BRAUDEL refers to medium-range time structures. And finally there is the *longue durée*, which means long-range developments in history. It is a purely empirical question to what extent any given *longue durée* may be a consequence of long-term existing institutions, or rather the result of geographic, climatic, demographic or other 'basic facts'. It is also an empirical question in which way such long-range natural processes may interact with basic human institutions, thereby setting the stage for institutions that give shape to the medium-range processes of the *conjoncture*. Even just the formulation of these questions makes clear how close we are again to the questions put forth by EE. And as soon as we think about plausible answers to these questions, we can easily see: Whatever gives shape as a structure of *longue durée* to the structures of any *conjoncture*, is an equivalent to those phylogenetic construction plans which establish the insurmountable prerequisites for each ontogenetic formation of an individual. These long-range phenomena correspond to those reality–constructing resources, by which 'society in man' demonstrates to 'man in society' his limits of any evolutionary stable, i.e., sustainable action.

In this context, the central question considering the nexus between institutionality and historicity is: Which structures do really have a chance to become generated as structures of order producing stability even in their surroundings, and reducing degrees of freedom in otherwise purely contingent processes? What, after all, implements structures into history and forms an evolutionary process that may be understood by teleonomical (in no way teleological!) considerations? Here we are at the very core of evolutionary theory and EE, whose topic is almost defined by this question. By pursuing this question with sufficient obstinacy, one will even get beyond the goal aimed at in the Dresden research project on 'institutionality and historicity' (cf. Technische Universität Dresden 1996). For there, we only find the question about the "nexus between change and 'continuance' of cultural objectivations, (i.e., institutions) their historic forms, and social structures".[11] However, one can also try to find the very prerequisites for the emergence of specific types of institutions. And we can turn a historian's fundamental question about the nexus between change and continuance, or about discontinuity and continuity, into a social scientist's fundamental question of how—as shown by history itself—very firm macrostructures arise from such micro-processes of reality–constructive actions, which—although in principle open for many paths—prove in practice to be highly selective and pre-formed.

There is even one more approach to the nexus between institutions and history that is opened up by crossing institutional analysis, studies of social reality construction, and EE. It is about how human cultural history itself became, through institutionalization and institutional selection, a history of learning. Then, historiography can try to write down accounts of this history of learning that are inspired by EE. A brief sketch may suffice to illuminate this point.

Let us start with understanding institutionality and institutions within the model of 'layers of social reality' discussed above. Then it becomes clear, that exactly by means of institutionality and forming of institutions, the evolution of human social behavior reaches beyond the 'sociality of kinship groups' which is presumably fixed genetically at the level of our ethogram. If we dare to use a much too simplistic formulation, we can even say: It is exactly institutions that transform the genetically fixed 'deep structure' of human existence into the 'surface structure' of that social reality which is perceived and used in everyday life; and it is nothing else but institutions that bridge the gap between the micro-level of relations in the human life-world and the macro-level of national or global structures. Institutions are, therefore, meso-structures that are condemned, from part of both their constituent and super-structures, to either *learning* or *perishing*. Hence, their analysis needs to be undertaken under two aspects: On the one hand in terms of their constituents on the 'lower' levels of social reality, such as the human ethogram, cultural-specific stocks of knowledge, and concrete individual persons; and on the other hand in terms of those supra-systems on higher layers of social reality into which institutions are embedded and for which they fulfill manifest or latent functions. Of course, the linkages between all three 'layers' have to be treated as a set of dynamic, interactive structures, in which institutionality only then coagulates to an institution, if—and as long as—a correspondence of that 'institution in being' with its environment in both the lower *and* higher layers in the strata of social reality can be realized and secured.

As is widely known, approaches based on the theory of evolution tend to be misunderstood, whether on purpose or for other reasons. Hence it is in order to clarify in *what* sense 'historicity' is conceived here. In no way, there is any 'teleological' meaning combined with this concept of historicity, nor any allusion to 'laws of history' (For a critical position cf. POPPER 1974). But this concept of historicity is always about *teleonomy*, i.e., about the question which institutional mechanisms and structures processes shape or even promote historic change and perseverance.[12] And it is evolutionary research, having done away so thoroughly with the idea that the creation of man might be the goal or end of natural history, that opens up the appropriate perspective on these processes. In this perspective, we see that there is really no firm goal in history, but much rather a lot of dead-ends and broken-off developments. But at the same time we do recognize clear structural conditions which favor development in one direction rather than a possible alternative development in an other direction, and which make—at least temporarily—one kind of morphological structures prevail over its alternative.

Therefore, it seems evident how we have to connect institutional analysis to an evolutionary analysis not only of natural, but also of social and political history. Its focus has to be the *interaction between institutions and social or political reality*. If formulated merely in this abstract manner, such a statement will only lead to a shrug of approval. However, if one thinks of the historicity of *concrete* institutions (especially institutions making up political systems), the importance—even for actual political problems—of such questions and investigations becomes obvious. The reason is that political systems—and their institutions—are central factors giving shape to society, economy, and culture. Therefore, their creation, change, and mode of acting has deep impacts on those areas of human history which, in turn, are most important for the way we see ourselves and think of competent actions. Quite undoubtedly, there are great differences between various political systems and institutions in relation to their capability to solve problems, with respect to their 'resistance' in view of dangers, and regarding their power of self-preservation. So the central question reads as follows: Which types of political institutions form history, and how? And which types of institutions are, according to the criteria mentioned before, superior to others? At this point of the argument, some remarks are in order on how such a research program can be put in concrete terms and be used for the comparative analysis of political systems.

An Example: Comparative Analysis of Political Systems in Evolutionary–Morphological Perspective

Why do single elements of political systems recur in completely different cultures, such as hereditary political positions, or a politically relevant priesthood? Why are, once brought into being, several elements of political systems—such as parliaments, political parties, and labor unions—incorporated into entirely different political systems over and over again? Why do certain configurations of components of political systems prove to be more stable than others under changing environmental conditions? And what can we learn from answers to these questions about various possibilities for the design and the development of political systems (PATZELT 1995)?

Such questions immediately link the comparative analysis of political systems to an objective so successfully carried out by botany and zoology. Their goal was to come up with a typology based on morphological considerations that might allow to discover, or even reconstruct, lines of evolution. Within the scope of such a 'genetic typology',[13] we should understand the limitation of degrees of freedom for further system development (taking place through the commitment by a species or type to a certain basic construction plan), and we should understand as well why the fulgurant[14] combination of components that have previously been separated opens up degrees of freedom for entirely new structures and new types of systems. Such developments can, and should, be understood as processes of systemic learning. Then, a 'genetic typology' may be read as a *protocol of learning* by systems that adapt to environmental changes or penetrate into new ecological niches.

But how can we really use this idea for the comparative analysis of political systems? For this purpose, we should trace in historical–intercultural comparisons, how (a) certain fundamental system elements come into being under particular environmental conditions, and how (b) these elements then determine the logic of functioning of the respective type of system in such a way that either prerequisites are opened up, or certain degrees of freedom are withdrawn, for future system development.[15] Then we should ask how (c) different variants of the same type of political system develop within such a basic construction plan under specific environmental conditions; how (d) these variants, or some of their components, either develop further under the influence of certain environmental impacts, or are specialized to a degree that they (e) can only endure in ecological niches, or (f) are condemned to vanish in their evolutionary dead-ends through 'regulating catastrophes'. Furthermore, we should study the presence of single system elements that survive even a change of the system (g), and the penetration of single system elements into new system environments either (h) through export of institutions or (i) through imitation. In the latter case, we would be back at our first task of analysis (a).

As examples for case (a), the arising of basic system elements that determine future developments, one could refer to the formation of the medieval feudal state or to the Greek polis. Given the political conditions under which such a basic type can really work, certain directions of future development can hardly be pursued any longer. For instance, the allocation of political functions by casting lots, as practiced in the Attic polis, can hardly be implemented, or be accepted, in a feudal system (case b). After the emergence of an at least periodically convening deliberative assembly, such as estate conventions in Europe since the Middle Ages, a constitutional, later even parliamentary democracy may arise within a political system that is ruled by a monarch. This happened in England. However, the emergence of deliberative assemblies may also lead to the formation of a dualism between crown and estates, as it happened in Germany. And the development can also take to an absolute monarchy, making estates powerless by renouncing their summoning, as was the case in France. These are three instances for case c. In England, the European feudal state has developed, with only a few dramatic crises, into the 'motherland of parliamentary democracy' since the High Middle Ages (case d). But the feudal state could also endure until today, as a constitutional monarchy, in 'political niches' like Liechtenstein and Luxembourg (case e). It disappeared in France through frequent changes of regime between the revolution of 1789 and the first decade of the Third Republic. In Germany, it was destroyed only through the 'regulating catastrophe' of the 1918 revolution (case f).

Political parties have developed in nearly all political systems where it was necessary to organize majorities in assemblies and to present candidates for elections. Once 'invented', parties remained present, at least in the form of a single hegemonial party, in all political systems where they ever came into being (case g). And together with labor unions, parliaments, bureaus of census and judicial systems, parties have been export articles of European and American imperialism in the 20th century (case h). Additionally, component parts of Western democracy have been implemented, or at least superficially imitated, by many states of the earth during the last five decades (case i). By such imitation, or by preceding imperialistic export of institutions, new starting points for institutional development were created in Latin America, Africa, and Asia (case a). In these instances, the course of system evolution will prove whether such structures will develop that fit to the new societal environments (cases b, c, d), or whether 'regulating catastrophes' will have to create space for the formation of new structures (case f). Also these new structures may fail in the short run, such as 'stone-age communism' in Cambodia, or they may prove to be insufficiently adapted and evolutionary unstable in the long run, as was the case in Eastern European state socialism. But they may be successful as well, West Germany's political institutions after 1949 being an outstanding example.

An evolutionary–morphological comparative analysis of political systems going beyond that sketch is nothing more than an EE-inspired and newly accentuated perspective on the classical topics of the comparative investigation of political systems. But just this change of perspective is valuable, since it imparts new insights and promises substantial theory transfer between various disciplines. By following the outlined tracks, and by carrying out a systematic secondary analysis of the ample literature of anthropology, legal history, general history, sociology and political science about the manifold historical and contemporary political systems, it should be possible to get to a genetic typology of political systems within a sound period of time. Such a genetic typology, comprising present and historical types of political order, is desirable all the more as it will serve as an instrument for the identification of basic patterns and regularities in the evolution of political systems. This goal can be reached only on the basis of a large number of very different cases of compared systems. However, such a large number will be comparable for theory–constructive purposes only by categories making up a sound typology; and such a typology will disclose its entire utility only by making comprehensible the *structurally anchored gains of the learning processes* of political institutions in the course of evolution.

Once created, such a genetic typology would shed light, on the broadest possible basis, on evolutionary stable and unstable elements of political systems and political system configurations. It would make clear degrees of freedom that may be either expected, or are simply not available, within a certain basic construction of a political system, or because of its reflection in the political tradition and culture of a given society. Such a genetic typology would further explain the possibilities to implement learning capacities already in the very structures of political systems. Additionally it would inform us about the ways, and the conditions, of dissemination processes of political structures. All of this information could be of great practical importance, since it is always necessary to reform constitutions, or to react to 'regulating catastrophes' by formation of new systems. And since it is a permanent task of political leadership to provide societies that change under varying environmental conditions with reliably adaptive political structures, political science—as a practical science in its ARISTOTELIAN tradition—should seek to provide politics with useful knowledge for this never ending task. But unfortunately, there is so far no discussion in political science about such possible uses of EE. Only within sociology, in the 'adaptive function' of social systems canonized in Talcott PARSONS' AGIL-scheme, there is a well-established theorem whose affinity to the central ideas of EE cannot be missed (cf. KÄSLER 1974, pp55ff; PATZELT 1993a). Sociology and political science, however, lack any impulse to link such interfaces to EE (cf. MEYER 1996)

This is deplorable all the more as it has been an ancient concern of political science to relate political structures to their biological foundations and to understand them on such a basis. About 350 years ago, Thomas HOBBES started his systematic exposition of a theory of political order in his books 'De Homine' and 'Leviathan' with a series of chapters like 'De generis humani origine' (Of the origins of mankind), 'De linea visuali et perceptione motus' (Of seeing and the perception of motion), 'De affectibus, sive perturbationibus animi' (Of emotions), or 'Of Sense', 'Of Imagination', 'Of the Consequence or Train of Imaginations', 'Of the Interior Beginnings of Voluntary Motions, commonly called the Passions; and the Speeches by which they are expressed'', 'Of the Virtues, commonly called Intellectual, and their contrary Defects'. HOBBES carried out these investigations for the

Author's address

Werner J. Patzelt, Dept. of Political Science, Technische Universität Dresden, D-01062 Dresden, Germany. Email: werner_j.patzelt@mailbox.tu-dresden.de

only purpose to create a firm basis for his political science argument. He expected to find this base in a correct theory of the biological and cognitive nature of man. Of course, he could undertake this enterprise only *more geometrico*, in his time's rationalistic way of thinking. But we can now pursue the same goal based upon the empirical findings embedded in EE. So it seems that political science, with the research program outlined here, is much nearer to its traditions and classics than it appeared at first sight.

Notes

1 This contribution is derived from work done in a major interdisciplinary research project on the nexus between 'institutionality' and 'historicity', carried out at the Dresden University of Technology within the Sonderforschungsbereich (SFB) 537 "Institutionalität und Geschichtlichkeit". Therefore, several references will be made to this research project and its guiding theoretical ideas (Technische Universität Dresden 1996).

2 Here and in the following line of arguments, 'environment' has to be understood in the sense of social systems theory. It means an often stable, but basically dynamic and time and again in part chaotic framework of social structures (above all organizations and institutions), technical structures (e.g., traffic- and communication networks), and natural structures (e.g., geographical and demographic facts). Even far-reaching interactions can occur between an environment conceptualized in this way and a social system. In their course, a social system needs not be restricted to just 'in-form' upon itself the respective environment, but it can—depending on the case—more or less take part in shaping, or even creating, this environment.

3 An analogy can be observed in large-scale and multinational enterprises. They are very similar to political systems with respect to their 'resistance' to environmental impacts.

4 POPPER argues in favor of the democracy—and against all forms of totalitarianism which base their systematic claims of authority upon ‚scientific grounds'—referring to the greater, since secured by institutions, *learning capability* of democracies. His argument is convincing at the starting point; it should however not block the look at the structural conservatism of saturated electorates.

5 For political systems, their 'ecological niches' are traditionally described by social–cultural, technical, geopolitical, economic, and other power-related factors.

6 See Technische Universität Dresden (1996), pp12–19.

7 'Reflexivity' refers here to re-acting effects of processes, mechanisms, and structures of action or of interpretation. This meaning must not be confounded with reflexivity as intellectual reflection. See PATZELT (1987), pp66ff.

8 It appears that ARISTOTLE's four causing principles—*causa efficiens, finalis, materialis, formalis*—would be extremely useful for the analysis of such processes. See RIEDL (1985), pp80–95.

9 In this respect, ethnomethodology contributes the ambiguous term of 'sense of social/political structure' to the analysis of institutionalization and institutionality. 'Sense of social resp. political structure' means, on the one hand, the sense invested (and later to be symbolized) in social structures, such as institutions. On the other hand, it means that 'sense' (in the meaning of 'interpreting competence') which has to be acquired if one desires to cope with everyday practical problems in a certain set of structures, and if one wishes to be successful with and in it. Cf. PATZELT (1987), pp49–51.

10 On the technique of theory construction practiced here, see PATZELT (1993b).

11 Technische Universität Dresden (1996), p7. This question is put in concrete terms by the following points listed there on p19: institutional processes; time constructions; historical narratives about the own institution; changes of institutionalization principles.

12 On the indispensable distinction between teleology and teleonomy see PITTENDRIGH (1958), pp390–416, and RIEDL (1985), pp156–159 and pp74–77, where the concept of historicity is connected with that of the reality strata.

13 The term *genetic typology* was chosen here because the 'genetic kinship' of system components shall be used a type-forming criterion.

14 The term 'fulguration' (from Latin *fulgur*, flash of lightning) was used by Konrad LORENZ as a term for the sudden arising of new system properties, which can happen at the simple coupling of systems, in which the new characteristics have not—and not even in disposition—been present. See LORENZ (1973), pp47–55.

15 Obviously we are talking here about two of ARISTOTLE's causing principles: the *causa formalis*, and the *causa materialis*. Cf. PATZELT (1986), pp174f, and RIEDL (1985), pp86–90.

References

Berger, P. L./Luckmann, T. (1971) Die gesellschaftliche Konstruktion der Wirklichkeit (The Social Construction of Reality). S. Fischer: Frankfurt a. M..

Braudel, F. (1958/1976) Geschichte und Sozialwissenschaften. Die 'longue durée' (History and Social Sciences. The 'longue durée'). In: Wehler, H.-U. (ed) Geschichte und Soziologie (History and Sociology). Kiepenheuer & Witsch: Köln, pp. 189–215.

Deutsch, K. W. (1973) Politische Kybernetik. Modelle und Perspectiven (The Nerves of Government). Rombach: Freiburg.

Garfinkel, H. (1967/1984) Studies in Ethnomethodology. Politiy Press: Cambridge UK.

Gehlen, A. (1986) Urmensch und Spätkultur. Philosophische Ergebnisse und Aussagen. 5. Aufl. (Primitive Man and Late Culture. Philosophical Results and Statements). Aula-Verlag: Wiesbaden.

Gehlen, A. (1993) Der Mensch. Seine Natur und seine Stellung in der Welt (Man. His nature and his position in the world). In: Rehberg, K.-S. (ed) Arnold–Gehlen–Gesamtausgabe Bd. 3. Klostermann: Frankfurt a. M., pp. 387–480.

Giddens, A. (1984) The Constitution of Society. Outline of the Theory of Structuration. University of California Press: Berkeley.

Göhler, G. (1988) Soziale Institutionen. (Social Institutions). In: Luthardt, W./Waschkuhn, A. (eds), Politik und Repräsentation. Beiträge zur Theorie und zum Wandel politischer und sozialer Institutionen (Politics and Representation. Contributions to the Theory and to the Change of Political and Social Institutions). SP-Verlag N. Schuren: Marburg, pp. 12–28.

Holzner, B. (1972) Reality Construction in Society. Schenkman: Cambridge MA.

Käsler, D. (1974) Wege in die soziologische Theorie (Ways Towards Sociological Theory). Nymphenburger Verlagsbuchhandlung: München.

Lorenz, K. (1941) Kants Lehre vom Apriorischen im Lichte gegenwärtiger Biologie (Kant's doctrine of the a priori in the light of current biology). Blätter für Deutsche Philosophie 15: 94–125.

Lorenz, K. (1973) Die Rückseite des Spiegels. Eine Naturgeschichte menschlichen Erkennens (The Back of the Mirror. A Natural History of Human Cognition). Piper: München, Zürich.

Luhmann, N. (1979) Schematismen der Interaktion (Schematisms of interaction). Kölner Zeitschrift für Soziologie und Sozialpsychologie 31: 237–255.

Meyer, P. (1996) Grundformen menschlicher Sozialsysteme aus der Perspektive der Evolutionären Erkenntnistheorie (Basic Forms of Human Social Systems in the Perspective of Evolutionary Epistemology). In: Riedl, R./Delpos, M. (eds) Die Evolutionäre Erkenntnistheorie im Spiegel der Wissenschaften. WUV-Universitäts-Verlag: Wien, pp. 200–224.

Patzelt, W. J. (1986) Sozialwissenschaftliche Forschungslogik (Logic of Social Science Research). Oldenbourg: München, Wien.

Patzelt, W. J. (1987) Grundlagen der Ethnomethodologie. Theorie, Empirie und politikwissenschaftlicher Nutzen einer Soziologie des Alltags (Foundations of Ethnomethodology. Theory, Empirical Research, and Uses for Political Science of a Sociology of Everyday Life). Fink: München.

Patzelt, W. J. (1993a) Einführung in die Politikwissenschaft (Introduction to the Political Science), 2nd edition. Rothe: Passau.

Patzelt, W. J. (1993b) Formen und Aufgaben von 'Theorieforschung' in den Sozialwissenschaften (Forms and Tasks of 'Theory Research' in the Social Sciences). Ethik und Sozialwissenschaft (Ethics and Social Sciences) 4: 111–123.

Patzelt, W. J. (1995) Vergleichende Parlamentarismusforschung als Schlüssel zum Systemvergleich. Vorschläge zu einer Theorie- und Forschungsdebatte (Comparative legislative research as key to comparative systems analysis. Proposals for a debate on theory and research). In: Steffani, W./Thaysen, U. (eds) Demokratie in Europa. Zeitschrift für Parlamentsfragen, Sonderband zum 25jährigen Bestehen. Westdeutscher Verlag: Opladen, pp. 355–385.

Patzelt, W. J. (1996) Zum politikwissenschaftlichen Nutzen der Evolutionären Erkenntnistheorie (The utility of Evolutionary Epistemology for political science). In: Riedl, R./ Delpos, M. (eds.), Evolutionäre Erkenntnistheorie im Spiegel der Wissenschaften. WUV-Universitäts-Verlag: Wien, pp. 273–288.

Patzelt W. J. (1998) Reality construction under totalitarianism. In: Siegel A. (ed) The Totalitarian Paradigm after the End of Communism. Towards a Theoretical Reasassment. Rodopi: Amsterdam, Atlanta, pp. 239–271.

Pittendrigh, C. (1958) Adaptation, natural selection and behavior. In: Roe, A./Simpson, G. (eds) Behavior and Evolution. Yale University Press: New Haven, pp. 390–416.

Popper, K. R. (1947) The Open Society and its Enemies, Vol. 2. Routledge & Kegan: London.

Popper, K. (1974) Das Elend des Historizismus. 4. Aufl. (The Poverty of Historicism). Mohr: Tübingen.

Popper, K. (1984) Objektive Erkenntnis. Ein evolutionärer Entwurf (Objective Knowledge. An Evolutionary Approach). Hoffmann und Campe: Hamburg.

Riedl, R. (1975) Die Ordnung des Lebendigen. Systembedingungen der Evolution (The Order of the Living. System Conditions of the Evolution). Parey: Berlin, Hamburg.

Riedl, R. (1980) Biologie der Erkenntnis. Die stammesgeschichtlichen Grundlagen der Vernunft. 2. Aufl. (Biology of Cognition. The Phylogenetic Foundations of Reason). Parey: Berlin, Hamburg.

Riedl, R. (1985) Die Spaltung des Weltbildes. Biologische Grundlagen des Erklärens und Verstehens (The Splitting of the World Picture. Biological Foundations of Explaining and Understanding). Parey: Berlin, Hamburg.

Riedl, R. (1986) Begriff und Welt. Biologische Grundlagen des Erkennens und Begreifens (Conception and World. Biological Foundations of Recognizing and Conceiving). Parey: Berlin, Hamburg.

Riedl, R./Delpos, M. (eds) (1996) Die Evolutionäre Erkenntnistheorie im Spiegel der Wissenschaften (Evolutionary Epistemology in the Mirror of Sciences). WUV-Universitäts-Verlag: Wien.

Riedl, R./Wuketits, F. M. (eds) (1987) Die Evolutionäre Erkenntnistheorie. Bedingungen, Lösungen, Kontroversen (Evolutionary Epistemology. Conditions, Solutions, Controversies). Parey: Berlin, Hamburg.

Technische Universität Dresden (1996) Antrag auf Einrichtung eines Sonderforschungsbereichs (Proposal for the Establishment of a 'Sonderforschungsbereich'); TU Dresden: Dresden.

Vollmer, G. (1983) Evolutionäre Erkenntnistheorie. 3. Aufl. (Evolutionary Epistemology). Stuttgart: Hirzel.

Wuketits, F. M. (1990) Evolutionary Epistemology and Its Implications for Humankind. State University of New York Press: Albany.

Elizabeth Ralevski

Aesthetics and Art from an Evolutionary Perspective

Abstract

Universal principles underlying aesthetics and art have been a topic of debate among philosophers and art historians for centuries. Similar debates have surrounded the issue of the functional uniqueness of art from an evolutionary perspective. This paper reviews the current experimental research in universal aesthetic judgment pointing to the conclusion that basic aesthetic rules may be common to all humans. In addition, the paper proposes a new model for the phylogenetic development of art with the basic premise that art has survived evolutionary history because it serves a very unique role in the formation of our personal and cultural identities.

Key words

Art, aesthetics, biological adaptation, evolution, sociobiology, theory.

Introduction

For the last few centuries the arts—music, poetry, literature, and the visual arts—have slowly shifted their primary aim from glorifying religion and royalty to worshiping individual self-expression. As a result, artistic creativity has been increasingly linked to the artist himself and his personal interpretation of the world. Artists' inspiration and its sources have become the main focus of debate that is centered on the question of whether true art should reflect inspirations that are spontaneous and instinctual (the romantic view—see ABRAMS 1984; WORDSWORTH/COLERIDGE 1963), or calculated and self-conscious (the view of modernism and post-modernism—see, for example, SASS 1999). This shift in artists' motivation has also been accompanied by alterations in artistic styles and experiential aesthetics. The change in styles has been so rapid in the last 100 years that the number of "isms" in the twentieth century may equal the total number of artistic styles introduced in the entire preceding history of western art. Moreover, in modern art the aesthetic experience—traditionally equated with an appreciation of the beautiful and the pleasurable—has been increasingly replaced with an admiration of the offensive and the abominable.

Faced with continuous dramatic changes in art and aesthetics, philosophers and art historians of the last two centuries in particular, have endlessly debated the true universal purpose of art and the traditional aesthetic theory based on the universal principle of pleasure. Frustrated in their inability to reach any definitive conclusions, some have gone as far as to propose an abandonment of the centuries-old philosophical search for universal purpose and universal aesthetics in art (PEKHAM 1973). In the last 30 years, only a few have dared to explore a universal theory of art and, interestingly, the most avid proponents have not been philosophers.

This paper has two main goals. The first is to review the research in aesthetics from an ethological perspective, and the second is to propose a new model for phylogenetic development in art. The first main section of the paper differentiates between art and aesthetics, making the distinction between the actual product of creativity (the art) and the concrete aesthetic experience (the aesthetic response) that results from the encounter with art. This first section also gives a brief historical overview of philosophical and psychological theories that attempt to explain aesthetic phenomena. Finally, this section summarises research in aesthetics, with a main focus on current advances in experiments on the visual system and the brain.

The second main section of the paper outlines a new model for the development of art by tracing the development of artistic products (works of art) and the development of aesthetic responses to these works. The core argument in this section centres on the proposition that the main purpose of art is to promote more effectively individuals' self-identity and to designate more clearly individuals' place in the collective culture and identity. Points of disagreement with existing theories[1] are also delineated.

Defining Art

The definition of art remains a modern problem, conceivably because it is rooted in the even more illusive construct of creativity. The most common approaches in defining its complexity have focused on aspects of: i) the artistic product, ii) the artistic process, or iii) the influence of the environment (TAYLOR 1988; WINNER 1991). The artistic *product* refers to the artistic expression itself, manifested as things, performances, or ideas. The artistic *process* refers to both the conscious and unconscious cognitive processes involved in the creation of art. The *environment* has been seen as an important stimulus to, and modifier of, the artistic process. All of the above constructs have been used in the empirical study of the art and have often been seen as complementary aspects of one, unified concept of art. In the context of the current discussion, the notion of art will be limited to the actual *product* of creative expression.

Defining Aesthetics

Art in its many forms (e.g., painting, music, dance, or poetry) often evokes feelings and emotions that are characterised as "aesthetic". The nature of this experience, and a clear definition of which specific reaction to art is distinctly "aesthetic", has remained elusive for centuries (DAVIES 1991). After nearly 2000 years of debates and controversy, the field of aesthetics is not any closer to defining the parameters of the aesthetic response than were PLATO or ARISTOTLE (WILD/KUIKEN 1992; WOODS 1991). Attempts to understand the *nature* of this phenomenon, by both philosophers and psychologists, can be classified into two diametrically opposed positions: one arguing that the aesthetic experience represents an unengaged, intellectual, distant reaction to the work of art (HOSPERS 1976; OSBORNE 1970, 1970a; STOLNITZ 1986), and the other claiming that the aesthetic experience is generated by our very involvement (visual and cognitive) in the work of art (BERLEANT 1986; DUFRENNE 1973; ROELOFS 1998).

Also controversial is the *definition* of aesthetic experience. A related question is whether this experience is a *pleasurable* sensation derived from an encounter with a *beautiful* object or event (HUMPHREY 1973; WOODS 1991) or whether the aesthetic experience is simply a matter of arousal relating to the level of complexity in the visual stimuli, without consideration of any hedonic properties (ARNHEIM 1992; BALTISSEN/OSTERMANN 1999; BERLYNE 1971). The lack of clear definition of what constitutes an aesthetic response has led to a great deal of diversity in methodology and measurement of the aesthetic experience (HEKKERT et al. 1994; HEKKERT/SNELDERS 1995; MCLAUGHLIN/MURPHY 1994; WILD/KUIKEN 1992; WOODS 1991). However, despite the discord in theoretical orientation, and disagreement in measures or conflict in methodology, the interest in aesthetic issues has remained.

In the last 100 years, thousands of experiments have been conducted investigating aesthetic phenomena. The most common have been preference studies with colours, shapes and objects, designed to discover common human propensities of the aesthetic response (ARNHEIM 1992; BOYNTON/OLSON 1990; SILVER/FERRANTE 1995). Although agreement exists that some aspects of aesthetic preference, such as colour, are universal (LIND 1993; SILVER/FERRANTE 1995), controversy endures as to the universality of preference for certain proportions (GREEN 1995), shapes (HOGE 1995) or pictures (CLEMMER/LEITNER 1984; KETTLEWELL 1988; WINSTON/CUPCHIK 1992). The accent is often on the apparent diversity of individual taste for aesthetic objects, emphasising the influence of culture, knowledge and taste on aesthetic judgement (UDUEHI 1995; WINSTON/CUPCHICK 1992). However, the fact remains that individuals from different time periods and cultures agree that the pyramids of Giza are magnificent, the sculptures of MICHELANGELO are divine, and the architecture of BERNINI is splendid (DE LA CROIX et al. 1991). It is difficult to account for such an agreement unless we assume that humans share some common properties or structures which allow us to concur in our judgement of what is preferable.

Historical Development of Aesthetics

In the West, discussions on problems of aesthetics have been around since the time of the Greek philosophers. Aesthetic debates, however, also go back centuries to the Indian, Chinese and Japanese civilizations (BERLYNE 1971). Since the 17th century a separate branch of philosophy was established to address specific issues on aesthetics (OSBORNE 1970; PAUL 1988). In the beginning, its main objective was to explain, in mostly abstract and speculative terms, the concept of beauty—for example, whether it is an objective property of things or a subjective human experience. The hope was that philosophical aesthetics would discover and demonstrate universal laws of beauty or aesthetic judgement, particularly in relation to art.

Philosophical theories

From a philosophical perspective the study of aesthetics has differed, depending on the epistemological (theory of knowledge) and ontological (theory of being) positions adopted. Epistemologically, philosophical aesthetics can be categorized into three main schools of thought: empiricism [LOCKE, BERKLEY, HUME] (HUTCHINS 1980a), rationalism [PLATO and LEIBNIZ] (HUTCHINS 1980b) and transcendentalism [KANT] (HUTCHINS 1980c; PARSONS 1992).

For the empiricists, knowledge (with the exception of mathematics and logic) is obtained mainly through the senses. Based on the empiricist notion that the mind is more or less a passive receiver of the outside world, (a so-called "tabula rasa"), agreement in aesthetic judgement is attributed to the inherent characteristics of the object itself. Pure philosophical empiricism is now almost extinct. At present, general agreement exists that knowledge builds upon predetermined predispositions and it is possible for things to be in the mind without passing through the senses.

Rationalists believe knowledge can be gained purely by reason. Rationalism emphasizes the mind with its predetermined functions and characteristics, although the rationalist philosophers differ in the degree to which they believe the mind is structured "*a priori*". PLATO perceived the mind as a passive receiver of ideas while others, such as LEIBNIZ, recognized the mind's fluidity and its ability to create. According to PLATO, beauty is characterised by the already determined *idea* of the beautiful, which is general and abstract as well as independent of time, place or personal judgement. Physical objects, for example, PLATO believed, are judged beautiful only if they are in agreement with the *idea* of the beautiful. Common to all rationalist thinking, however, is the notion that the environment plays a minimal role in the structuring of the world. Therefore, from a rationalist perspective aesthetic judgement can be inferred to be universal, although not determined on the basis of spatio-temporal attributes. The judgement of the beautiful rests on the *ideas* that are true representations of the reality. Although rationalism can account for the strong objective validity in aesthetic judgement, it fails to account adequately for its subjectivity.

Transcendentalists claim that human knowledge is derived from experience but is dependent on the structure and function of the human mind. KANT's transcendentalism conceives an aesthetic judgement as universal for everyone and yet unexplainable by any universal concepts (PARSONS 1992). He takes the intermediate position which recognizes both that an aesthetic judgement is a subjective experience and that objects themselves contain beauty as an intrinsic property (ADLER/GORMAN 1980). In other words, KANT recognizes that an aesthetic judgement is influenced by external objects perceived through the senses, but at the same time, it is molded, limited and made possible only through the faculties of the mind, which in terms of function and character, are the same for all human beings.

Although theories of knowledge acquisition provide a basis for understanding the validity of aesthetic judgement, ontological theories address the aesthetic object itself. Ontological theories are important for understanding the role of the external object in the process of aesthetic judgement. Idealism, mostly favoured by rationalist philosophers, postulates that true beauty exists only in the idea of the beautiful thus placing emphasis on content rather than form. In contrast, transcendental ontology favours form over content. Radical materialism, favoured by classical empiricists, which conceives the aesthetic object as purely matter, is now extinct.

Therefore, evaluation of the philosophical theories allows us to conclude that the most valuable position to adopt in the study of universal aesthetic judgements is the philosophical transcendentalism. Although philosophy has been providing fruitful theoretical ground for the study of aesthetics for centuries, empirical investigation of this phenomenon has remained outside of the philosophical discipline.

Psychological theories

Early in the 20th century, psychoanalysis and Gestalt psychology provided the groundwork for the theoretically based investigation of aesthetics. The major psychoanalytic contribution to the field of aesthetics has been its emphasis on individual artists and the focus on the content of artists' work (ADAMS 1993; COMFORT 1962). Art in psychoanalytic terms was explained in very much the same way as other products of our imagination, such as dreams, psychotic fantasies, jokes, fairy tales, etc. Art was, according to FREUD, another unconscious vehicle through which unfulfilled wishes were being expressed (ADAMS 1993). The main problem with the psychoanalytic approach to aesthetics is its main focus on content. It gives no guidance or rules on how to approach aesthetic phenomenon in terms of structure and form. This is the gap bridged by the Gestalt school.

Gestalt psychology brought two original and very useful ideas to the field of aesthetics (HAMLYN 1957; SMITH 1988). The first was the notion of 'physiognomics' which explained why human behaviours and postures are identified in non-human objects. For example, the weeping willow looks sad because its branches are hanging and are structurally expressive of the same emotional state in humans. The second, and ultimately the most useful to the study of aesthetics, was the idea of "goodness of configuration". According to the Gestalt psychologists our perceptual system is structured in a way that prefers certain configurations and organizations over others, and follows the 'law of Pragnanz' (WERTHEIMER 1967). According to the Gestalt theory, the distinguishing characteristics of this 'internally set good structure' are the tendency towards regularity, symmetry and simplicity. Although the notion of good configuration represents a useful paradigm for the study of aesthetics, it was criticized for having two main problems (BERLYNE 1971). First, it had no way of measuring the degree to which this property of goodness exists in different individuals, and second, it led to the assumption that the best art will contain, to a maximum degree, regularity, symmetry and simplicity.

Although both the psychoanalytic and the Gestalt traditions provided theoretical foundations for the study of aesthetic phenomena, neither school performed any direct experiments on aesthetics. Their theorizing concerning aesthetic issues was mostly based on inferences drawn from experiments concerning other psychological phenomena.

Early Experimental Work in Aesthetics

In the late 19th century, FECHNER (1871) began the first empirical investigation of aesthetic phenomena. Recognized as the 'father' of experimental aesthetics (BORING 1950), FECHNER's early experiments established the foundation for the empirical investigation of aesthetics. His first aesthetic experiments focused on the study of the 'golden section'. The 'golden section' refers to a number, approximately equal to 0.618, that relates to a point on a line at which the line is divided as to create the most pleasing arrangement. When a line is split in the golden section (60/40), the ratio of the shorter part to the longer part is the same as the ratio of the longer part to the whole line (GREEN 1995). FECHNER's early work focused on preference studies using rectangles that followed the golden section rule (FECHNER 1871). His results provided evidence that the golden rectangle was the most preferred shape.

FECHNER also conducted studies in preference for paintings. He asked visitors to the Dresden museum to indicate which one of two paintings they preferred more, an original "Madonna and Child" by HOLBEIN (also known as the Darmstadt Madonna) or a possible forgery (also known as the Dresden Madonna) which was exhibited at the same time (FECHNER 1876). FECHNER was interested in testing the widely held view at the time that the authenticity of the paintings could be determined simply by comparing the *beauty* of the two works of art. Although the experiment was not a complete success due to a very small number of art-goers that took part in the experiment (forty-three out of thousands), the results of his analysis revealed a clear preference for the Darmstadt Madonna. A recent replication of FECHNER's experiment (MARSHALL et al. 1995) found no clear preference for either paintings on the basis of *beauty*, although art-naive subjects judged the two paintings as different in terms of artistic merit, giving higher ratings to the HOLBEIN's Madonna. Other studies conducted by FECHNER included analysis of famous works of art in which he investigated whether great paintings followed the proportion of the golden section (FECHNER 1876). Contrary to his expectations, great paintings were no more likely to illustrate the golden section.

Following FECHNER, in the first half of the 20th century, experiments that were designed to search for universally common principles in aesthetic preference were mainly limited to simple geometrical shapes and the golden section rule. Numerous preference experiments on the golden section were conducted (for a review see GREEN 1995). The general findings, however, were inconclusive. Although some evidence indicated a clear preference for the golden section (FARNSWORTH 1932; WITMER 1894), other data provided evidence that symmetrical division (PIERCE 1894), or figures with different proportions (i.e., 2:1; 3:1; 4:1), were preferred over the golden section (DAVIS 1933; WEBER 1931).

The predominant finding in psychology in the first half of this century has been that there are *no* universal rules guiding aesthetic preferences. These findings were supported by research in the later half of the 20th century (PIEHLE 1976).

Aesthetics in the later half of the 20th century

In the early seventies, after centuries of debate by philosophers and a great number of inconclusive studies by psychologists, the discovery and demonstration of universal laws of beauty that govern aesthetic judgement were as elusive as ever (PECKHAM 1973). Philosophers and psychologists have failed to produce a unified definition, not only of what constitutes 'beauty' (as a pleasurable stimuli), but have increasingly debated whether beauty even has any relationship to art. In addition, they have not been able to agree on what function art plays in our society in general. Some philosophers have argued that the intellectual debate on questions of aesthetics and art can only be considered a failure and have wondered if these questions have any meaning at all, some going so far as to suggest abandoning all together the philosophical discipline of aesthetics (PECKHAM 1973). Others have avoided such pessimistic conclusions, considering the effect that it might have on the society as a whole. "By the standard of our culture, if a painter can no longer use the glamorous word 'artist' to identify himself, his whole basis of self-evaluation and public esteem is gone. So tremendous is the status of art in our times that it has become a substitute for religion, a revelation of the true meaning of human life" (PECKHAM 1973, p3).

More prevalent among modern philosophers is the opinion that philosophical aesthetics has lost its relevance to art because modern art does not aim at the pleasurable and the beautiful (PAUL 1988). This view has been gaining ground under the influence of modern art theory and art criticism which argue that a new aesthetics, more relevant to modern art, needs to be explored (MARTIN 1995). This new aesthetic must account not only for the pleasant and the beautiful but more so for the ugly and the abominable, for the shocking and the contemptible. However, even if one accepts the argument that modern aesthetic theory is required, the reality remains that traditional philosophical aesthetics is as relevant today as it was in the time of PLATO. Individuals continue to prefer beautiful things in art, and, considering that modern art ("modernism") is only a small and, most likely, transitory period in the history of art, attempts to unravel the mysteries of the beautiful continue to remain sound.

Two main factors have been responsible for this division of opinion and disappointment among the scholars of aesthetics: i) the glorified status that art has in our culture (EVERETT 1991), and ii) the heterogeneity of what constitutes the category art (BERLYNE 1971; DISSANAYAKE 1988).

Early in history, and in less technologically advanced societies, art has had a major role in social and religious ceremonies and rituals (DISSANAYAKE 1988). Despite their influence, artists had been regarded simply as craftsmen who had been valued for their work but held no particularly high place in society. The idealization of artists as supreme symbols of humanity began in ancient Greece, continued throughout the Renaissance and reached its pinnacle during the Romantic period (DE LA CROIX et al. 1991). Since ancient Greece, artists have been glorified for their divine minds and exceptional wisdom, as keepers of knowledge which was considered beyond the reach of the average person. For this reason, empirical investigations of artistic products by philosophers and psychologists have been seen, not only as undermining the individuality of the particular artistic expression, but almost as sacrilegious.

Another factor that has hindered the progress in aesthetics is the great heterogeneity of the category "art". Philosophers have for centuries tried to develop a unified definition of art guided by two basic principles: i) that everything in the category "art" must have some unique property that is common and essential to all entities in the art category, and ii) art must serve some unified function (OSBORNE 1970). However, the main conclusions, after centuries of debate, have been that all art is not in essence homogeneous and frequently serves no unified function. The category "art" contains classes that are essentially very different; it contains traditional artistic enterprises such as painting, sculpture, dance or music, as well as ventures such as cooking, fashion, and car making. The functional significance of art has been even more confusing. It has changed drastically throughout history. At one time or another, art was thought to serve the following functions: presentation of pleasing forms, production of intense sensations, promotion of morale, promotion of religious piety, provision of knowledge, encourage loyalty to a regime, strengthening social bonds, or encouragement progress. The list can easily be expanded (BRUNIUS 1967; DEWEY 1934; HUMPHREY 1980; STORR 1972).

One way to overcome the above problems is to go back to the early roots of philosophical aesthetics and start with the assumption that art appreciation and art creation are not unique and isolated phenomena. In fact, some works of art are continuously appreciated by different people from different cultures at different periods of time. This implies that there are common factors and principles on the basis of which aesthetic judgements are being made. Also, art in some form or another is practised in every culture on earth (DISSANAYAKE 1988). This also suggests that art creation is a universal behavioural phenomenon that serves some useful function in order to have survived millions of years of evolutionary history.

Current Research in Universal Aesthetic Preference

"Whether mankind considers himself the son of God or a successful ape will make a clear difference in his behaviour towards actual facts: in both cases he will also hear very different commands from within himself." Arnold GEHLEN (1940, p1)

PAUL (1988) has argued that KANT's transcendentalism provides a sound theoretical basis for studying aesthetic judgement as a universal phenomenon. Although subjective aesthetic judgements are formulated by experience, universal and objective aesthetic rules *of form* exist in all human beings. They are based on fundamental and intrinsic universal properties of human perception and function. Therefore, the study of universal aesthetic judgements should focus on the investigation of the human mechanisms involved in the construction of beauty, rather than dealing with the infinite variety of beautiful objects (PAUL 1988).

In 1871, FECHNER (1987) proposed three methods for experimental investigation of aesthetics which remain in common use today: the method of choice, the method of production, and the method of use. The method of choice is the most commonly used technique in which subjects are presented with a variety of objects and asked to select, often by comparison to other objects, the ones that are most pleasing. In the method of production, subjects are asked to create, either by drawing or combining elements, configurations that they find most pleasing. The method of use focuses on the investigation of art works or other objects, with the assumption that the most common configurations in the works of art represent items favoured as the most pleasing in a particular culture.

Current research in universal aesthetic preference combines philosophical transcendentalism and FECHNER's methodology with a focus on mechanisms involved in the construction of beauty, concentrating mainly on the study of the visual system and the brain.

The visual system

The recent search for universal aesthetic rules has focused on the investigation of what our species are biologically *capable* of appreciating. It has been guided by the assumption that all humans are born with perceptual biases that occur on three different levels: i) basic ii) species-specific, and iii) cultural (EIBL-EIBESFELDT 1989). *Basic biases* are the ones we share with other vertebrates. The most fundamental one is the pursuit of visual clarity, the search to make sense of our world. For example, in Rubin's cup, where one can see either a dark vase or two light profiles, we tend to see the dark figure on the lighter background. This search for clarity has been demonstrated in infants as well as adults (EIBL-EIBESFELDT 1988). Human preference for clarity and distinction of forms is well illustrated in a study by BOSELIE/CESARO (1994). In this study subjects were presented with drawings that contained either unambiguous or ambiguous objects in one or two frontal planes. Clear preference was shown for the unambiguous drawings.

Humans also show a strong preference for regularity (SCHNEIDER et al. 1993), order (GOMBRICH 1979), and symmetry (ENQUIST/ARAK 1994), while trying to suppress any imperfections and irregularities (BOSELIE/CESARO 1994). EIBL-EIBESFELT (1988) has argued that we seek and enjoy order as a general principle because it is efficient and adaptive. It is also limited by the small capacity of our short-term memory (according to EIBL-EIBESFELT we process 16 bits per second). DÖRNER/VEHRS (1975) asked subjects to place red and green squares on a grid to create either a beautiful or unpleasant arrangement. They found that the arrangements identified as aesthetically pleasing tended to form geometrical designs. Our search for visual order is relentless even if order does not exist. If the order in the object of our perception is too easy to discover (less than 16 bits per sec) we judge the arrangement as boring. If the order is too complex, we lose interest. BERLYNE (1960, 1974) has argued that the relationship between complexity and the pleasurable aesthetic response is an inverted U-shaped function. In a number of experiments, MARTINDALE and his colleagues (1990) presented subjects with various sizes of polygons that differed in the number of turns from 5 to 20. The most interesting findings were that the size of the polygons had no effect on preference. Judgement was only influenced by the number of turns. Polygons having 10 turns were preferred the most, followed by polygons with 8 and 5 turns. Shapes that had more than 10 turns were rated as significantly less preferable, indicating that with increased complexity of the stimuli (beyond 10 turns), aesthetic preference decreased. However, the more recent tests of BERLYNE's theory did not support his assumptions (MESSINGER 1998). In fact, an inverted U-shaped function was found when interest rather than complexity was compared to pleasure (MESSINGER 1998).

Humans also show a distinct preference for certain geometrical proportions. Squares and rectangles are perceived as aesthetically beautiful if their proportions are 1:1.63, or follow the golden section (EIBL-EIBESFELDT 1988, HEKKERT et al. 1994). Other research has shown that symmetry rather than the golden section is judged as more beautiful (DAVIS/JAHNKE 1991). Yet, others have argued that rectangles with proportions 1:1.5 are in fact preferred over the golden section rectangle (BOSELIE 1992). It is important to remember, however, that these results are not in disagreement with the argument of universal aesthetic judgement. Symmetry, as well as proportions of either 1:1.5 or 1:1.6 are similar enough to suggest that dimensions that are around the golden section, or proportions that follow the golden section are preferred more often than dimensions that dramatically diverge from those proportions (GREEN 1995). Even more intriguing are the findings by MCMANUS and colleagues (1993), who suggest that there is a universality of aesthetic preference for certain patterns and geometrical forms. The authors presented art-naive subjects with computerized pictures of both original Mondrian and modified versions of his paintings, in which the compositional lines were moved slightly. Mondrian's paintings can best be described as geometrical arrangements of lines in vertical and/or horizontal space. Subjects were asked to indicate which painting in a given pair they favoured. Overwhelmingly, subjects showed a preference for the original Mondrian, suggesting that art-naive individuals can somehow detect even minute differences in compositions that encapsulate some general principle of aesthetic order. MAGRO (1999) used line drawings that represented exaggerations of primitive (prehistoric) and modern human anatomical proportions—without being immediately obvious that the drawing were derived from human anatomy. The exaggerated drawings of the modern human proportions were overwhelmingly rated as more attractive. According to the author, the findings suggest that general aesthetic principles may be based on the anatomical shapes and proportions of the modern human.

Animals, as humans, prefer order, regularity and symmetry (MORRIS 1962). Studies of ape paintings, first begun by Desmond MORRIS (1962), show that chimpanzees create "art" that is aesthetically appealing to humans. In fact, when chimpanzee paintings were anonymously exhibited in a London gallery among those of modern artists, a number of action painting critics praised the works by the chimpanzees for their vitality and importance.

Therefore, we can conclude that on a basic level, humans share with animals preferences for clarity, distinction of forms, regularity, order, symmetry, simplicity of information and propensity towards certain geometrical proportions.

According to EIBL-EIBESFELDT (1988), *species-specific biases* have developed through evolution to serve some adaptive function. Humans in many cultures show preference for child-like features, because those features serve to release nurturing responses and behaviours. Delicate features (large eyes, small nose, small beard, high forehead) are associated with beauty across many cultures (CUNNINGHAM 1986; FAUSS 1988). Regardless of race, a fine nasal bridge is considered attractive (RENSCH 1963). This is true even in cultures where only a few people conform to this ideal. Babies as young as 4 months look longer at photographs of female faces that are attractive than at faces that are unattractive (LANGLOIS et al. 1991; SAMUELS et al. 1994). Our consumer society makes perfect use of our biases by creating cartoons, dolls and toys of females with large heads, small bodies and delicate features. Cross-cultural agreements have also been found in judgements of the beauty of human form (MAGRO 1999). Presented with drawings of human anatomical traits arranged in terms of anatomical evolution in primitive traits, intermediate traits and derived traits, individuals from different cultures, race, and age concur in their agreement by preferring proportions and characteristics that are intermediate or derived. There is also a strong aversion towards primitive traits and proportions in humans (MAGRO 1999). Female attractiveness has often been linked to health and vigour (KOGAN 1994; SINGH 1993) as well as the higher likelihood of acquiring a "resource-rich" mate (BUSS 1994; KENRICK/KEEFE 1992).

Of all the features on the human face, however, in many cultures the eyes take the place of great importance. Across cultures, staring eyes are associated with threat (EIBL-EIBESFELDT 1975). The sense of discomfort we feel when being stared at probably has its roots in our primitive past where a predator would fixate its prey before an attack. Staring eyes are often used either to ward off evil or to frighten an enemy. In everyday human communication, staring eyes are often associated with an aggressive encounter. COSS (1970) presented subjects with eyespots and raised eyebrows in three different positions: horizontal, vertical or oblique. He observed the greatest arousal and interest when the eyespots were in the horizontal position. This could explain why in friendly face-to-face encounters people often tend to tilt their head to minimize arousal or threat.

Cross-cultural aesthetic agreements also exist on bodily presentation and proportions. Across cultures, the ideal male body tends to be broad shouldered, small waisted, muscular and slender (EIBL-EIBESFELDT 1989). In many societies throughout history, men have shown the tendency to wear clothing that emphasises their shoulders. EIBL-EIBESFELDT (1988), following MORRIS (1962), traces this aesthetic preference, once again, to our primitive ancestors, who had tufts to emphasize the width of their shoulders.

In contrast to the male, the female ideal of beauty (in terms of bodily proportions) has been more susceptible to change. Cross-cultural studies show that until recently the ideal female body was a plump female with large buttocks and breasts (ANDERSON et al. 1992; BROWN 1991; CASSIDY 1991). According to BROWN (1991), nine out of ten societies—among cultures in the Human Relations Area Files—show a preference for women with fat hips and legs. Historically, the high correlation between fat and social status can be easily explained in evolutionary terms. Humans faced with frequent food shortages in the preindustrial societies had developed a variety of physiological and behavioural mechanisms to store fat during times of plenty. Since the upper classes had more food available and did less physical work, they could store more calories. Until the later half of the last century in United States, fat explicitly was associated with wealth (SEID 1989).

The idealisation of the thin female body among Western cultures is a recent phenomenon that began at the end of the nineteenth century and has grown from a "prejudice" in the 1950s to a "religion' in the 80s (SEID 1989). Numerous theories have been proposed to account for the current trends. Cross-cultural analyses has shown that slim female body preferences are to some degree related to climate or the status of women in their society (ANDERSON et al. 1992). SMUTS (1992) has argued that the recent high correlations between high-status and slimness among women is a result of both chronic food surplus and the change in the male/female work ratio. According to SMUTS (1992), in a society where food shortage is no longer a problem, body size is no longer a high-status symbol. On the contrary, thinness has become a symbol of high status and wealth. Slimness, coupled with women's increasingly higher earning power, increases her likelihood of getting a better mate. The often quoted phrase by the Duchess of Windsor, "No woman can be too slim or too rich" (ANDERSON et al. 1992; POLIVY et al. 1986; SMUTS 1992), illustrates SMUTS' point well.

Certain gestural elements, traceable to our animal ancestry, are also universal: for example, vertical open palms, exposed teeth, gaping mouth, frowned, phallic displays (showing dominance, aggressive encounters such as rape) (EIBL-EIBESFELDT 1988).

ULRICH (1986), as well as ORIANS/HEERWAGEN (1992) have demonstrated that, across cultures, humans show strong preferences regarding the way their environment is arranged. People prefer dwellings that give them a sense of comfort and security. There is a strong cross-cultural preference for thicker walls. Seating arrangements in niches or corners, with back to the wall, are also strongly preferred. Houses are often chosen with an unobstructed view to insure territorial security. Also, humans generally prefer natural rather than built landscapes to live, work and visit (PURCELL et al. 1994).

As humans, we also share other biases. The fear of heights is a universal phenomenon, and light is always associated with positive things (EIBL-EIBESFEDT 1989). Colour preference research shows that humans show a strong preference for the colour blue (SILVER/FERRANTE 1995). One exception may be the populations of Asia, who show a strong selection bias for the colour white (SAITO 1994). Colours such as red and orange are associated with warmth, while other colours, such as blue and purple, are linked with coldness (ITTEN 1961).

We can conclude that specific to our species are preferences that include delicate, child- like facial features, certain bodily presentations and proportions, some gestural elements, cosy environmental arrangements, or the colour blue.

Culture-specific biases, according to EIBL-EIBESFELDT, are best discussed in terms of style. In Europe, style in art and architecture played an important role in bridging national and ethnic boundaries. Style reflects the needs of the society and thus serves many different functions. Style serves to define cultural values and norms. For example, portraits of bravery and honesty in one's countrymen emphasise the importance of those values in a particular society. Individual identity and worth are defined by style or by the things the members of a society gather and consider to be valuable. Style can also be used to culturally encode messages, as in Egyptian hieroglyphs.

In summary, our visual system has been constructed on different levels with biases that reflect phylogenetic adaptational needs. Perceptually we strive towards clarity, regularity, order, and symmetry. Also, general agreement exists on preferences for colour, facial and bodily features, gestures, as well as natural, open environments. Across cultures, differences in aesthetic preferences reflect variations in values and norms that are expressed in a particular society in terms of style.

The brain

Biological investigation of aesthetics would be incomplete without considering how the brain processes aesthetic information. LEVY (1988) has investigated the role of the two hemispheres in the evaluation of aesthetic experience and expression as it relates to visual art. He studied artists and ordinary individuals with damage to either the left or the right hemispheres.

It seems intuitive to assume that aesthetic experience is characterized by collaboration between the right and the left hemisphere. In non-artistic individuals, damage to either the left or the right hemisphere results in defects in drawings. By contrast, similar damage to either hemisphere in talented artists results in no such consequences. They either experience no deficit or, with recovery, regain the premorbid ability to create art. LEVY concludes that processes necessary for artistic creation are bilaterally represented, so that damage to either side of the brain has less effect on the ability to create. Perhaps artists have a superior ability to communicate between the left and right hemispheres. Alternatively, the processes essential for artistic creativity may be represented more equally in both hemispheres, facilitating communication between the two sides. Highly accomplished musicians show the same bilateral superiority when compared to less talented musicians.

These findings suggest that in, non-artistic individuals, aesthetic experience is asymmetrically represented in the brain. In fact, asymmetric activation of the brain occurs depending on the stimuli presented. This asymmetric activation can therefore lead to a perceptual bias. LEVY (1988) reports that most right-handed people, when presented with a face judgement task, judge that the smile occurs on the right side of the face when in fact the figures are mirror images. Left-handed people show a similar perceptual bias. Other studies using scenes, random shapes or other nonverbal visual stimuli show the preference for the subject of interest to be to the right of the picture (e.g., LEVY 1976; POPPEL/SUTTERLIN 1983). When a picture is symmetrical, there is no preference. However, when the focal point of the picture is displaced either to the left or to the right, the preference is for the focus to be on the right. It is important to note, however, that the right visual preference for scenes was studied by LEVY with right-handed people only (LEVY 1976). The question to answer is why a left hemispheric perceptual bias exists, LEVY argues that the bias is there to harmonize and balance the level of engagement of both hemispheres. The right hemisphere is activated when a visual stimulus is presented. When, in addition, an asymmetric visual stimulus is given by placing the object of interest to the right, the left hemisphere is activated also, allowing for joint engagement of both hemispheres.

According to the forgoing hypothesis, one would assume that affect will influence the representational placement in pictures. LEVY and colleagues (1983) have shown that for right-handed individuals, happiness results in high activation of the right hemisphere, whereas sadness leads to low activation of the right hemisphere. Studies with eight-year-old children show that, when asked to make happy or sad pictures, children place the euphoric content on the right side and the dysphoric content on the left (HELLER [unpublished doctoral manuscript], in LEVY 1988).

REGARD/LANDIS (1988) investigated aesthetic judgement in terms of lateral brain differences. Their early studies, done with a tachistoscope (an apparatus that very quickly flashes information to each hemisphere separately), showed that the two hemispheres process visual information differently. For example, the right hemisphere processes information faster, more globally, and often with less verbal awareness than the left hemisphere. Knowing these hemispheric differences in processing visual information, the authors were interested in two things: first, which aspects of visual stimuli influence aesthetic preferences, and second, how are the left and right hemispheres independently involved in the aesthetic judgement.

In their first study, right-handed males and females were presented with figure pairs in a vertical position for 150 milliseconds. The pairs of figures were presented to either the right or the left visual field, and subjects were asked to indicate a preference for one of the figures by pressing a button. Pairs were selected in such a way that one member of each pair always obeyed the rule of *pragnanz* (the tendency toward order, simplicity and completeness). The results showed that the *pragnanz* figure was preferred when presented to the *right visual field* (left hemisphere). Latencies were longer for both sexes, regardless of hemisphere, when the pragnanz figure was chosen, suggesting slower decision-making when presented with aesthetically pleasing stimuli. In addition, the interaction between judgement and sex was significant. Males showed significantly greater preference for the non-pragnanz figure when presented to the left visual field. For women, by contrast the preference was in the opposite direction, although not significant. Based on these findings, the authors reached a number of conclusions: i) aesthetic selections seem to have an affective component, as shown by significantly longer latencies for the pragnanz figures; ii) affect seems to be processed differently by the two hemispheres; iii) males and females make different judgements when information is presented to the same hemisphere; and iv) the quality of the stimuli (whether the figure illustrates pragnanz or not) influences perception and aesthetic judgement.

REGARD and LENDIS' second study (1988) was designed to investigate the hypothesis that the right hemisphere is involved in subliminal perception. It is believed that the right hemisphere processes visual information faster and more globally. Subjects were presented for 1 millisecond with a pair of stimuli that consisted either of two human faces or a human face and a picture of a potato. The potato was the same size and shape as the human face. The visual pairs were shown to either the right or the left visual field. After the presentation of the face-potato pair, subjects were exposed to the face-face pair, one of which was previously seen in the face-potato pair. At the end of the session, subjects were given a recognition test (Which face have you seen before?) and an affect judgement test (Which face do you like better?). The results revealed that the affective judgement was influenced by the subliminal presentation. Faces presented to the left hemisphere were preferred, and those presented to the right were disliked, which implies that the left hemisphere prefers the known whereas the right prefers the new.

The authors conducted a further experiment allowing longer exposure times (3–5 ms and 20 ms.) with different subjects. They found that when the presentation was 3–5 ms (allowing only discrimination between face and potato), no effect on recognition or preference was found. However, in the 20 ms condition, recognition was affected. The well-documented right hemisphere superiority in recognition was replicated, but no influence on preference was found. The authors concluded that affective judgement involves both hemispheres, although the left hemisphere prefers the known whereas the right prefers novelty.

Their third study was designed to investigate hemispheric processing of ambiguous stimuli. Twenty-four subjects were presented with 10 Rorschach inkblots in random order for 150 ms. The inkblots were presented randomly to each hemisphere. Subjects' interpretations were taken and coded in terms of responses to form, colour, shading, and content. The responses for each inkblot were coded by a specialist, and "psychograms" for each visual field were derived. The results revealed that 57% of the interpretations differed when the same inkblot was presented to both hemispheres independently. Twenty-seven percent of the responses were the same but verbalized differently, and 16% of the responses were identical. The Rorschach interpretations were grouped in terms of their mode of approach and further analyzed. A trend analysis revealed no difference in mode of approach; most subjects discussed the inkblots as wholes. However, a majority of the subjects differed in the quality of responses they gave to the same inkblot when presented to different hemispheres. On the basis of those differences, the authors were able to develop distinct personality profiles for both hemispheres. They concluded that the right hemisphere seems better at combining and identifying details, although sometimes at the price of missing the whole; it is better at comprehending what is near and immediately given, even when the given information is incomplete. It is also better at making inferences from incomplete information. The right hemisphere interprets forms with more fantasy, greater precision (even with lack of information) and with greater originality. It is more guided by affect. The left hemisphere, however, efficiently handles the understanding of the subject matter but shows little combinatorial ability. It processes information in a very cool and controlled manner, by sanctioning affective impulses, and keeping emotional responsiveness to a minimum. In summary, both sides of the brain process aesthetic information differently, depending on gender, familiarity of information and the affective aspects of the stimuli presented.

As a general conclusion, the research in universal aesthetic judgement allows us to infer that basic aesthetic rules, uninfluenced by knowledge or training, are common to all humans. These aesthetic rules can be found in the fundamental mechanisms of human perception and brain function. Common to our species are preferences for harmony, symmetry, and unity. Universal preferences for facial features, colours, or gestures also vouch for mutual roots in our biology. General rules that underlie the function of our brain also define how aesthetic judgements are being made. For example, the right hemisphere, which is better at recognizing stimuli, is activated when pleasant events are presented and seems more interested in things that are new. The left hemisphere, on the other hand, shows preference for pragnanz forms, reacts more to negative stimuli, and seems more interested in things that are known. However, although the rules of universal aesthetic judgement are fixed, the flexibility of our mental architecture allows for subjective aesthetic variation. The multitude of aesthetic expressions and definitions of the beautiful in the history of art, as one example, bear witness to our extraordinary ability to always create something new and different.

The Function of Art and Its Development

The above evidence of universal aesthetic judgement is persuasive. However, this brings us only a small step closer to answering the fundamental question: Why was art (as behaviour) selected for in the first place? The problem is that universal rules of aesthetic judgement are relevant to art as well as nature or ordinary everyday objects. The experience of beauty or aesthetic pleasure is a part of numerous aspects of everyday life. For example, we find flowers and landscapes beautiful, we aesthetically evaluate our wardrobe, or choose our purchases often on the basis of what is exquisitely appealing to our senses. The universality of mechanisms that allow us to experience the beautiful or the aesthetically pleasurable can explain two things: i) the reasons why we select certain objects as more pleasurable than others, and ii) the reasons why we select to express ourselves artistically in terms of shapes and colours that evoke pleasurable experiences. However, it tells us nothing about the reasons we engage in the production of art in particular or how this behaviour evolved in the first place. It is possible that the behaviour of making art serves a unique function in order to survive millions of years of adaptive evolution.

Within the last 30 years, only a few theories have been proposed to explain the functional uniqueness of art and its development from an evolutionary perspective. All, however, have been criticized for either not addressing or fully missing some links in the developmental sequence of the artistic experience and behaviour.

New Model for the Phylogenetic Development of Art

The development of art as a human adaptation will be discussed in terms of two separate, yet related issues: the aesthetic response (emotional reaction) and the aesthetic expression (the work of art). The aesthetic reaction to a work of art will be traced back to the basic aesthetic experience of pleasure when encountering certain patterns, sounds or colours. The aesthetic expression will be sketched through four distinct developmental adaptations of ritual, play, imagination and self-awareness. The new model rests on the basic premise that art has survived evolutionary history, because it serves a very unique function in the construction of our personal and cultural identity (for a similar argument relating to music, see FRITH 1996). Although art is often discussed in terms of how it affects and reflects the culture in which it is produced, the argument here will focus on how the artistic expression, regardless of whether it refers to a work of art or a bodily decoration, has permitted us to: i) differentiate ourselves from the group in a unique manner; ii) project the image of our imagined self, and iii) designate our place in the collective culture and identity.

The four basic adaptations of ritual, play, imagination and self-awareness are all integral parts of both the development of art and the evolution of self-identity. Ritual is similar to art in its use of metaphors and symbolism, in its ability to evoke out-of-the-ordinary experiences, or in the capacity to take everyday objects and experiences and charge them with an extraordinary meaning. Rituals as art are also essential in uniting and binding the culture in a common belief, explaining and providing meaning to the world around us, documenting and certifying practices that are important to the culture, and, by understanding our place in the group, promoting stable social functioning. Play and art are similar in being unpredictable, fun, free, spontaneous, metaphorical, and exaggerated. Play is also important in discovering who we are by adapting different identities, experimenting and mastering many social skills, or discovering and perfecting one's abilities in a non-threatening environment. Imagination and self-awareness are essential to the artistic process because they allow for a rich world of symbolism to be transformed from the realm of the ordinary to the third dimension of neither subjective nor objective work of art. Imagination and self-awareness are vital in the development of self-identity, attachment and socialization. In summary, we experience and produce art because it helps us to understand *who we are, what we want to be,* and where *we stand in the relationship with the rest of our culture.*

The aesthetic response is rooted in what DARWIN called the sense of the beautiful. Therefore, the development of the aesthetic response can be traced back to our animal ancestry. This is clearly articulated in the writings of DARWIN (1979, originally written in 1859). DARWIN, in his detailed observation of the animal species, noticed that some birds and animals have ornamentations and bodily features which, at first glance, did not seem to serve any useful function (e.g., peacock feathers). Those ornamentations were usually present in the male of the species. After long deliberation, DARWIN concluded that those features developed to attract the female and thus, relate only to the opposite sex and mating. He postulated that they have developed as a way of competition between males, the most vigorous of which will attract more females, be more successful in warding off potential rivals, get the better mates and have a greater chance of producing more offspring. DARWIN wrote: " When we behold a male bird elaborately displaying his graceful plumes or splendid colours before the female, whilst other birds, not thus decorated, make no such display, it is impossible to doubt that she admires the beauty of her male partner" (FULLER 1983, p5). He named this response of the female *the sense of the beautiful* and defined it as the simple pleasure experienced from certain colours, shapes or sounds. DARWIN was unable to explain why only certain sounds or configurations were appealing to one particular species and not to others. He later proposed that differences were due to habits. Since DARWIN, the emphasis has shifted from competition to recognition (COMFORT 1965). The neo-DARWINIANS recognised that the displays of colour and movement in animals serve as recognizable signals to other members of the same species. For example, two males communicate through threat postures that are recognizable to one another. This implies that the displays serve a useful, functional role. Taking all the evidence into account, the neo-DARWINIANS proposed that the sense of the beautiful is merely a congenital response to a signal stimuli. However, others have noted that some responses such as complex bird songs are difficult to explain as a congenital response. Jackdaws, after being trained to certain rhythms, are able to recognize them when played by different instruments, with a different tempo, pitch or interval. This must require an explanation different than a congenital response. Perhaps DARWIN's original idea of an aesthetic feeling of pleasure would be a better explanation.

The explanation of a congenital response was further complicated by the observations that many birds did not have physical displays or ornamentations but rather very complex courtship rituals (dance movements, shaking-bouts, displays) that had no apparent functional significance in the original DARWINIAN sense. It was not clear what the significance of this complex yet seemingly unnecessary behaviour was. Julian HUXLEY in his book "The Courtship Habits of the Great Crested Grebe" published in 1914 was among the first to speculate on this phenomenon. HUXLEY concluded that rituals served to express emotions among male and female birds. Thus, certain movements lost their original adaptive function, merged into a smooth rhythmic sequence, and became symbolic expressions (symbolic communication, ceremonies).

His conclusions are now widely accepted. In fact, it is known that in all vertebrates, rituals are used, not only to express emotions in courtship but also to indicate territorial occupancy, threat, interest, strength, and endurance. They represent clear communication signals designed to release the appropriate response. Such unambiguous signal sending leads to strengthening of sexual and social bonds among the members of the group.

Humans engage in numerous ritualistic behaviours (greeting signals, dance). In primitive societies, rituals and ceremonies are a major and integral part of everyday life. Although they are culturally based, human and animal rituals share the same form and function. In terms of form, they are alike in that ordinary movements and gestures develop a new charged special meaning with certain intensity, and independent movements are merged in a single smooth sequence. Functionally, they are identical, in that rituals represent communication devices that promote social cohesion, and help us understand our place in the social structure of the group. As in animals, ritualistic behaviours in our species are effective, because they rely heavily on evoking a pleasurable experience or the aesthetic response.

In the late 1800's, a number of art critics and philosophers wrote on the relationship between artistic activity or expression and play, in both animals and humans. The basic premise was that art and play are, in essence, identical by being beyond the realm of function and necessity. Play as a ritualistic behaviour is characteristic of most mammals and is observed in over 40 species of birds. It is usually found in creatures that have long periods of immaturity and are sociable. Play encompasses both innate and learned patterns of behaviour, and, it is often a combination of the two. Characteristics of play include 'nonfunctionality', unseriousness, and ability to be self-rewarding. Play is often associated with exaggeration, novelty, entertainment, imitation, sociability, symbolism, curiosity and exploration. All of these attributes are present in children's play, and more importantly, are related to *art*. It is important to note, however, that in animals, play is essential in perfecting instinctive behaviours important to survival, such as catching prey. In humans, play is more pursued for its own sake and involves more imagination and creativity. The development of play became necessary with the increased socialization among the species. As certain species became more socially advanced, longer periods of immaturity became necessary to perfect the more complex skills of communication, identity, socialization and survival. "The problem of personal identity, one may say, arises from play-acting, and the adoption of artificial voices; the origins of distinct personalities, in acts of personation and impersonation" (REE 1990).

However, with the complexity of life in higher primates, there were increased pressures on the animals to classify their experiences in an orderly, cohesive manner; to associate and appreciate abstract concepts and configurations. Thus, imagination was born.

Imagination can be defined loosely as the ability to form an image (or concept) when the image is not present to the senses. Primates show the ability to classify and appreciate abstract concepts and configurations (WEINER 1971). They have an extraordinary ability to learn symbolic associations. This is very well demonstrated by their ability to learn sign language. This ability to symbolize and imagine, some have suggested, led to the development of facial and bodily expressions as a means of conveying (expressing) emotions, communicating, and regulating social behaviours. "Desmond MORIS has suggested that the complexity and motility of human facial muscles evolved to promote identity and recognition of the individual within the group by facial expression, and to create group identity through physiognomic types, and so to abet group hunting procedures, group loyalties, and the formation of social hierarchies which could support hunting, defence and cultural development" (KAGAN 1983, p2).

Primates also dream, although it is suspected that their dreams lack complex symbolization. The dream imagery, is for the most part, related to basic drives and instincts. One can speculate that primates have rich, imaginative lives. Similar to human children, young chimpanzees employ their imagination in various ways: they like to dress up, beat on drums, scratch in mud, and paint. They even identify their paintings as representations of physical objects (bird, tree). Are these aesthetic expressions works of art? To answer that, one must examine the similarities between paintings of apes and children. Both young primates and children become interested in the activity at the age of one-and-a-half. Their paintings start with just scribbles and develop into more and more organized and rhythmic patterns. Shapes are drawn in the final stages at about the age of three. From then on, the paintings begin to differ. Children's paintings become representational (face, body, tree) and lose the abstract, pattern-like quality. Apes continue to grow, but their paintings remain the same. Although their abstract patterns are symbolic representations of the world around them, they are not called *works of art*. A work of art "implies a complete symbolic world, which can exist independently of the organism's own body, but which belongs, as it were, neither to the organism itself nor to the existing external reality" (FULLER 1983, p11). Chimpanzees may recognize the bird in their abstract representation, but it would seem that they do not *know* that they are aware of this knowledge. Such knowledge is reserved only for humans and it is generally known as self-consciousness or self-awareness.

Self-consciousness is best defined as the ability to describe oneself to oneself as a self. It requires complex cognitive functioning, and it has developed with increased intellectual powers and more refined social interactions. It is fundamental in the development of self-identity and culture which, in humans, allows for the development and refinement of the genetic predispositions for language, attachment, or socialization. It is this awareness that allows for the rich symbolism to be represented and transformed into that third dimension that is neither purely subjective nor purely objective, into the work of art.

This self-awareness or recognition of oneself as both part of the group and as a separate entity must have reinforced, early in human history, the *need to be different*. It is not difficult to imagine that, for a small group of hunters/gatherers, differentiating tools and weapons would have minimized possession conflict among the group. In case of unexpected attacks, weapon distinction would have secured efficiency and increased the group coordination so necessary for survival. Clothing used by humans in climates that do not require protection from the elements could have originated to define hierarchies in the social structure, to enhance individual differences, and to allow successful communication between and among friendly or unfriendly tribes. There is no better way to signal status immediately, and from a distance, than with attire; identify the enemy's arrows from your own by giving them a specific mark; express affection with the gift of a special stone. The need to distinguish oneself would have been most effectively achieved by *specialness* (DISSANAYAKE 1988), by elaborating and making things and oneself more than the everyday and the ordinary. In contemporary Western societies, we use a full plethora of artistic expressions in the service of self-definition. We surround ourselves with symbols that reflect status and wealth (e.g., expensive cars, homes, art, and cloths). Men and women spend a great deal of effort and resources on bodily adornments and displays that often only serve the purpose of showing status and wealth (LOW 1979). "People like to modify their bodies. We paint our faces, pierce nose and ears, circumcise the penis, enclose the neck, feet or waist in confining rings or shoes or corsets, bleach or brown or tattoo or even carve our flesh. Often the entire body is the focus of modification" (CASSIDY 1991). SANDERS (1989) has discussed the art of tattooing as a manipulation of one's self-image with a higher purpose of defining how an individual will be treated by the society. Tattooing, he argues, is a mark of both alienation from the mainstream and belonging to a unique group membership. In a similar vein, SMUTS (1992) has maintained that females' pursuit of thinness—achieving an aesthetically preferred image of attractiveness—is a strategy adopted by modern women to show status and wealth as well as to secure mates from higher social strata. There is, in principle, no reason to suspect that all of our efforts aimed at artistically expressing ourselves through our bodies, our possessions or our lifestyles could not serve a single purpose of promoting our understanding of who we are and where we stand in relations to the rest of the social group.

Therefore it can be hypothesised that art, as behaviour, has been shaped by natural selection to solve the problem of effectively differentiating oneself from the group. Promotion of self-identity could have facilitated the definition of hierarchies in the social structure, redefined sexual competition and secured smooth group coordination.

In summary, the phylogenetic development of art as a behaviour can be traced through the development of the aesthetic experience and the artistic expression. The aesthetic experience is rooted in basic physiological and psychological processes common to all animals. It relates to an experiential response of pleasure in the presence of familiar, recognizable stimuli. A complex aesthetic experience is only typical of the human species. It requires cognitive facilities that allow the incorporation and understanding of complex symbolic relationships. The artistic expression represents a combination of four distinct developmental adaptations: ritual, play, imagination and self-awareness. From an evolutionary perspective, art has been shaped by natural selection to promote individual identity through making things special and out of the ordinary.

Comparisons to Other Theories

PECKHAM (1973), in his book "Man's Rage for Chaos", was among the first to discuss artistic activity as a behavioural adaptation. He proposed a physiological basis for artistic behaviour, although he was unable to find any sound physiological evidence to support his claim. He postulated that humans possess a basic drive towards order, which is manifested in a continuous desire to perceive the environment in a comprehensible manner. This desire for order stems from a need to anticipate future events. In every new situation, humans bring 'an orientation', which represents a predisposition to filter relevant from irrelevant information from any perceptual encounter. This tendency to select only certain information leads to a problem of eliminating stimuli that in the future may become useful and relevant. For that reason, constant modifications of the orientation are necessary for successful adaptation. "Thus arises the paradox of human behaviour: the very drive to order which qualifies man to deal successfully with his environment disqualifies him when it is to his interest to correct his orientation" (p. 11). This paradox is resolved by artistic perception, which can be interpreted to imply greater sensitivity and flexibility to the environment and serves to relax and hinder the domineering drive for order.

Therefore, according to PECKHAM, artistic behaviour can be reduced to a biological mechanism. His approach represents an attempt to link a basic mechanism that is essential for survival with an artistic predisposition, thus arguing for universality of the artistic behaviour. However, reducing art to a basic defence structure hardly explains the complexity of artistic experience and expression. This somewhat FREUDIAN approach to biological drive and defence mechanisms raises a number of problems.

Drive towards order is not a uniquely human experience. It is common to most mammals. The necessity of this mechanism is obvious. Coding of familiar surroundings is essential in detecting danger when surroundings are changed. Such a mechanism has an essential survival value. Breaking up of the tendency for comprehension and consistency by perceiving the environment in new and unusual ways is also adaptive in that it allows one to explore novel solutions to problems when old ones are no longer effective. In that sense, artistic perception should not be an exclusively human endeavour, be related to the aesthetic feeling or be restricted to art. Order also relates to non-aesthetic symbols and experiences. Other behaviours, such as play, fantasy, dreams and rituals, make use of the anti-habitual quality of artistic perception (DISSANAYAKE 1983). Therefore, if artistic perception is not unique to art, it is difficult to explain why was art selected for as a separate behaviour.

If one is to use the concept of drive to explain artistic behaviour, it is perhaps much closer to both FREUD and DARWIN to use sexual drive as a link to artistic expression. From a DARWINIAN perspective, artistic expression can be explained as a behaviour exhibited to attract the opposite sex (by elaborate displays, ornamentations or singing). From a FREUDIAN perspective, art which abounds with latent sexual themes (naked bodies, heterosexual rendezvous) can be explained as unconscious symbolic expressions of the sexual drive or the Oedipal complex (COMFORT 1962).

Departing from the notion of a basic drive, ALLAND (1977, 1989) has argued that art has developed progressively through evolution to serve as a sophisticated communication device. He acknowledges that essential in the development of art are exploratory behaviour and play, preference for particular forms and shapes, perceptual discrimination associated with greater memory storage, and transformation-representation. The first three adaptations we share with other animals. The last one is exclusive to our species.

Play and exploration in animals is functional in that allows the young to learn the necessary skills for hunting and survival. In human babies, play is less functional and more pursued for its own sake, for which reason it is often related to art. It is self-rewarding, and allows for play acting and testing, in a non-threatening environment, new ideas, unconscious desires and imaginative approaches. Humans, as animals, show preferences for different patterns, shapes, special configuration and colours. Higher primates are even able to paint. However, ALLAND is quick to point out that "painting apes are not artists" (p. 32). Humans differ because they exhibit an extraordinary ability to store minute details about their environment, have sizable storage for faces, and have a greater mnemonic capacity to store information.

Also unique to humans is the ability to speak. ALLAND links language to our ability to represent the world around us symbolically. Language allows for dissociation of the symbol from the object, and it permits temporal flexibility (we can think about the past and the future). It also grants transformation of symbols, which is most vividly illustrated by the use of metaphors. ALLAND calls this process of transformation and representation of symbols 'transformation-representation' and he hypothesises it to be the basis for artistic activity because both creativity and innovation are dependent on this process.

Therefore, art, with its roots in biology, has bloomed in our species with the development of the transformation-representation process. It is a communication device between the artist and the viewer, or the symbol maker and the receiver. The success of the communication, as in language, will depend on purely cultural grounds. ALLAND also recognizes, but does not pursue in detail, the experiential side of art. Art arrests and seduces the viewer, he says, by evoking an aesthetic experience. This emotional experience is unique to art and can be either primitive (natural) or complex (man made).

What is problematic in ALLAND's theory is the premise that art represents yet another communication device. It is unlikely that art could have been selected for as a communication tool, for two simple reasons. First, it is much less effective than language. In fact, in most instances, it is unclear and even ambiguous. Second, more effective communication is achieved through non-aesthetic communication.

An alternative model has been proposed by KNIGHT and his colleagues (1995) in which the authors define the nature of the 'symbolic revolution', or the blossoming of human art, dance, singing and ritual that occurred among the human species approximately 50,000 years ago. They eloquently trace this revolutionary event through a series of genetic and cultural changes that took place early in human history. According to the authors, the increased cost on females, which resulted from the increase in brain-size of infants and escalating demand on the mother to look after her offsprings, required an increase in male investment. One strategy adopted by females to counter male philandering tactics prevalent at that time was to conceal any reliable signals of fertility. By concealing the time of ovulation and loss of oestrus, the males would have no way of judging when the female could be impregnated. Therefore, a male would be more likely to remain with a single female to insure fertilisation. This camouflaging of ovulation coupled with continuous receptivity by females would insure prolonged male courtship and increased parental investment.

Another strategy adopted by females to counter male philandering tactics was cycle synchrony. When the onset of menstruation is synchronised among the females in a single group the risk of male philandering is reduced, and since no single male can cope with guarding and impregnating a whole group of females, new male members are attracted into the group and become available to the potentially fertile females.

Once cover up of any fertility signals was achieved, menstrual bleeding was the only signal left that was indicative of imminent fertility among females. Since menstrual bleeding assured increased mating efforts from males, females that were not bleeding joined in the 'collective deception' with other females by painting themselves with red pigments to signal imminent fertility. According to the authors, this early body painting is indicative of the 'symbolic revolution' or the root of the human symbolic representation in art, dance, singing and ritual.

The model has been both praised and criticised by others (see the comments section in KNIGHT et al. 1995). It has been praised for its ability to incorporate a variety of available archaeological data into a coherent model, and it has been criticised for its "feminist perspective" that supposes changes based on female "strike" and female "conspiracy".

In the context of the current discussion it is relevant to note that the theory *mainly* describes a process—the emergence of art as a ritual—based on a series of *genetic changes* rather than cultural ones. Defining art as an expression of "ritual priorities" that "should include figures interpretable as menstruating, pregnant and/or nursing; as gender-ambivalent; and as pantomiming ' animal' courtship behaviour or attributes" (KNIGHT et al. 1995, p96) is perhaps adequate for defining the origins of the 'symbolic explosion', but a comprehensive model has to also account for the purpose of art continuously over the human history. For that, the effects of culture have to be considered.

Two other writers (DISSANAYKE 1988; FULLER 1983) trace the development of art as a behavioural adaptation in a very comprehensive and systematic manner. Essentially, they agree that art should be discussed in terms of two separate yet related issues: the aesthetic response (emotions), and the aesthetic expression (the work of art) through the emergence of play, ritual, imagination and self-awareness. They disagree about the functional role and significance art plays in the society today. FULLER, an art historian, strongly favours the position that in recent history art has lost its functional role and is now pursued for its own sake. In his essay, entitled "Art and Biology", FULLER (1983) proposes that art arose as a result of three phenomena in evolutionary history: the appearance of the aesthetic response, human ability to 'work on the external world', which he named culture, and the uniquely human skill to transform the reality through symbolism and imagination. The aesthetic response, FULLER maintains, has its roots in animal rituals. For him, ritualistic manifestations of emotions are the beginnings of symbolic expressions and serve purely a functional role. From an evolutionary perspective, biological changes often transform from purely functional to the symbolic. The aesthetic response, FULLER believes, moves in the same direction, although not always in a linear fashion.

The aesthetic expression (artistic activity), FULLER argues, is closely related to play. Play as an aesthetic response begins in animals by serving a functional role. In primates, however, we first see examples of non-functional play expressed in apes as dressing up, beating drums and even painting. In humans, as in primates, FULLER contends, play has lost its functional significance and *art* as play is now *non-functional and unnecessary*, therefore, pursued for its own sake. FULLER however, is very careful to distinguish between ape painting and a work of art. A work of art, he postulates, implies a world of symbolism that is detached from the physical reality of the self. It implies a process of symbolic transformation that is not bound to time or space, but it is dependent on *culture*. Culture "is an outgrowth of man's capacity for labile symbolization, and his ability to detach his symbols from himself into a third area of experience which is neither quite 'objective' nor quite 'subjective'" (p. 12). Two processes in evolutionary development made the appearance of symbolism, culture, and ultimately art, possible: Neoteny (slowing down of the growing process) and consequently, in humans, the long-time dependence on the mother. Due to the long period of immaturity, infants learn through interaction with the mother, to create illusions, a world of images, therefore learning to use dreams and transform symbolic reality. This rich world of imagination is transformed into material expression of art, such as painting or sculpture.

In essence, FULLER argues that art, although rooted in various functional behaviours such as play or ritual, has lost its functional role, and is pursued for its own sake. Although on the surface this assumption may seem correct, three questions remain unanswered. First, what is it about art that is so unique to secure its survival through millions of years of evolution? Second, if art is pursued for its own sake, as are play and ritual, what is unique about art to have been independently selected for? Third, why, in certain, less advanced societies, does art continue to be diversely expressed in paintings, decorations, music or dance?

In contrast to FULLER, Ellen DISSANAYAKE (1988) explores the evolution of art by asking the ultimate question "What is art *for*?" Her aim is to define art as behaviour and outline its functional role in human life. Her search for a common denominator in Western as well as primitive art has led her to conclude that art can be defined as the behaviour of 'making special'. 'Making special', according to DISSANAYAKE, relates to the human universal ability to *recognize* and *make* things special, or take things from the realm of the every day and make them extraordinary. She recognizes that art (objects and activities of recognized higher quality) can not be ultimately reduced to 'making special'; she proposes the notion as a starting point in the investigation of universals in 'the arts' in general.

Similarly to others (ALLAND, FULLER) she acknowledges that two other human behaviors, ritual and play, are directly related to art. In fact, DISSANAYAKE speculates that, at some point, all three were intimately associated, only to emerge later in evolution as separate entities. Unique to DISSANAYAKE is the recognition that all three behaviours share the propensity to 'make things special', a tendency which has evolved from the human need to elaborate and in some situations create an expression or experience that is out of the ordinary. The need to make things special, DISSANAYAKE argues, has its roots in our biology, and it has evolved, like complex bird songs and rituals, from the need to demonstrate more persuasively and efficiently the seriousness of what is being presented. Delivering a message in a different, out-of-the-ordinary way, has two benefits: it reinforces the behaviour, and it increases the likelihood of the message being noticed. The behaviour of 'making special, according to DISSANAYAKE, was preceded by the ability to recognize specialness. This ability, once again, is not very different from the tendency in many animals to notice and explore novel objects and environments. As in animals, for humans, this ability, similar to novelty seeking, has a great adaptive benefit.

Defining art in terms of 'making special' has the advantage of incorporating all of the heterogeneous classes comprising the category art. Making special can apply to traditional art as it can apply to car making or cooking. Because of its overinclusiveness, one can argue that it is difficult, on the basis of these criteria, to distinguish a work of art from a diesel engine. And that is exactly the point. From an evolutionary perspective the behaviour of making art must be broad enough to account for its diversity. Another, more important question, however, still remains, "Why are we compelled to make things special?" For DISSANAYKE, this behaviour results from our need to more persuasively and effectively deliver a message. However, explaining the functional significance of art in those terms hardly differentiates animals from humans. In fact, it is difficult to see the difference between the peacock's feathers and an Armani suit. Even more importantly, it is difficult to explain why are we constantly compelled to deliver *the same message* in variety of ways. It is at this very point that the new model diverges dramatically from the existing theories. It provides a more plausible explanation for the development and continuous facilitation of artistic behaviour and expression. I have argued that art, as behaviour, has been shaped by natural selection to solve the problems of effectively relating to, and differentiating from, the rest of the group. It is this need to understand who we are and how we relate to the rest of our society that fuels artistic expression. Identity is not static but an always-evolving process so necessary for our relationship to the always evolving culture. We are in a constant need to invent and re-invent ourselves in relationship to the group that is our reference point. The promotion of self-identity facilitated in the past, and continues to promote, the definition of social structure, redefine sexual competition and secure smooth group co-ordination. For that reason, aesthetic expression in the form of art will remain an eternal and integral part of human behaviour.

In conclusion, the new model shares a number of similarities and differences with the five existing theories. It diverges significantly from PECKHAM's theory, which reduces art to the realm of basic biological drive, by elaborating on the development of artistic behaviour through other behavioural adaptations such as ritual, play, imagination and self-awareness. It deviates from the model proposed by KNIGHT and colleagues by recognising that the development of artistic behaviour should be explained by a series of cultural changes rather than *mainly* genetic once. It is similar to the other three authors in tracing the behaviour of art in a similar fashion through the above-mentioned behavioural adaptations. It differs from ALLAND in two ways: 1) by clearly defining the boundaries between the behaviour of art and the experience in art, and 2) by arguing that the main adaptive role of art is not to serve as a communication device. The main contrast with FULLER is to contend that art *does* serve a purpose, and it is not pursued for its own sake. The basic contradiction with DISSANAYKE is to argue more convincingly that art has not developed only to more effectively deliver a message but serves a unique purpose: it is very useful in promoting self-identity, by allowing a clear differentiation of one's self from the rest of the human species.

Acknowledgements

I owe a debt of gratitude to Irwin SILVERMAN for inspiring me and encouraging me to write this paper. Thanks are also due to Dennis KINNEY and Bill THOMPSON for their helpful comments on the earlier versions of the manuscript; and Sharon TRAMER for her help in editing.

Author's address

Elizabeth Ralevski, McLean Hospital, 115 Mill Street, Belmont, MA 02478, USA.
Email: ralevski@mclean.harvard.edu

Note

1 Although in this paper the discussion on aesthetics will mostly refer to fine art, and visual art in particular, there is no reason to expect that the universal rules of aesthetic experience and expression are any different for other forms of artistic expression.

References

Abrams, M. H. (1984) The Correspondent Breeze: Essays on English Romanticism. New York: W. W. Norton.

Adams, L. S. (1993) Art and Psychoanalysis. Harper Collins Publishers: New York.

Adler, M. J./Gorman, W. Eds (1980) The Great Ideas: A Syntopicon of Great Books of the Western World. Encyclopedia Britannica Inc: Chicago.

Alland, A. Jr. (1977) The aesthetic sense. In: Alland, A. (ed) The Artistic Animal: An Inquiry into the Biological Roots of Art. Anchor Books: New York.

Alland A. Jr. (1989) Affect and aesthetics in human evolution. The Journal of Aesthetics and Art Criticism 47: 1–14.

Anderson, J. L./Crawford, C. B./Nadeau, J./Lindberg, T. (1992) Was the Duchess of Windsor right? A cross-cultural review of the socioecology of ideals of female body shape. Ethology and Sociobiology 13: 197–227.

Arnheim, R. (1992) But is it science? In: Cupchik, G. C./ Laszlo, J. (eds) Emerging Vision of the Aesthetic Process: Psychology, Semiology, and Philosophy. Cambridge University Press: Cambridge.

Baltisse, R./Ostermann, B. M. (1999) Are the dimensions underlying aesthetic and affective judgement the same? Empirical Studies of the Arts 16: 97–113.

Berleant, A. (1986) Experience and theory in aesthetics. In: Mitias, M. (ed) Possibility of the Aesthetic Experience. Martinus Nijhoff: Dordrecht.

Berlyne, D. E. (1971) Aesthetics and Psychobiology. Appleton-Century-Crofts: New York.

Boring, E. G. (1950) A History of Experimental Psychology (2nd Edition). Appleton-Century Croft: New York.

Boselie, F. (1992) The golden section has no special aesthetic attractivity. Empirical Studies of the Arts 10: 1–18.

Boselie, F.,/Cesaro, A. (1994) Disjunctive ambiguity as a determinant of the aesthetic attractivity of visual patterns. Empirical Studies of the Arts 12: 85–94.

Boynton, R. M.,/Olson, C. X. (1990) Salience of chromatic basic colour terms confirmed by three measures. Vision Research 30: 1311–1317.

Brown, P. J. (1991) Culture and the evolution of obesity. Human Nature 2: 31–57.

Brunius, T. (1967) The uses of works of art. In: Beardsley, M. C./Schueller, H. M. (eds) Aesthetic Inquiry: Essays on Art Criticism and the Philosophy of Art. Dickenson Publishing Company Inc: Belmont.

Buss, D. M. (1994) The Evolution of Desire. Basic Books: New York.

Cassidy, C. M. (1991) The good body: When big is better. Medical Anthropology 13: 181- 213.

Comfort A. (1962) Darwin and the Naked Lady. George Braziller: New York.

Coss, R. G. (1970) The perceptual aspects of eye-spot patterns and their relevance to gaze behaviour. In: Hutt, C./Hutt, S. J. (eds) Behaviour Studies in Psychiatry. Pergamon Press: Oxford, pp. 121–147.

Cunningham, M. R. (1986) Measuring physical attractiveness: Quasi-experiments on the sociobiology of female facial beauty. Journal of Personality and Social Psychology 50: 925–935.

Davies, S. (1991) Definitions of Art. Cornell University Press: Ithaca.

Davis, F. C. (1933) Aesthetic proportion. American Journal of Psychology 45: 298–302.

Davis, S. T./Jahnke, J. C. (1991) Unity and the golden section: Rules for aesthetic choice? American Journal of Psychology 104: 257–277.

Darwin, C. (1979) The Origin of Species. In: Pippleman, P. (ed) Darwin. W. W. Norton & Company Inc: New York.

De La Croix, H./Tansey, R. G./Kirkpatrick, D. (1991) Gardner's Art Through the Ages. Harcourt Brace Jovanovich, Inc: New York.

Dewey, J. (1934) Art as Experience. Capricorn: New York.

Dissanayake, E. (1988) What is Art for? University of Washington Press: Washington.

Dörner, D./Vehrs, W. (1975) Ästhetische Befriedigung und Unbestimmtheitsreduktion. Psychological Review 37: 321–334.

Dufrenne, M. (1973) The Phenomenology of Aesthetic Experience.Northwestern University Press: Evanston.

Eibl-Eibesfeldt, I. (1989) Human Ethology. Aldine de Gruyter: New York.

Eibl-Eibesfeldt, I. (1988) The biological foundation of aesthetics. In: Rentschler I./ Herzberger B./Epstein D. (eds) Beauty and the Brain: Biological Aspects of Aesthetics. Birkhäuser Verlag: Basel.

Eibl-Eibesfeldt, I. (1975) Ethology-The Biology of Behaviour. Holt, Reinhart and Winston: New York.

Enquist, M./Arak, A. (1994) Symmetry, beauty, and evolution. Nature 372: 169–172.

Everett, S. (1991) Art Theory and Criticism: An Anthology of Formalist, Avant-Garde, Contextualist and Post-Modernist Thought. McFarland & Company Ltd: Jefferson.

Farnsworth, P. R. (1932) Preference for rectangles. Journal of General Psychology 7: 479–481.

Fauss, R. (1988) Zur Bedeutung des Gesichts für die Partnerwahl. Homo 37: 188–201.

Fechner, G. T. (1871) Zur Experimentalen Aeshetik. Hirzel: Leipzig.

Fechner, G. T. (1976) Vorshule der Aesthetik. Breitkopf und Hartel: Leipzig.

Frith, S. (1996) Music and identity. In: Hall, S./DuGay, P. (eds) Questions of Cultural Identity. Sage Publications Inc: London, pp. 108–127.

Fuller, P. (1983) Art and Biology. In: Fuller, P. (ed) The Naked Artist: 'Art and Biology' & Other Essays. Writer and Readers Publishing Cooperative: London.

Gehlen, A. (1940) Der Mensch, seine Natur und seine Stellung in der Welt. Athenaeum: Frankfurt.

Gombrich, E. H. (1979) The Sense of Order: A Study of the Psychology of Decorative Art. Phaidon: Oxford.

Green, C. D. (1995) All that glitters: A review of psychological research on the aesthetics on the golden section. Perception 24: 937–968.

Hamlyn, D. W. (1957) The Psychology of Perception: A Philosophical Examination of Gestalt Theory and Derivative Theories of Perception. Routledge & Kegan Paul Ltd: London.

Hekkert, P./Peper, C. E./Van-Wieringen, P. C. W. (1994) The effect of verbal instructions and artistic background on the aesthetic judgement of rectangles. Empirical Studies of the Arts 12: 185–203.

Hekkert, P./Sneilders, H. M. J. J. (1995) Prototypicality as an explanatory concept in aesthetics: A reply to Boselie (1991). Empirical Studies of the Arts 13: 149–160.

Hoge, H. (1995) Fechner's experimental aesthetics and the golden section hypothesis today. Empirical Studies of the Arts 13: 131–148.

Hospers, J. (1976) Meaning and Truth in the Arts. University of North Carolina Press: Chapel Hill.

Humphrey, N. K. (1973) The illusion of beauty. Perception 2: 429–439.

Humphrey, N. K. (1980) Natural aesthetics. In: Mikellides, B. (ed) Architecture for People. Studio Vista: London.

Hutchins, R. M. (1980a) Locke, Berkeley, Hume. William Benton Publishing: Chicago.

Hutchins, R. M. (1980b) Plato. William Benton Publishing: Chicago.

Hutchins, R. M. (1980c) Kant. William Benton Publishing: Chicago.

Huxley, J. (1914) The Courtship Habits of the Great Crested Grebe (Podiceps cristatus) with an Addition to the Theory of Sexual Selection. Proceedings of the Zoological Society of London, pp. 491–562.

Itten, J. (1961) Kunst der Farbe. Subjektives Erleben und objektives Erkennen der Wege zur Kunst. Maier: Ravensburg.

Kagan, A. (1983) Some reflections on the biological significance of style. Arts Magazine 58 (3): 118–121.

Kenrick, D. T./Keefe, R. C. (1992) Age preferences in mates reflect sex differences in human reproductive strategies. Behavioural and Brain Science 15: 75–91.

Kettlewell, N. (1988) An examination of preferences for subject matter in paintings. Empirical Studies of the Arts 6: 59–65.

Knight, C./Power, C./Watts, I. (1995) The human symbolic revolution: A Darwinian account. Cambridge Archaeological Journal 5: 75–11.

Kogan, N. (1994) On aesthetics and its origins: Some psychobiological and evolutionary considerations. Social Research 61: 139–165.

Langlois, J. H./Ritter, J. M./Roggman, L. A./Vaughn, L. S. (1991) Facial diversity and infant preferences for attractive faces. Developmental Psychology 27: 79–84.

Levy, J. (1976) Lateral dominance and aesthetic preference. Neurophychologia 14: 431 -445.

Levy, J. (1988) Cerebral asymmetry and aesthetic experience. In: Rentschler I./ Herzberger, B./Epstein, D. (eds) Beauty and the Brain: Biological Aspects of Aesthetics. Birkhäuser Verlag: Basel.

Levy, J./Heller, W./Banich, M./Burton, L. (1983) Are variations among right-handed individuals in perceptual asymmetries caused by characteristic arousal differences between hemispheres? Journal of Experimental Psychology: Human Perception and Performance 9: 329–359.

Lind ,C. (1993) Psychology of colour: Similarities between abstract and clothing colour preferences. Clothing & Textiles Research Journal 12: 57–65.

Magro, A. M. (1999) Evolutionary-derived anatomical characteristics and universal attractiveness. Perceptual and Motor Skills 88: 147–166.

Marshall, P. H./Worthen, J./Brant, L./Shrader, J./Kahlstorf, D./Pickeral, C. (1995) Fechner redux: A comparison of the Holbein Madonnas. Empirical Studies of the Arts 13: 17–24.

Martin, F. (1995) The horse and the cart: the contemporary artist and the aesthetician. Empirical Studies of the Arts 13: 119–130.

Martindale, C./Moore, K./Borkum, J. (1990) Aesthetic preference: Anomalous findings for Berlyne's psychological theory. American Journal of Psychology 103: 53–80.

McLaughlin, J. P./Murphy, K. E. (1994) Preference for profile orientation in portraits. Empirical Studies of the Arts 12: 1–7.

McManus, I. C./Cheema, B./Stoker, J. (1993) The aesthetics of composition: a study of Mondrian. Empirical Studies of the Arts 11: 83–94.

Messinger, S. M. (1998) Pleasure and complexity: Berlyne revised. Journal of Psychology 132: 558–560.

Morris, D. (1962) The Biology of Art. Methuen: London.

Orians, G. H./Heerwagen, J. H. (1992) Evolved responses to landscapes. In: Barkow, J. H./Cosmides, L./Tooby, J. (eds) The Adapted Mind: Evolutionary Psychology and the Generation of Culture. Oxford University Press: Oxford.

Osborne, H. (1970) The Art of Appreciation. Oxford University Press: Oxford.

Osborne, H. (1970a) Aesthetics and Art Theory. E. P. Dutton: New York.

Parsons, C. (1992) The transcendental aesthetics. In: Guyer, P. (ed). The Cambridge Companion to Kant. Cambridge University Press: Cambridge.

Paul, G. (1988) Philosophical theories of beauty and scientific research on the brain. In: Rentschler, I./Herzberger, B./Epstein, D. (eds) Beauty and the Brain: Biological Aspects of Aesthetics. Birkhäuser Verlag: Basel.

Peckham, M. (1973) Man's Rage for Chaos. Schocken Books: New York.

Piehle, J. (1976) The "goldern section": An artifact of stimulus range and demand characteristics. Perceptual and Motor Skills 43: 47–50.

Pierce, E. (1894) Aesthetics of simple forms. Psychological Review 1: 483–495.

Polivy, J./Garner, D. M./Garfinkel, P. E. (1986) Causes and consequences of the current preference for thin female physiques. In: Herman, C. P./Zanna, M./Huggins, E. T. (eds) Physical Appearance, Stigma, and Social Behaviour. Hillsdale: Erlbaum, pp. 89–112.

Poppel, E./Sutterlin, C. (1983) Wahrnehmungsasymmetrie und Gestaltung von links und rechts in Bildern. Neuropsychologische Aspekte der ästhetischen Wahrnehmung. Symposium "Biology of Aesthetics" Reimers-Stiftung: Bad Homburg.

Purcell, A. T./Lamb, R. J./Peron, E. M./Falchero, S. (1994) Preference or preferences for landscape? Journal of Environmental Psychology 14: 195–209.

Ree, J. (1990) Funny voices: stories, "punctuatuin" and personal identity. New Literary History 21: 1055.

Regard, M./Landis, T. (1988) Beauty may differ in each half of the eye of the beholder. In: Rentschler, I./Herzberger, B./ Epstein, D. (eds) Beauty and the Brain: Biological Aspects of Aesthetics. Birkhäuser Verlag: Basel.

Rensch, B. (1963) Versuche über menschliche Auslösermerkmale beider Geschlechter. Journal of Morphological Anthroplogy 53: 139–164.

Roelofs, M. (1998) Aesthetic experiences and their place in the mind. Dissertation Abstracts International 58: 4300.

Saito, M. (1994) A cross-cultural study on colour preference in three Asian cities: Comparison between Tokyo, Taipei and Tianjin. Japanese Psychological Research 36: 219–232.

Samuels, C. A./Butterworth, G./Roberts, T./Graupner, L. et al. (1994) Facial aesthetics: Babies prefer attractiveness to symmetry. Perception 23: 823–831.

Sanders, C. R. (1989) Customizing the Body: The Art and Culture of Tattooing. Temple University Press: Philadelphia PA.

Sass, L. A. (1999) Psychoanalysis, romanticism, and the nature of aesthetic consciousness, with reflections on modernism and postmodernism. Psychoanalytic Review 85: 717–746.

Schneider, J. F./Krumov, K./Andrejeva, L./Kibarova, E. (1993) Authoritarian attitudes and aesthetic preferences: A Bulgarian replication. Perceptual and Motor Skills 77: 255–258.

Seid, R. P. (1989) Never too Thin: Why Women Are at War with their Bodies. New York: Prentice Hall.

Silver, N. C./Ferrante, R. (1995) Colour preferences. Perceptual and Motor Skills 80: 920–922.

Singh, D. (1993) Adaptive significance of female physical attractiveness: Role of waist- to-hip ratio. Journal of Personality and Social Psychology 65: 293–307.

Smith, B. (1988) Foundations of Gestalt Theory. Druckhaus Berltz: Hemsbach.

Smuts, R. W. (1992) Fat, sex, class, adaptive flexibility and cultural change. Ethology and Sociobilogy 13: 523–542.

Stolnitz, J. (1986) The actualities of non-aesthetic experience. In: Mitias, M. (ed) Possibility of Aesthetic Experience. Martinus Nijhoff: Dordrecht.

Storr, A. (1972) The Dynamics of Creation. Secker & Warburg: London.

Taylor, C. W. (1988) Various approaches to and definitions of creativity. In: Sternberg, R. J. (Ed) The Nature of Creativity. Cambridge University Press: New York.

Uduehi, J. (1995) A cross-cultural assessment of the mainland graves design judgement test using U. S. and Nigerian subjects. Visual Art Research 21: 11–18.

Ulrich, R. (1986) Human response to vegetation and landscape. Landscape and Urban Planning 13: 29–44.

Weber, C. O. (1931) The aesthetics of rectangles and theories of affection. Journal of Applied Psychology 15: 310–318.

Weiner, J. S. (1971) Man's Natural History. Weidenfeld & Nicolson: London.

Wertheimer, M. (1967) Gestalt theory. In: Elliss, W. D. (ed) A Source Book of Gestalt Psychology. Humanities Press Inc: New York.

Wild, T. C./Kuiken, D. (1992) Aesthetic attitude and variations in reported experience of a painting. Empirical Studies of the Arts 10: 57–78.

Winner, E. (1991) The puzzle of art. In: Feinstein, S.C./Esman, A. H. et al. (eds) Adolescent psychiatry: Developmental and clinical studies. University of Chicago Press: Chicago IL, pp. 148–159.

Winston, A. S./Cupchik, G. C. (1992) The evaluation of high art and popular art by naive and experienced viewers. Visual Art Research 18: 1–14.

Witmer, L. (1894) Zur experimentellen Ästhetik einfacher räumlicher Formverhältnisse. Philosophische Studien 9: 96–144.

Woods, W. A. (1991) Parameters of aesthetic objects: Applied aesthetics. Empirical Studies of the Arts 9: 105–114.

Wordsworth, W./Coleridge, S. T. (1963) Lyrical Ballads (1798 ed., with additional 1800 poems and prefaces). Methuen: London.

Zusammenfassungen der Artikel in deutscher Sprache

James Barham

Biofunktionaler Realismus und das Teleologieproblem

Im Bereich der Überlegungen zur biologischen Funktionalität stellt die Frage nach der Intentionalität einen besonderen Fall dar. Um Intentionalität zu „naturalisieren" muß auch die „biologische Funktionalität" einer Naturalisierung unterworfen werden, wobei letztere vor allem die Frage nach deren teleologischem Charakter aufwirft. Weder Kybernetik, noch natürliche Selektion können Teleologie naturalisieren, weil beide diese voraussetzen. Dahingehend ist die derzeit herrschende kybernetisch–selektionistische Weltsicht grundlegend unvollständig und das Problem der Teleologie nach wie vor existent.

Im vorliegenden Artikel wird eine alternative, dynamisch–emergentistische Weltsicht dargestellt, die sich auf unterschiedliche Bereiche gegenwärtiger Wissenschaften stützt. Diese als „biofunktionaler Realismus" (*biofunctional realism*) bezeichnete Sichtweise interpretiert den teleologischen und normativen Charakter biologischer Funktionalität als Manifestation einer bestimmten Organisationsform von Materie und stellt sich damit gegen die funktionalistische Sichtweise einer vielfältigen Realisierbarkeit biologischer Funktionen.

Dementsprechend wird besonders dem Phänomen der Teleologie verstärkte Beachtung zuerkannt, wobei sich herausstellt, daß diese aus zwei unterschiedlichen, jedoch eng aufeinander bezogenen Merkmalen besteht: Begehren (*conation*) und Kognition. Diese Merkmale bilden die Hauptbestandteile eines Modells biologischer Funktionalität, welches sich auch auf Konzepte der nicht-linearen Dynamik stützt.

Winfried Behr

Homo Ridens: Eine Spekulation über Ursprung und biologische Funktion von Humor

Auslachen, Lachen und Lächeln sind funktional verwandte Verhaltensweisen mit teilweise unterschiedlichem evolutionären Ursprung.

Auslachen ist das älteste Phänomen. Es begann als eine Äußerung von Tadel und Mißbilligung ungewöhnlichen Verhaltens eines Individuums durch die Gruppe (*mobbing call*). Ausschluß des getadelten Mitglieds verbesserte die Effizienz von Gruppenunternehmungen und damit die individuelle Fitneß der verbleibenden Gruppenmitglieder. Auslachen hat Bezug zum Komischen. Kriterien des Komischen sind ungewöhnliches (dummes) Verhalten oder Aussehen, Glücklosigkeit (Pech) und Nachahmung.

Lachen hat Bezug zum Witz. Der Witz schafft komische Effekte mit verbalen Mitteln. Aggressionsopfer in Witzen sind nunmehr imaginär; Lachen als Folge von Witzigkeit ist deshalb – anders als Auslachen – kaum noch offen aggressiv. Voraussetzungen für das Vergnügen am Witz ist dessen Verständnis. Den Humor von Gruppen zu verstehen ist Bedingung für die Zulassung zu den Gruppen. Humor fördert so die intellektuelle Homogenität. Die Fitneß der Gruppenmitglieder gewinnt teilweise durch höhere Gruppeneffizienz und teilweise durch die Gelegenheit zu Sexualpartnerschaften mit Individuen gleicher intellektueller Fähigkeit (*assortative mating*).

Lächeln, als Reaktion auf Witzigkeit, ist entstanden als eine Kommunikation zwischen potentiellen Sexualpartnern, die damit intellektuelle Fähigkeiten demonstrieren. Es fördert Sexualpartnerschaft zwischen intellektuell gleichwertigen Partnern.

Ladislav Kováč

Fundamentale Prinzipien der kognitiven Biologie

„Kognitive Biologie" beabsichtigt eine Synthese von Daten unterschiedlicher Wissenschaften innerhalb eines einheitlichen Rahmens, wobei – in Übereinstimmung mit der „Evolutionären Erkenntnistheorie" – die biologische Evolution als Prozeß zunehmender Informationsakkumulation interpretiert wird. Dieses „Wissen" ist in den Konstruktionsplänen von Organismen verkörpert, wobei die strukturelle Komplexität dieser Konstruktionen deren epistemischer Komplexität entspricht. Im Gegensatz zur EE räumt die „kognitive Biologie" der molekularen Ebene einen zentralen Stellenwert hinsichtlich kognitiver Prozesse ein. Sie bedient sich dabei des

Prinzips der „minimalen Komplexität", welches von der Voraussetzung ausgeht, daß die zielführendste Art des Studiums von biologischen Phänomenen immer in der Untersuchung der einfachsten Organismen und Prozesse besteht, in denen das zu untersuchende Phänomen auftritt. Die erstaunlichen Parallelen zwischen physikalischen Prinzipien und Prinzipien der kognitiven Biologie weisen in Richtung auf eine enge Beziehung zwischen Ontologie und Erkenntnistheorie.

Hinsichtlich des menschlichen Bewußtseins bzw. bewußter emotionaler Erfahrungen wird das „Prinzip der Minimierung von Leiden" entwickelt, welches wie alle anderen Prinzipien der kognitiven Biologie deskriptiv und nicht normativ aufzufassen ist.

Wolfgang Patzelt

Institutionen als wissensgewinnende Systeme: Was können Sozialwissenschaftler von der Evolutionären Erkenntnistheorie lernen?

Dieser Artikel geht von der Fragestellung aus, was die Institutionenforschung bzw. die vergleichende Analyse politischer Systeme aus der Beschäftigung mit der Evolutionären Erkenntnistheorie lernen kann. In fünf Argumentationsschritten wird dies erläutert. Zunächst werden dabei einige Grundelemente der EE erörtert. Sodann wird gezeigt, daß soziale und politische Systeme ganz im Sinne der EE als erkenntnisgewinnende Systeme aufzufassen sind. Schließlich wird erklärt, wie Institutionalität als Lernergebnis und Geschichtlichkeit als Lernprozeß verstanden werden kann.

Den Abschluß bildet eine Konkretisierung dieser Ausführungen am Beispiel einer evolutionistisch–morphologischen Vergleichsanalyse von politischen Systemen.

Elizabeth Ralevski

Ästhetik und Kunst aus evolutionärer Perspektive

Für Philosophen und Kunsthistoriker waren universelle Prinzipien, welche der Kunst und der Ästhetik zugrunde liegen, über Jahrhunderte hinweg Gegenstand besonderen Interesses. Ähnlich umfassende Debatten befaßten sich – ausgehend von der historischen Entwicklung von Kunst – mit der Frage nach deren Funktion.

Diese Arbeit beabsichtigt einen Überblick über gegenwärtige experimentelle Forschungen hinsichtlich universeller ästhetischer Urteile zu geben und kommt dabei zu dem Ergebnis, daß elementare ästhetische Präferenzen kulturinvariant in allen Menschen auftreten.

Weiters stellt dieser Artikel ein neues Modell der phylogenetischen Entwicklung von Kunst vor und geht dabei von der Voraussetzung aus, daß die Kunst im Verlauf der Phylogenese deshalb einen zentralen Stellenwert einnahm, weil sie für die Bildung personaler wie auch kultureller Identitäten von besonderer Bedeutung war bzw. ist.