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.kla.univie.ac.at/ **Chairman:** Rupert Riedl **Managing Editor:** 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 and the Austrian Federal Ministry of Science, Research and Culture.**

Rupert Riedl

Cognition of Evolution: Can Causal Explanation Overrule Cognition?

Abstract

It is the purpose of this contribution to disentangle the processes of cognition and causal explanation, to investigate them epistemologically, and to uncover the consequences of their habitual entanglement.
In order to do so I shall have to reconstruct the major steps in the history of the discovery of evolution. For some of us this will probably be no more than a recapitulation. Yet in the current context it is necessary, for history tends to repress or distort historical facts to justify the present situation, to dispense itself from the responsibility of its recalcitrant paradigm, i.e., to shake off its contradictions.

Key words

Causal explanation, cognition, conjecture, evolutionary epistemology, experience, unterstanding.

I. Introduction and terminology

(1) *Can cognition be substituted* by explanation? How would this be possible? What could be explained without being known?! Yet, we will find that the process of cognition of the 'natural system' of living beings is composed of, and may be recognized by, such subphenomena as taxa, Bauplan, type, and homologies. After having been challenged by causal considerations, however, the conception of cognition was screened off by explanatory arguments, and, more specifically, by causal explanation.

(2) This investigation will be enabled by the *recognition of unconscious forms of problem solving*. This approach is supported by Evolutionary Epistemology (EE) in the sense of Popper (1957), Lorenz (1978), and Riedl (1984). EE has shown that an unconsciously acting mental capacity to recognize and solve simple, but life-supporting problems — a product of genetic learning and adaptation—is inherited. (Detailed literature will be cited.) The ways in which this capacity copes with complex systems are very differentiated; they cannot easily be pressed in the common terminology. This capacity cannot be changed; it can only be overcome by 'experience', in common terminology (Riedl 1995a).

(3) *A clarification of the terminology* is required to differentiate the processes of cognition and causal explanation. It is recommended always to check the illustrations. This will be the more obligatory as there is no literature dealing with this differentiation. Contemporary biologists do not seem to be aware of it. Badly enough, the common terminology has to be differentiated itself; I will try not to invent new terms, and not to distort the meaning of the common ones.

I talk of *cognition* in the sense of step-wise taking notice, becoming aware, and appreciating the lending weight to, the comparing, and the grouping of phenomena. Cognition starts from inherited physiological processes and accumulates *knowledge*. In contrast, I take *causal explanation* in the common sense of purposeful collection and adducing reasons.

The terms 'explanation' and 'understanding', coined by Dilthey (1883), have been used to differentiate how we grasp phenomena in the sciences from how we do so in the arts: Natural phenomena are to be explained, artifacts are to be understood. In common usage these terms have remained interchangeable. I take 'understanding' as the wider concept, viz., as the product of causal explanation, an operation which, as we shall see, moves back and forth between detection and explanation.

In both processes the observer depends on memory and observed cases that are provided by a highly redundant world, followed by the assumption that cases which he takes to be similar in some respects can be put into a category, which, after repetition, allows to make predictions. Repeated observation is facilitated by searching for comparable situations (also in memory and background knowledge) or by producing comparable situations in experimentation.

Both processes include a *conjecture* (in Popper's [1957] sense) that is based on the already known cases, an *expectation* that similar features can be sorted out from the dissimilar ones, such as the different

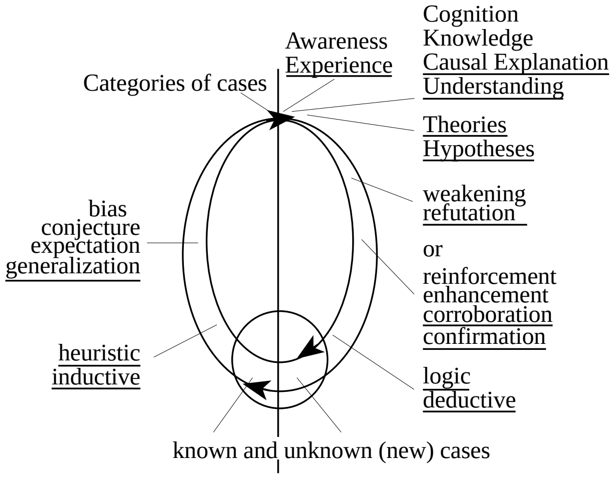

Figure 1: *General terminology* of cognitive processes. Operations dominated by consciousness are underlined.

times of observations. This corresponds to a heuristic or inductive attempt, driving towards a *generalization*, a conceptual grasping, a hypothesis, or a theory, according to the level allowed by the approach.

Such a preliminary *category of cases* will be tested on new cases. And the expected generalization can be either refuted by a new case or reinforced, enhanced, or confirmed. This process develops towards deduction and logic, in the sense that the boundaries of the expected generalization will be more and more consciously reflected and defined. We call the outcome *experience*. In occasions with more and with surrounding experience, the expectation will be adjusted, i.e., the boundaries of the generalization will be altered by excluding cases and putting them in another category (see Fig. 1).

I described the underlying inherited processes as the 'inborn hypotheses' of 'apparent truth' and 'comparability' (RIEDL 1984), and traced them back to their origins, to conditioned responses (RIEDL 1992), and Gestalt perception (RIEDL 1987). Both are interdependent and iterative processes drifting towards an optimum of confidence. And the two halfe-circles of the spin never turn back to their origin; they do not form a logical circle, because every case—every refutation or confirmation—alters the expectation (RIEDL 1984, OESER 1976). We know this rise of the spiral as the gain of experience, be it knowledge or understanding.

In neither process (in the systems of cognition and causal explanation) do the helical spins that gain experience remain isolated. On the contrary, their optimized generalizations or theories form the cases to develop a conjecture for a next level of generalization. Consequently, such generalizations ('laws') of theories are the basis for a superimposed generalization of theory. Together they form a pyramid or hierarchy of conjectures and experiences with the 'state of the art' on top. This principle may be likened to HEMPEL and OPPENHEIM's (1948) 'deductive-nomological' or subsumption schema for explanation, in the sense that detailed experiences and the observation of correlations are always required in order to describe and predict these in the framework of an over-arching correlation. One could also say that theories constructed out of observed cases become themselves cases of an over-arching theory (RIEDL, SPALT).

Later on, I have demonstrated (RIEDL 1985) that what is being formed is actually a double pyramid, with its bases at the mesocosmic level of our observations, and with its tips ending in the remotest theories concerned with the most comprehensive phenomena in the microcosm and megacosm. Both pyramids include branching *paths of experience*, running from the tips to the manifoldness of the observable phenomena of our world, and converging *paths of conjectures* in the other direction (Fig. 2). In this paper I will demonstrate that these patterns hold for both the system of causal explanation and that of cognition.

The path of conjectures in the system of cognition will be called the *path of perception*, because it is based on Gestalt perception and conditioning that form invariants by generalization within perceived 'fields of similarities' derived from comparable cases. The gain of experience, on the other hand, will be called *path of confirmation*, because it starts with neuronal and later cerebral memory, allowing reinforcement and enhancement, consciously experienced as confirmation. I call the product of the iteration of the two *gain of knowledge*, in the sense that dogs (and occasionally men as well) require knowledge about the way home, but no (deeper) understanding. As at the same time, understanding presupposes the possession of knowledge.

In the system of causal explanation, the path of conjectures will be called *path of detection*, because it is based on the detection of the right selection of assumed identical reasons by eliminating variants and fixed boundary conditions. The path of experience, on the other hand, will be called *path of explanation*, taken in its psychological sense, as we take the description of a correlation (as time and speed in 'free fall') as explained—if it can be—together with comparable phenomena (e.g., a trajectory), be deduced from a superimposed correlation (as mass and

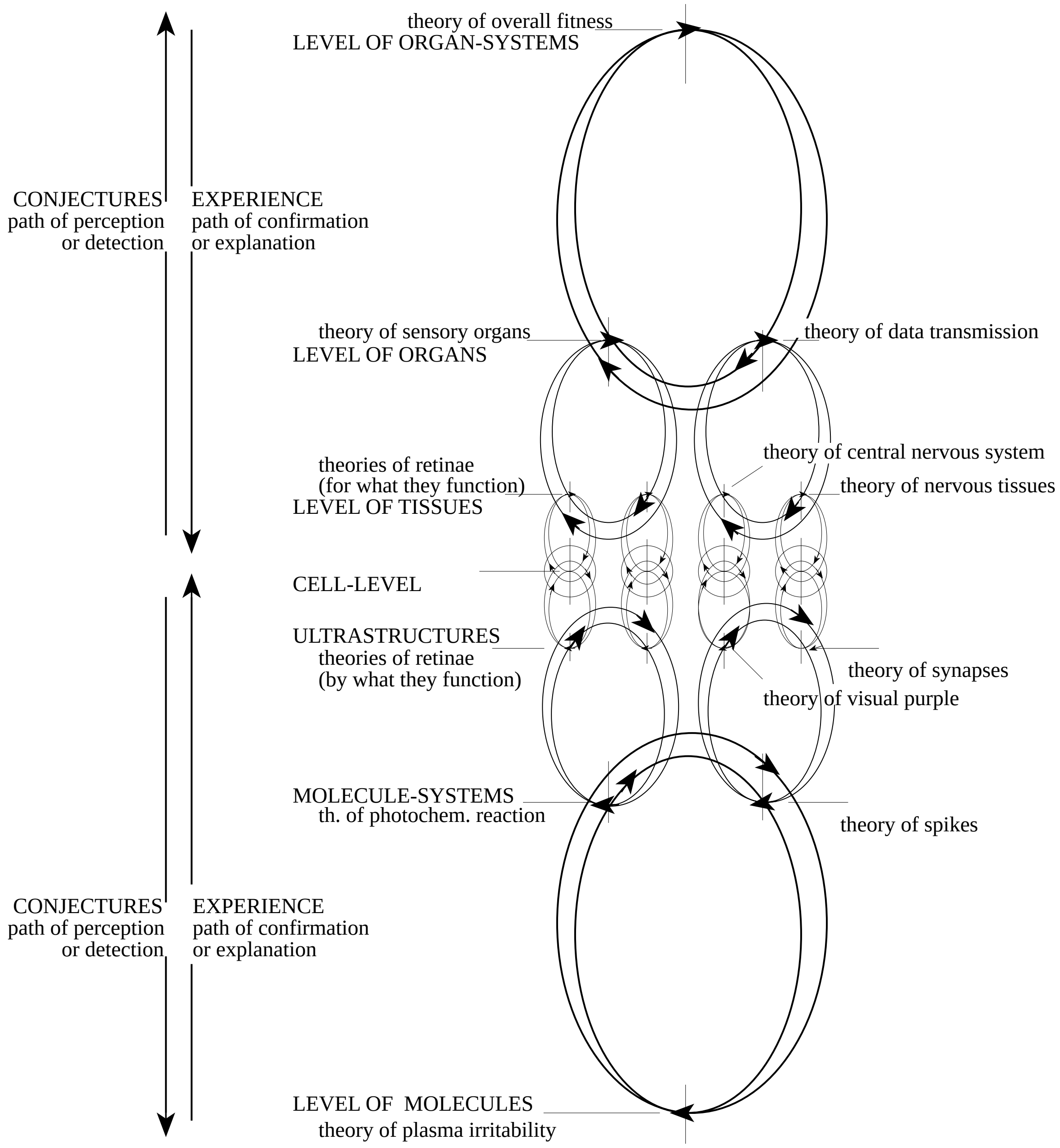

Figure 2: *The double pyramid*. Examples of theories concerned with the retina tissue; downwards to the molecular level, upwards to the level of the organism, related to the paths of conjectures and and experiences. To simplify the graphic, only two examples of subtheories are given at every level (symbols as in Fig. 1).

distance; named gravitation). I call the product of the two *gain of understanding* in the sense given above; a conscious process, presupposing knowledge, including the inspection of the process which produces this understanding (Fig. 3).

Our critical investigation requires that a term as common and well understood as *discovery* will be used in this context only when the gain of knowledge and understanding are taken together. The term *cognition* must be avoided for similar reasons (one may want to have Fig. 3 within reach to adjust to this differentiated terminology).

(3) Finally, *common features in cognition and causal explanation* will be assembled in order to contrast them with their differences.

Both methods depend, first, on a lawful, i.e., predictable, yet highly redundant world, the repetitions being not quite identical, and, second, on some kind of available memory, be it in a nervous system only or, in addition, in a culture.

Both operate by helical, iterative processes of observing cases, touring between induction and deduc-

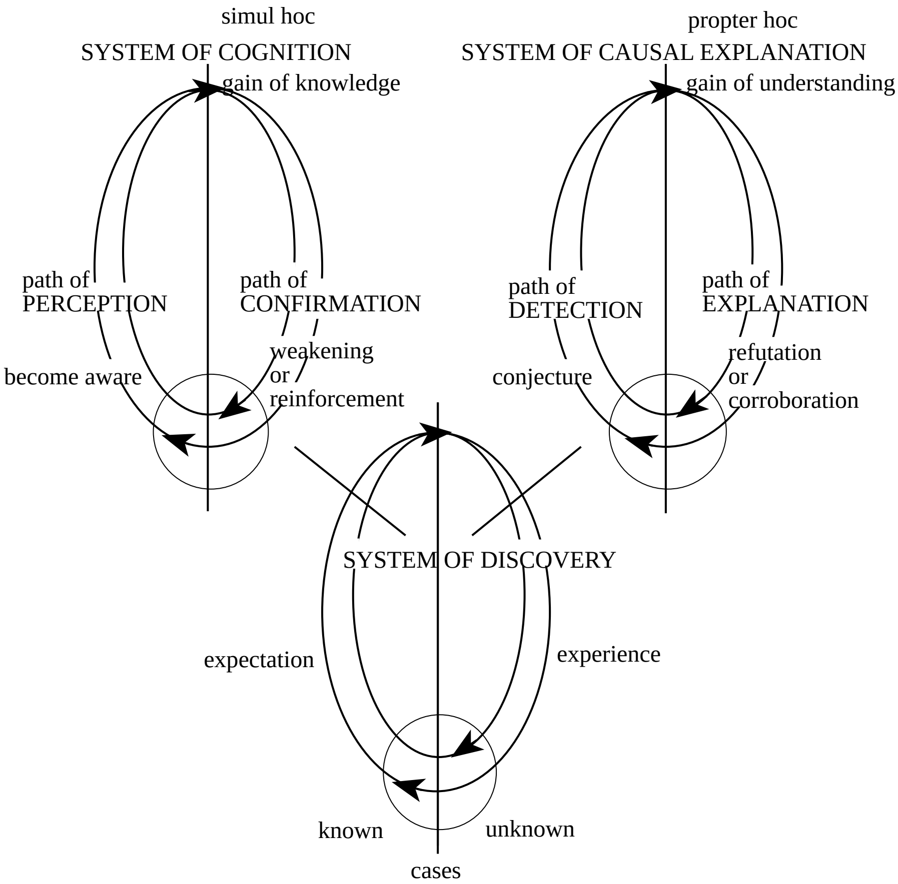

Figure 3: *The differentiated terminology*, divided into processes of cognition and processes of causal explanation and understanding (symbols as in Fig 1).

tion, conjectures of the individual and gaining experience from the environment, generalizing phenomena as well as groups of phenomena to which they may belong; forming, as we will see, structural and class hierarchies.

Both are controlled by exposing conjectures derived from experienced cases to new cases, optimizing the reliability of the expectation. The helicals never turn back to their origin, and the rise of the helical corresponds to the gain of experience, be it qua facts or qua reliability.

Together the two helicals form a double pyramid. Optimized expectations on one level provide the cases for the next one. Both pyramids are based on a level of basic observation, ending in theories concerned with phenomena of the microcosm and the megacosm. The helicals control one another mutually. Their extension and overall optimization correspond to the state of the art, the empirical truth that has been secured.

The combined parts of conjectures start from the basic observations and congregate towards the tips of the pyramids. The path of experience, in contrast, starts from those tips and ramifies toward the manifoldness of available observations (Fig. 2).

Both methods aim to enlarge and improve the predictability that is to be produced by a limited brain confronted with unlimited facts, for the simple reason that safe prediction supports life and improves its quality. In general, the two-sided paths of explanation correspond to the 'path of origination' of objects in this world, confirming the fact that all complex systems develop as insertions between constituents and an environment.

(4) *Cognition in contrast with causal explanation*. Despite the common features of both methods, which have to do with basic structures of the world and principles of man's knowledge-gaining apparatus, there are profound differences as well, which concern methods, results, and their estimated position in the sciences.

Methodologically, the two methods differ as our inherited problem-solving mechanisms differ from those on which consciousness, language, and culture have been superimposed (the conscious processes being based on the inherited ones). They differ, qua *post hoc*, as HUME (1748) defined the precondition of observations, which actually splits into two subsequent expectations: in *simul hoc* (a term coined here), standing for laws of simultaneity, and *propter hoc*, standing for the laws of causality. The inherited process is still grounded in observation, the second, as recognized by HUME and inspiring KANT, must be added as a speculation.

cognition	causal explanation
primarily inherited	culture and language dependent
predominantly unconscious	predominantly conscious
ratiomorphic[1]	rational
Gestalt perception	logical operation

still observable	to be added by speculation
cybernetic	probabilistic
recursive processes	mainly linear, 'if-then 'arguments
hermeneutic[2], morphological[3]	scientistic[4]
synthesizing systems	dissecting systems

Results differ correspondingly; as the predictability of the regularities of complex patterns differs from that of analyzed functions, as reliability grounded on synthesis of a manifoldness differs from reliability grounded on the elimination of variants, or as the luxuriant wealth of phenomena differs from its mechanistic skeleton. This is the more so since causal thinking to date has become restricted, as we will see, to take only one of the four forms of causes as scientifically acceptable: ARISTOTLE's *causa efficiens*, which can be best translated as 'power'.

cognition	**causal explanation**
patterns of order	patterns of 'natural laws'
constraints of forms of simultaneity	constraints of funct.(forces) of actions
simul hoc	propter hoc
as results of history and irreversible	as eternally given, and generally reversible
predominantly qualitative	predominantly quantitative
not easily formalizable	predominantly formalized
hard to rationalize	based on rational decisions

Rank and merits as given to both methods by the public also differ markedly. Causal explanation seems to operate with eternal laws—actually they are long-ranging ones (THIRRING/STÖLTZNER 1994), while the laws that govern orderly patterns of high complexity tend to have a shorter history. In this context, 'causal explanation' is commonly used to refer to products of conscious speculation. As a consequence, it can easily be made plausible. But in cognition it is the other way around: laws are grasped irrespective of their history. It is little known that laws of simultaneity can be as reliable as causal laws (RIEDL 1977a; cf. ELSTER 1983, p26 on non-causal laws such as the BOYLE-CHARLES law). Causal explanation in science must alter as soon as a change of cognition leads to different insights. In contrast, cognition, as we will find out, does not alter when explanation is changing. Only if the explanation is of a metaphysical type—say, a proof of the existence of God—is it the other way around. Causal research narrows its focus to fields in which the levels of complexity can be reduced sufficiently so as to allow formalization and experimental interaction; or, alternatively, to fields in which phenomena of corresponding levels may be traced back to the former, such as astrophysics and geophysics. Cognition, on the other hand, extends far into fields of high complexity where interaction is not explanatory. Experimentation interferes with nature (HACKING 1983), thus providing power over nature and, consequently, power for the experimenter and his sponsor. Hence experimentation is firmly supported wherever power might be gained. Cognition, in contrast, scarcely interferes with nature.

cognition	**causal explanation**
prerequisite	subsequent
independent of explanation	dependent on cognition
does not change in case of altered explanation	must change in case of altered cognition
challenges explanation	seems to replace cognition
scarcely allows experiments	strives for experimentation
extends to high complexity	narrows, reduces complexity
taken as only descriptive as prescientific	taken as sufficiently explanatory as the very science

This is the differentiated situation from where I start this investigation.

II. The history of cognition

In this section I will outline the more remarkable steps made towards the cognition of evolution. I will concentrate on the principles that are being considered, pointing out which concepts have been rationalized, made suspect, or unconsciously taken for granted. It attempts an epistemological investigation, not an organized epistemology; because it must follow the path knowledge has gained in history.

Surveying the centuries one finds that the path of cognition, at one time taken as a miracle, lost the interest of scientists, whereas the path of explanation became attractive. This may parallel a shift of interest from inductive and heuristic to deductive and logical procedures, a growing trust in rationality and analysis.

(1) *The prescientific situation* shows that 'primitive' cultures are equipped with a remarkably differentiated knowledge of taxonomy. They know exactly which animals are quadrupeds, birds and 'crawlers' (amphibians and reptiles). They are less interested in the intermediate taxa, such as families and orders, but they differentiate hundreds of species, in ways that largely correspond to our current scientific taxonomy (BERLIN et al. 1966). Something similar is true of the uneducated people of our own population.

Since none of these people have rationalized the principles used for this achievement, that the capacity must be given in advance: inborn Gestalt perception of and faculties for comparing and weighing characters as well as estimating the reliability of a prognosis. I have described these faculties as the hypotheses of 'comparability' and 'apparent truth' (RIEDL 1987, 1992).

(2) *A concept of participation.* Epistemologically, the process of cognition starts at the roots of pre-SOCRATIC philosophy. Gods and muses were engaged in comprehending how we might come to grasp the phenomena of this world. But as soon as philosophy discovered that humans possess reason, the question 'How come that man is able to perceive the world in a reasonable way?' changed. In principle, it was thought that there must be a connection between the senses and the intellect of human beings and the universe. Starting from pre-SOCRATIC ideas, PLATO developed the concept of participation, assuming that a superimposed principle may govern both the human soul and the universe. ARISTOTLE (ROSS 1928) expected that the 'particles' in our senses and the world may be equivalent. He also distinguished between form and substance, a distinction that may be compared to Gestalt and matter. Both occupied philosophy until today.

As I take the standpoint of EE (cf. CAMPBELL 1974), we may understand the correspondence of our mind to the phenomena of the outer world, not only on the basis of the similarity of the particles (as Aristotle thought), but also on the basis of the successful perception and computation of natural phenomena as the consequence of adaptation.

(3) *The founding of biology* can be traced back to ARISTOTLE as well. Taxonomy, comparative anatomy, and embryology all have their roots here. Thus ARISTOTLE (ROSS 1928) divided organisms into 'animals with blood' and 'bloodless animals'—a distinction that corresponded approximately to our 'vertebrates' and 'invertebrates', since he recognized only red body fluid as blood. But dolphins, exhibiting a fish-like form and a cold skin, were still—and for a long time—recognized as fish. Yet, from this kind of perspective, the first large taxa were recognized. Obviously, ARISTOTLE also held other, more important characters in the back of his mind. In his comparative anatomy he operated with similarities, compared fingernail and horseshoe (which are today taken as homologies), but also hand and claw of a crayfish (which are today taken as analogies). Such lack of distinction continued until GOETHE's time.

The differentiation of homology and analogy will become important for the cognition of evolution, as it represents contrasting interpretations of similarity. Homology traces back to 'inner principles' or systemic conditions, whereas analogy traces back to chance or to corresponding functions forced on unrelated taxa because they accidentally ran into the same environmental conditions (functional analogy, RIEDL 1977a). This was sorted out in recent times only. The concept of 'analogy', however, was soon recognized as an important source of thinking in heuristic, inventive processes. This may explain why it was used beyond GOETHE's time when actually 'homology' was meant. Otherwise, the natural order could not have been uncovered before this time, as it actually was.

(4) From here on, *the perception of four remote principles* has guided the history of biology along its long way. This has to do with the knowledge of what we must differentiate as class hierarchies and structural hierarchies in the living world, and their remote ends, such as the species and the animal kingdom. The species concept (A) had to be developed, separated from lumps of species on one side, and races and variations on the other. The species problem remained controversial until the end of the 18th century (e.g., BONET 1762). LINNAEUS (1735) developed the useful binary nomenclature, and his numerical attempt demonstrates how difficult it still was to flesh out the principles of comparing structures. ADANSON (1752) may have been the first to make a technical suggestion. He recommended to list all the characters of a group of organisms; by stepwise extraction of the redundant ones, one would proceed from families to genera and species (see FOUCAULT 1966). This was logically correct but still left open the question, 'What does it take to recognize a character?'.

Towards the other end of the class hierarchy (B), the discovery of the five kingdoms of organisms

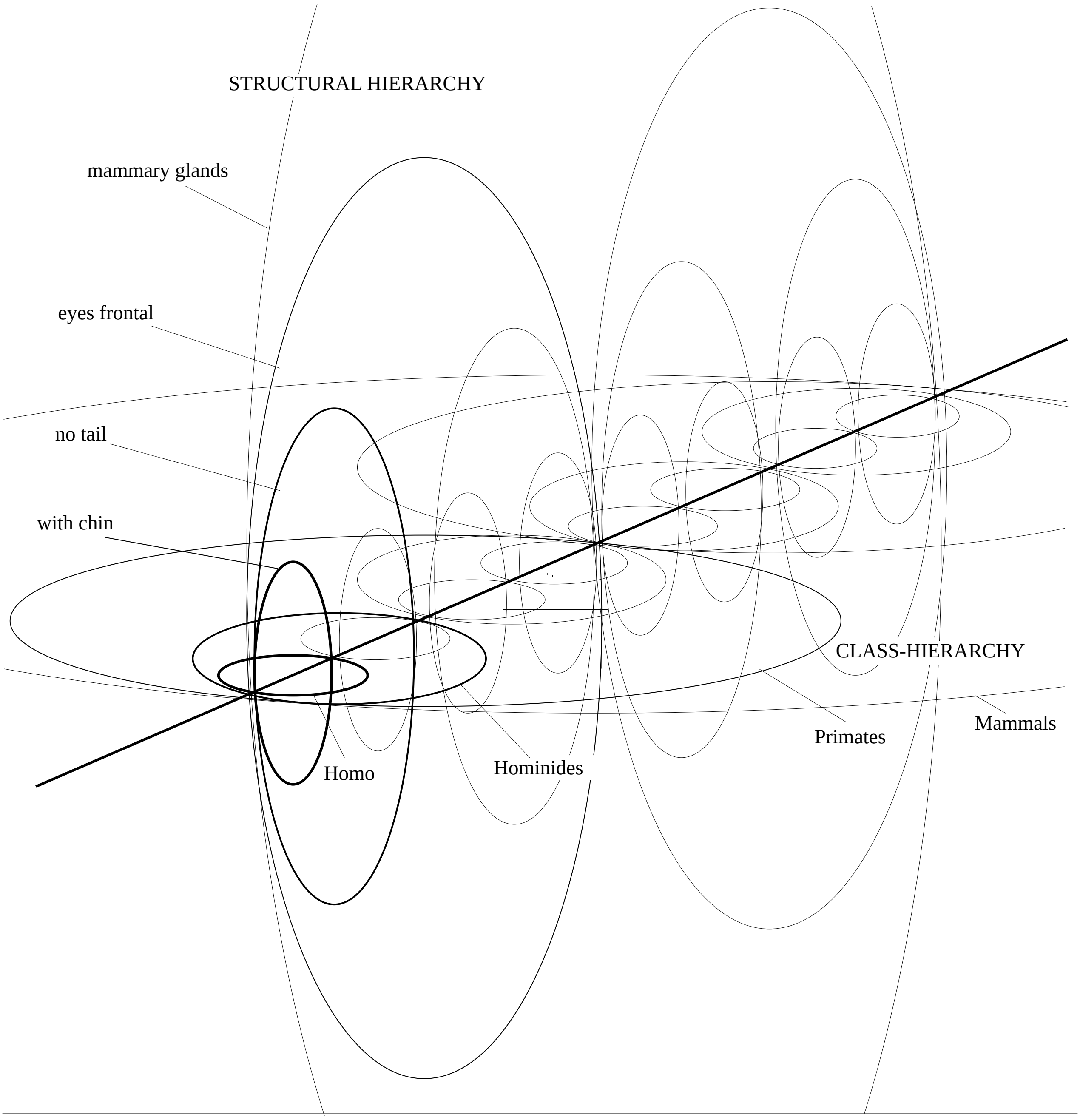

Figure 4: *The intersection of structural- and class-hierarchies*. The example is taken from the definition of the genus Homo. The cycles symbolise the circumference of taxa, as well as of the representation of characters. For simplification, only one alternative is given in each layer, and only the path leading to Homo has been labeled.

(*Prokaryotes*, *Protista*, higher plants, higher fungi, and *Metazoa*) as the largest units progressed. But in being not as essential, this division remained controversial far into the 20th century.

On the other hand, the remote principles of structural hierarchies—not yet called so at the time—had to be dissected down (C), first to cells (in the 19th century), then to protoplasm and the substances in which heredity and variability are grounded (in the middle of the 20th century). The synthesis of the uppermost principle (D), the Bauplan, is more puzzling and forced a much earlier controversy. I will return to this soon.

Although the connection of the two hierarchies of class and structure and their interconnection is of interest, it attracted little epistemological attention. Class hierarchy is represented by half a million taxa of the 'natural system', from more than two million species up to the five kingdoms. Even this empirically highly tested system, which for some goes without saying, seems to be for others a *contradictio in adjecto*. As a product of nature, it was held, it could not be a system, and as a man-made system, it could not be of nature (this was stated explicitly, e.g., by HASSENSTEIN 1951, but silently assumed by some learned physiologists). Hence, why is it puzzling, and why is it a natural system?

It might be puzzling, because class hierarchy does not provide the onset—it provides the result. And it is natural, because it is empirically woven with increasing knowledge and confirmation from the orderly similarity patterns of the hierarchy of organismic structures. The definitions of class hierarchies are made by terms of structural hierarchies. So, the possession of a notochord (*Chorda*) includes creatures with a backbone, then those with regional differentiation of their backbone, and so on (chordates, vertebrates, quadrupeds). These structures are the necessary matrix for one another in development, and in the same sequence.

One should also not forget that a concept like 'feather' makes sense only if two conditions are met. First, that one knows the alternatives, e.g., hair, scale and armor; and secondly, that it is part of the skin, and the skin is of a bird. Structural hierarchies compose the system of comparative anatomy, the derived class hierarchies compose the natural system of overall similarities.

(5) *The Bauplan discussion* towards the end of the 18th century attracted the fourth problem (D). The question, put in my terminology, would have read: do higher principles of structural hierarchies exist, and how could this be proved? Most obviously, beetles and butterflies have more in common than insects and spiders or arthropods and segmented worms (*Articulata*). The discussion, provoked by vast collections of specimens by BUFFON (1749–1789), LAMARCK (1809), and studies carried out by CUVIER (1817/1828) and GEOFFREY SAINT HILAIRE (1818–1822), led into a deadlock. The learned debate centered around an unending quest: whether the investigation had to proceed analytically (from the whole to its parts), or synthetically (the other way around). But the existence of similarity patterns at different levels of generalization (as we would say today) was accepted eventually.

Epistemologically, we will find this problem of investigation prepared for a remarkable step: the insight into alternating processes, as we will find it soon in GOETHE. However, the principles of typogenetic and typostatic phases (new structures emerging within a short span of time versus structures that are being fixed stepwise) in phylogeny, necessary to understand the transient phases from one Bauplan to another, became topics for paleontologists only recently (e.g., SCHINDEWOLF 1936, SCHMALHAUSEN 1949). But they were already inspired by questions of explanation. And the discussion of the underlying principles, such as constraints (RIEDL 1977a), equilibria (GOULD 1984), or the chance of success of variations (WAGNER 1988) still goes on today.

(6) The *similarity patterns* were completed in the meantime. Throughout the 17th and 18th centuries, insights into the patterns of similarities in the organic world were completed to a remarkable degree. The acceptance of the Bauplan phenomenon justified the hierarchical arrangement. The reign of animals was in principle differentiated correctly. More than a thousand species were arranged in families and orders, up to fifteen stems (or phyla) of animals, in a way that fully corresponds to our modern taxonomy.

This is all the more remarkable as almost no interest was shown concerning the mental procedure behind this achievement; it was still done unconsciously. Only three stems remained in an uncertain position; infusorians, hydropolyps, and echinoderms. The situation was crying out for an explanation. We will come to this later.

Underlying this kind of thinking must have been the use of a class hierarchy, represented in the taxonomy versus a structural hierarchy, such as in cells, tissues, and organs, or multicellularity (*Metazoa*), segmentation (*Articulata*), armored skin and legs (*Arthropods*), and further on by a three-partite regionalization (*Insecta*), and so on. But no word is found in the literature about this differentiation and the interconnection of the two hierarchies. It was important for all further steps, but obviously taken for granted.

(7) *The type problem produced the methodology* at the same time. Around the turn of the 18th to the 19th century, interest in the process of discovery arose. GOETHE, encouraged by and enthusiastic about the discussion in the French Academy, particularly between GEOFFROY SAINT HILAIRE and CUVIER, started to investigate the methodology of comparative anatomy. "As in other sciences," he summarized, "one did not sufficiently use refined kinds of imagination. One either took the case too trivially, viz. getting stuck with simple appearance, or one tried to aid one's thinking with teleological causes. One thus drifted further and further away from the concept ('Idee') of a living creature." (written 1795, amendments 1796, position 8; 26. Translation cited from RIEDL 1995b).

This was also the time when LAMARCK published his *Philosophie Zoologique* (1809), in which he was the first, as we will elaborate later, to ask explicitly for the explanation of such a hierarchical order. But GOETHE did not take any notice of LAMARCK.

(8) *The core of the method is based on mutual optimization*, which is in principle a 'hermeneutic' process based on observations, as will be shown now (RIEDL 1985): Compose a theory by induction from the cases, namely the specimens of species, holding the concept of what they have in common, being the conjecture for a type concept. Proceed from there by deduction of the assumed basic features of the species (and so on) until an optimum of understanding seems to be achieved. "The particular case cannot contain the pattern of the whole", GOETHE states (position 10;16), "and consequently, there is no chance of finding the pattern of the whole in a particular case. The classes, genera, species, and individuals are as cases to the law, they are contained in it but they do not contain or give it."

In principle the methodological problem was solved. The process of mutual optimization between the imagination and the cases of a Typus concept by iteration was understood. The insight that this process must extend into the hierarchy of taxa was suggested in outline (RIEDL 1994).

However, at the turn of this century, the feeling for the qualification of such processes was still taken for granted in arts and letters. The reductionism required by positivism had not yet spread into all sciences, and, on the other hand, empirical hermeneutics was not yet led away to idealistic concepts by philosophy. This may have been the reason why GOETHE trusted in it also for science.

In a way, hermeneutics as an empirical method was already in the air. E.g., August BOECKH's teaching on the analysis of Greek and Latin texts not only included mutual optimization by iteration, but also, clearly, hierarchy, i.e., the mutual enlightenment of understanding which operates between word and sentences, between sentences and context, context and author, and author and his time. This understanding must have been simply in the mind of learned persons at the time, because BOECKH never considered publishing his courses on this subject, which he essentially started giving in 1811 at the university of Berlin. They were not published until two generations later, by BRATUSCHECK (BOECKH 1877, [2]1966).

After this episode, the interest in the process of discovery fell into a gap between the positivisms in France and in England, which focused on a straightforward mechanistic approach to explaining nature on the one hand, and on an idealistic (later romantic) conception in Germany on the other, leaving it to an unchangeable spirit or mind to judge an ever uncertain and changing world. Consequently, the concept of type was later on felt to be an art but not a science (HASSENSTEIN 1958).

(9) *The homology concept was developed further* independently of these changes. In the middle of the 19th century a coincidence of two discoveries characterized the situation. RICHARD OWEN coined the term 'homology' (1848), and the knowledge he acquired allowed to separate the term from 'analogy', under which the phenomenon was formerly subsumed. OWEN applied homologies as the components of the type (Archaetypus) in the sense of GOETHE. On the other hand, WALLACE and DARWIN (1859) used MALTHUS'S and SPENCER's principle of selection to understand the origin of species—the process of speciation.

The debate around the homology theorem as the method to uncover identical principles of form became more vigorous towards the middle of the 19th century. OWEN distinguished special, general and serial homologies. The last are called 'homonomies' today. The first were defined as homologies between recent species, the second as homologies tracing back to the 'Urform' or archetype. We do not use this differentiation anymore, but it contributed in two ways to subsequent thinking.

In one respect this idea supported DARWIN (1859, chapter 14) in separating inherited homologies from adaptively required analogies, which he called analogous modifications "with respect to comparable environmental conditions". He felt that homologies may be explained "to a certain degree" by his theory of selection.

(10) *Mixing the principles*. Not unlike OWEN, DARWIN saw and used the importance of the adaptive background of analogies versus the counteradaptive one of homologies. In another respect, this distinction eventually led to confusing knowledge and understanding. OWEN contributed the confirmation about the existence of similarity principles, DARWIN the explanation for the process of speciation. Both insights are indispensable for grasping evolution; cognition of the patterns of similarity as well as their causal explanation. But, paradoxically, a quarrel about priority arose between the two, in which the principles used for the gain of knowledge and for the gain of understanding—in our divided terminology—were mixed up (the story is in VOIGT 1973), probably because both were right in insisting on having discovered the principle of evolution.

Because of his controversies with TH. H. HUXLEY, OWEN was regarded as an old-fashioned 'typologist' and even PLATONIST, and mistakenly as anti-DARWINIAN. OWEN, however, took his 'general homology' phylogenetically and methodically as the source to discover the archetype, in the sense of 'common ancestry' (cf. ASMA 1996), while DARWIN and HUXLEY had explanation in mind.

From this time on, the visualization, i.e., the mental picture of the path of causal explanation of organismic patterns of similarities, started to cover, to subsume and eventually to substitute the path of cognition. The term 'morphology', originally reserved for the specific process of the mutual optimization of the reliability of cognition by comparison, became pejorative or meaningless. It has since been used either pleonastically for 'comparative morphology' or contradictory for 'functional morphology', where epistemologically nothing but 'comparative' or 'functional anatomy' can be meant.

(11) *Pinning down the mistake*. Most scientists of the time also confused the matter or followed the thinking of DARWIN and Thomas Henry HUXLEY. Shortly after DARWIN's *Origin*, scholars with great influence such as Ernst HAECKEL (1866), a devoted admirer of DARWIN, or Carl GEGENBAUR (1870), hardened the mistake, substituting concepts of the confirmation of a cognition with concepts of the explanation of an understanding. HAECKEL took common origin as the ground for homology, GEGENBAUR embryonic *Anlagen*. If they were right, one would have to ask by which criteria HAECKEL did know something about common origin, and GEGENBAUR about common embryonic origin. Yet biologists followed such authorities. As we will come to see, this mistake persists to date.

(12) *An empty century*. Remarkably enough, and this is the sore point of this history, the challenge posed by the problem of cognition was not met in the next century, during which the system of cognition continued to be substituted by explanation. The situation would not draw our attention if the interest in evolution had faded away.

Loss of interest in a phenomenon and its rediscovery a century later is not, by itself, remarkable in the history of science. But here it is the other way round! The interest in evolution increased remarkably, and in the century between 1850 and 1950 an explanatory theory of evolution was articulated that is taken to be comprehensive and sufficient, satisfying the needs of handbook and textbook writers.

(13) *Criteria of homology*. It would take a century after OWEN before interest in the path of discovery reemerged. It was Adolf REMANE who, in 1952 and then again in 1971, came up with criteria of homology again. But in contrast to his time, his criteria were again criteria for the process of cognition. As far as I am acquainted with REMANE's personal thinking, he was interested in operational methods for systematics and comparative anatomy. His criteria were operations, viz. either to compare many features starting from two specimens, or to compare single features of the species of a whole taxon (my interpretation).

For the first case he defined his 'main criteria' for the discovery and justification of homologies, for the second case he defined 'supporting criteria' (*Hilfskriterien*). I have found the two to be equivalent, such as flipping a coin many times compared to flipping many coins once. Actually all his criteria may be synthesized.

REMANE divided his 'main criteria' into three kinds: 'special quality', 'position', and 'transition'. Considering the hierarchical patterns in which homologies are positioned in relation to each other, I have found that 'position' gives an indication from the side of the upper homology, 'special quality' from the side of the lower homologies. If one focuses, e.g., on the vertebral column of the collar region of a mammal, 'vertebral column' is the expected position, that is to say, the homology in mind must be placed between skull and thorax. The 'special quality' leads us to expect seven vertebrae. Focusing the second vertebra (spine), the '*epistropheus'*, the collar region is the upper homology, the possession of a tooth (cog), a *dens epistrophei* is among its special qualities, and so on.

What was not noticed in the mid-20th century is that what I ascribed above is a hermeneutic process as known and thought about by BOECKH a century earlier. As in the deciphering of an unknown text, the reliability that a homology be detected definitively or with high probability depends on the number of continuously confirmed predictions within the whole hierarchical pattern of expected homologies.

(14) *The muddle however remained* for many years after, and is still common sense for most of my fellow biologists today. Actually the most learned and admirable scientists, such as George Gaylord SIMPSON in paleontology, Ernst MAYR in systematics, and Dieter STARCK in comparative anatomy, fell into the same trap. When MAYR (1969, p. 68) endorses SIMPSON's (1961) view in stating: "representatives of a taxon are similar because they are related, and not... (related) because they are similar," he explicitly substitutes cognition with explanation. What would

then justify the assumption of their relationship? In the same vein, STARCK writes: "to establish homologies... requires the knowledge of phylogeny" (1978, p. 1–2). From what sources, we must ask, would his knowledge about phylogeny then derive?

(15) *The demonization.* Adolf REMANE's book (1952, new edition 1971), and those of a few others, such as BLJACHER (1960), were never translated into English, because American scientists with command of German or Russian demonized those approaches as 'German idealistic philosophy'. As in most such cases, no written documents can be cited. This supposition stems from discussions the author has had. But the result may confirm it. And most of my English-speaking fellow biologists assume that 'what is important is translated anyway'.

The effect was either disregard or attempts to rebut a presumed anachronism. A century of remarkable progress in explaining evolutionary phenomena is still overruling and screening off the process of cognition.

(16) *A theorem of reliability.* Coming back to REMANE's criteria, we recall that the criteria of 'position' and 'special quality' were to be synthesized. If, furthermore, one takes the 'transition criterion', i.e., searching for information by intermediate stages, since they are only composed of a sequence of position and quality criteria, then the homology theorem can be justified on the grounds of a theorem of reliability (RIEDL 1977 and 1980 a + b).

What was said above specifically traces back to the starting point of my investigation. The cognition of evolution, common descent, is based on the recognition of the hierarchical pattern of similarities, the "*Order in Living Organisms*" (RIEDL ,1977), composed of a class hierarchy of 'Baupläne', which in themselves are composed of a structural hierarchy of homologies. The reliabilities of the cognition of each of these homologies, derived by a hermeneutic method of mutual enlightenment, support one another, and they all have the laws of simultaneity as their basis.

We may also recall that the reliabilities of the laws of simultaneity, the *simul hoc*, depend on the number of confirmed expectations in observations, as the probability of the laws of causality, the *propter hoc*, depends on the number of confirmed expectations in, e.g., experiments.

My report (RIEDL 1975, 1977a) arrived at a time when the interest of scientists, particularly biologists, for the last hundred years, had shifted to explanation. With the exception of a few insights such as 'cerebral hermeneutics' (STENT 1981), the interest in the hermeneutical process had dissolved.

(17) *Attempts to escape the proplem of complexity* of biological systems were made several times but remained unsuccessful. The only attempt that produced a fair amount of literature and discussion is numerical taxonomy. Since there is no computer program to test homology, numerical taxonomists suggest, the systematist should not speculate but simply take exact measurements (SOKAL and SNEATH 1963). This may be acceptable if the joints of a leg of the species within a genus of, e.g., beetles are considered. It turns to absurdity if the leg of a beetle is, by the same method, compared with one of a spider. In other words, there is no way around homology. And the discussion (e.g., MAYR 1965, SNEATH and SOKAL 1973), extending into the weighting problems related to homologies ended in a deadlock and has remained so in principle.

The problem of weighting is easier to solve with regard to compounds of chemical similarity, called 'isologes' in chemistry. Although most homologies are composed of isologous compounds, we accept (since FLORKIN 1966) that isologes can only be taken as homologous if they match beyond doubt with the homologies of the species that are being compared. This again depends on a process of mutual enlightenment.

(18) *The destiny of hermeneutics.* Hermeneutics, banned or forgotten in the sciences, took a different path in the humanities (in arts and letters). Learned persons like GOETHE or BOECKH were not deeply engaged in philosophy but more interested in scientific methodology. Thus not much cross-cultural transfer occurred, while SCHLEIERMACHER and WILHELM VON HUMBOLDT, their contemporaries, independently elaborated a 'universal hermeneutic'. These authors influenced DILTHEY, who in 1883 coined the term 'Geisteswissenschaften' (humanities) to detach them from 'Naturwissenschaften' (natural sciences) by a methodological assumption. Since the objects of the humanities are man-made, and those of natural sciences are not, he connected the first with a method of a hermeneutic understanding, and the latter with a method of explaining. The humanities, which at this time were already competing vigorously with the natural sciences, adopted this differentiation with relief, thus causing intractable problems for three reasons. First, understanding and explaining remained interchangeable terms in ordinary language as well as in most scientific language games (cf. TOULMIN 1961). Second, it obscured the relation between the processes of confirmation and explanation. Last and worst, it made most philosophers think that explanation is all there is to science,

while the gain of knowledge as deeply rooted within the abilities of human beings is beyond any need and/or possibility of methodological enlightenment (an important exception being APOSTEL 1959). Consequently, this stream of thoughts led into two dead-end roads. 'Philosophers of science' (e.g., STEGMÜLLER 1974) banned hermeneutics from the sciences as a logical circle; 'philosophers of the humanities' (e.g., HEIDEGGER, GADAMER or HABERMAS), still within the currents of 'German idealism', 'Phenomenology', or 'French existentialism', changed the old empirical approach into 'philosophical hermeneutics' (cf. GRÜNBAUM 1984, 1986; GUTTING 1984; PUTNAM 1990). As a result, their views remained more entangled between their own controversies without accomplishing anything in clear methodology (cf. KALERI 1982).

One of the consequences is that even some humanists shun the term hermeneutics today (it also disappeared in the *Encyclopaedia Britannica*, where it is found only under "Exegesis", viz., *hermeneutica sacra'*.

At this point of my historical survey, two questions must be answered: What was accomplished by the explanation of evolution? This will be laid out in the next chapter. What harm results when the gain of knowledge is substituted by the gain of explanation? This is the topic of the third part of this contribution.

III. The history of explanation

(1) *The will of the Creator* was taken as sufficient to understand why the world is as it is. This is evident from pre-SOCRATIC philosophy, and even from HOMER's poetry. Most of the incidents influencing a human being's life were understood as the aims of a purposeful world of gods. And we must confess that parts of these assumptions are still in evidence in all of today's religions.

But men's big leap from myths to reasoning also took place early. It was in the fifth century BC. that ANAXAGORAS stated that "the traditions of the Greeks are too multifarious and too childish" to be trusted. As a consequence, the tradition which consists in searching for the truth is almost as old in our culture as the other one. And the question: 'Why is a fact what it is?', soon begot its pragmatic and psychological influence. Insights, as we expect, may improve predictions, and correct predictions, as we know, support life and its quality. In addition, knowing and understanding also provide intellectual satisfaction.

EE has disclosed the inherited instructions. We are equipped with the expectation that like or similar structures or processes have identical causes and correspond to identical purposes or functions (RIEDL, 1984, 1985). The process is still based on generalization and similarity patterns, but employs, in addition, assumptions beyond the visible ones.

(2) *A division of four causes* was perceived but a century later. ARISTOTLE (ROSS 1928) distinguished the four forms, in which causality appeals to our intellectual faculties; *causa efficiens, causa materialis, causa formalis*, and *causafinalis*, corresponding, in today's terminology, to 'power', 'material components', 'form-giving selection', and 'goal-seeking processes' respectively.

From the standpoint of EE, the division is caused by two cognitive dualisms. First, power and goal-seeking are perceived as processes, in contrast to material and form, which are viewed as structures; they are represented by verbs and by nouns, respectively. Second, by focusing the hierarchical order of complex systems, power and materials enter from the lowest, 'form-giving' and goal-seeking processes from the highest layers (RIEDL, 1985, 1994).

In the following centuries, however, and in ignorance of this recent insight, attempts did not end to assume one of the four causes as the primary one, causing, or even substituting the others. Accepting such a diversity as apparent seemed to be unsatisfactory. And the idea that this universe must have had a 'single first reason' seems still to be at work. It was not understood that profound logic governs this division.

One of the following traditions, labeled 'idealism', took *causa finalis*, the other, known as 'materialism', envisioned *causa efficiens* as the primary source of causing the phenomena of this world. Either of them, paradoxically, thought to be sufficient to understand the phenomena of this world. In principle this has not changed since. From the Early Church Fathers to the Humanities one tended to the first; in the sciences, under the leading paradigm of physics, the second assumption was favored. This was either ignored by biologists, or was later to cause new contradictions and controversies for them.

This is all the more strange as the function of materials as well as the goal-seeking processes must have been evident for the biologists of the following centuries. Also, the principle of selection, corresponding to the *causa formalis*, was envisioned since the times of LUCRETIUS.

(3) *The 'GALILEAN revolution'* pinned down the materialistic concept; in contrast to the idealistic one

that was perpetuated by PLATONISM and the church. GALILEO was striving against the contemporary muddles in philosophy, restricting his view to the *causa efficiens* (in English unhappily translated as 'efficient cause', as if the rest would be 'inefficient causes'). He knew about ARISTOTLE's distinctions. But, on the one hand, material and form of his objects, as well as the goal of his experiments where certainly variables of his own decisions. And, on the other hand, caution was recommended with respect to the doctrines of the church; particularly *causa finalis* had to be avoided.

In this way physics was born as a science which reduces qualities until it makes sense to describe the remaining quantities in mathematical terms. From here, great influence spread over all of the inorganic sciences. The remarkable success that followed promoted the assumption of the materialistic trend, that power is not only a necessary but also the sufficient cause to explain the world. A major hurdle was thus prepared for biology.

This development would hardly have influenced contemporary biology were not it for the circumstance that in the Renaissance (e.g., in LEONARDO) causal questions became prominent, which led to functional and later on to physiological questions (e.g., with SPALLANZANI). This would not be worth mentioning, but physics later on gained a kind of leadership in science which caused the aforementioned problems for biology. The roots of the 'physics envy' problem are found here.

(4) *A puzzle of taxonomic order* emerged on the way from the 17th to the 18th century. Sciences, at that time Botany and Zoology, were occupied by the discovery of the order of similarities in the diversity of organisms, and made great progress in the second half of the 18th century.

This situation marks the new turning point for introducing explanation in biology. For progressive scholars of that time accepted that even sensibility, irritability and emotions must derive from a material basis; a 'physiological' one, as we say today.

(5) *A need to ask for a cause* of the natural order come up in the second half of the 18th century. As reported, zoology and botany collected large materials during the 17th and particularly the 18th centuries. The deep insight into the "arrangement and division of animals" was felt to be accomplished, and it confirmed "a most natural order". This was mainly due to BUFFON's and LAMARCK's extensive investigations in comparative anatomy. And from today's perspective we have to acknowledge that the genealogical tree of animals was in principle visible.

Hence the important question arose of how to explain this 'most natural order' that was clearly beyond every chance effect. And LAMARCK, in the preface of his *Philosophie Zoologique* (1809), marks explicitly the turning point leading from widely confirmed knowledge of this natural order to explanation and understanding, when he writes: "How could I ever have observed this peculiar (remarkable) graduation of the organization of animals …, without asking for the cause (causal reason) of such a positive (evident) and important fact" (p.1).

With this step, with the improved reliability of the discovered 'most natural order', the genealogical tree, the 'theory of descent' was born. Similarity was explained by the degree of relationship, deviation by the necessity of adaptation to the environment—and this included man. In analogy to human introspection, LAMARCK extends this to 'need' and 'want'.

(6) In addition to his discovery of common ancestry, *active adaptability* was envisioned by LAMARCK as well. These two concepts—his discovery of descent and his attempt to explain inheritable variation—have to be strictly separated. Starting from the fact that use and disuse can change organs (e.g., body building increases muscles, while splinting reduces them) he expected such changes to be in a way inheritable. This would have been a most welcome way of evolution, if our efforts to keep body and soul healthy and trained could be an immediate benefit for our children. But, as we know, evolution does not work this way.

GOETHE did not participate in this discussion; he did not even take notice of LAMARCK. His interest remained in the process of cognition. His few remarks about explanation will be cited in a later connection. Yet many scholars followed LAMARCK. Erasmus DARWIN (1731–1802), for one, did, and (focusing on the following mainstream) so did his grandson CHARLES (1809–1882), as well as one of CHARLES' teachers, Robert Edmund GRANT. Charles DARWIN followed active adaptability for two reasons.

First: it was current thought, there was no other concept for the cause of a 'reasonable way' of variation at hand; the change of organs according to their use was obvious (and wrongly assumed to be inheritable), and this concept of active adaptation, which he developed later, clearly depended on a kind of 'adaptive' inheritable variation. I speak explicitly of reasonable and adaptive variation, because random variation as we understand it today was on no one's mind at that time. Variation was viewed as benefiting the process of adaptation. Geese kept in captivity, DARWIN reports reduce the strength of their wings, and, as the altered way of living leads to expect, their legs grow stronger. In a way DARWIN was more LAMARCKIAN than was LAMARCK himself, for he believed, e.g., that in human populations where the penis is circumcised, the prepuce became shorter over the generations. This would not even have been believed by LAMARCK.

Secondly: a mechanism was posited which may be called an 'inner principle' of evolution. It is based on the observation of a whole set of facts in structural biology, which cannot be understood at all by adaptation. I come to this in point 8 of this chapter. The reason why DARWIN put the two phenomena together is that he considered inheritable variation itself an inner principle.

DARWIN remained engaged by this problem until his late years; he published his "pangenesis theory" in 1862.

In the very same way as LAMARCK, he first justified his move from discovery to explanation: "Everybody will try to explain, even in a most incomplete way, … how it comes that the effect of increased or reduced use of a limb can be transmitted to the child." And later: "As far as I know, no other attempt has been made to unify these classes of facts under a common viewpoint, as incomplete, I commit, as this endeavor is" (1875). These 'classes of facts' included not only heritability of variations (which, as we see now, were no facts), but hard facts suggesting an 'inner principle' of evolution.

In short, DARWIN suggested that microscopic particles produced by each organ would communicate throughout the body, e.g., by means of the body liquor. And if an organ changes size, the number of particles would change and be noticed by the germ cells. As we know today, this is not the case. Yet the necessity of an inner principle was envisaged to complete evolutionary theory.

(7) *The concept of passive adaptation* has another, better known history, on which I can be short. Foreseen by LUCRETIUS (+ 55 BC) in Rome, anticipated by MAUPERTUIS (1698–1759) in Berlin, the concept underwent a remarkable development in England. MALTHUS (1766–1834) made selection an object of scientific thinking; SPENCER (1820–1903) coined the term "survival of the fittest"; WALLACE (1823–1913) and DARWIN applied this concept to the living world in general.

After DARWIN's *Origin of Species* (1859), the principle of passive adaptation spread quickly in the scholarly world, while active adaptation remained widely the assumed source for variability.

(8) *Inner principles of evolution* have been suggested early and were forgotten early. The earliest notion stems from HIPPOCRATES (460–375). "Every part of the body", he states, "seems to club and contribute to the seed", and was already known to DARWIN (1875,Vol.2,p.37o).

GOETHE, as we recall, was mainly interested in the system of cognition. But a few hints are preserved. "In the process of erecting the type…", he states (1795), "we presuppose that nature follows a certain consequence. We trust that nature follows certain rules in every case". Later he called such principles 'esoterisch', which was subsequently understood—wrongly—as 'mysterious' (at that time, in German one used 'esoteric' in contrast with 'exoteric' in the sense of ARISTOTLE 'for insider' and 'for outsider', as well as for inner and outer factors). 'systemic conditions' would be the best translation today. This is the direction in which evolutionary theory will eventually go.

With his early discoveries in embryology, Carl Ernst von BAER (1828) took the position of guiding principles in ontogeny, because there are only a few types of development in the animal kingdom, corresponding widely to the largest taxa—an insight to which DARWIN strangely did not refer.

A whole set of phenomena that could not be explained by adaptation was known to a learned person like DARWIN, however, pointing towards an 'inner principle of evolution'. Spontaneous atavisms, surplus organs, cysts in the skin and the ovarium, containing hair and teeth, and 'heteromorphoses' were among them, showing complex organs (e.g., antennae) produced at the wrong places.

(9) *To complete the theory*. After the problem 'What selects variation?' was solved, DARWIN's genius tried to solve the three open questions with the single assumption above. Again it is of no importance in this context whether he discovered the right mechanism or not. What is important is that he identified the three remaining questions: What causes variation? Which inner principle makes the germ cells sensible to select the requirements of the organism's organization (as demonstrated by spontaneous atavisms or heteromorphoses)? and, What makes changes inheritable?

Today the first and third problems are solved. Variability is ultimately caused by mutability; and mutations in the genome are inherited. The second question was more and more forgotten—a misfortune for the further development of evolutionary theory.

(10) *A first reduction* on the broad view in the LAMARCK-DARWINIAN sense was produced by WALLACE. His successful book, *Darwinism* (1889), published seven years after DARWIN's death, was conceived as a homage to DARWIN and intended to update what he thought was pertinent in DARWIN's thinking. Variability was still thought to be caused by changing habits, in a sort of active way, and inheritance was taken for granted, its mechanism being left open. In all, he took selection by the environment as a straightforward and sufficient explanation for the phenomena of evolution. Pangenesis, with its expectation of an inner principle, was left out. He felt it to be fussy and too troublesome.

No doubt, the contribution of passive adaptation was, in outline, explained. This concept foresees a one-sided process producing 'correspondence' of structures and functions to the requirements of the environment. But by this step a first factor of mere chance becomes dominant in the explanation of evolution: because the encounter with new, selective acting environmental factors must be widely due to chance effects.

The phenomenon of *coherence*, which makes the inner parts coherent to one another, now called 'organization', was left open—particularly so because coherent processes are of a feedback kind, reciprocal, ending up in top-down/ bottom-up-causality, which seemed troublesome or irksome to comprehend. In short, evolutionary theory become adaptationist in a first stage.

(11) *DARWINISM and LAMARCKISM were invented* by this step. WALLACE's DARWINISM (actually a 'WALLACISM') seemed to have lesser open questions, was easier to grasp, and therefore welcomed, and the term 'DARWINISM' spread abroad, but, paradoxically, was not honoring DARWIN's full view, but the reduction of DARWIN by WALLACE. Only schools of thought that were very close to DARWIN himself, such as HAECKEL's and his successor, PLATE's, where aware of this change and labeled themselves the 'Old DARWINISTS'. But this distinction was also soon forgotten.

Whoever still believed in a predominance of active adaptation was labeled 'LAMARCKIAN'. And in opposition to the amputated DARWINISM a different tradition developed that later on led to vitalism and kinds of metaphysical thinking such as psycho-LAMARCKISM.

(12) *A second reduction laid ahead*, eventually caused by a definitive separation into two parties, so to speak: active and passive adaptationists, i.e. LAMARCKISTS versus DARWINISTS, taking definitely opposing standpoints. This developed together with the theory of the German August WEISMANN, a contemporary of WALLACE. This theory, later known as the 'WEISMANN doctrine', assumes that no information transfer can be expected from the body ('soma', later called the 'phenes') to the germ cells (the genes). By taking this seriously, the questions 'What causes variation?', and 'What makes variations inheritable?', gained in importance.

(13) *The change from DARWINISM to Neo-DARWINISM* occurred when both problems seemed to be solved simultaneously. Mutability was discovered, and mutations in the germ cells where found to be inheritable. This changed DARWINISM (WALLACISM) to neo-Darwinism, and, correspondingly, the opposing LAMARCKISM to psycho- and neo-LAMARCKISM.

Regardless of all further accomplishments in classical and molecular genetics, this paradigm has in itself met wide acceptance and remained in principle unchanged. This is the more interesting as amendments in the concept of genetics occurred at the same time. In early genetics, a kind of bead-string model was envisioned, assuming that one gene acts for one character in the phene system. Yet it did not take long for the phenomena of polygeny (several genes for one phene), and pleiotropy (one gene for several phenes) to be discovered. But this did not change the paradigm.

This state of the art of the neo-DARWINIAN paradigm accepts a second stochastic term, because selectivity is still thought to originate from accidental encounters with new requirements of the environment, as the actual cause of variation was found to act totally accidentally, in the sense that it shows no correlation whatsoever to the adaptive requirements of the species. What is now called 'overall fitness' must be the result of the chains of two stochastic processes—mutability being absolutely blind, selectivity being totally short-sighted, not foreseeing the slightest possible advantage for the future. It must now be assumed that the orderly patterns of the living world, being far beyond chance effects, may be the outcome of chance effects after all. The result of the paradigm is now a pure or casual adaptationism.

(14) *The 'synthetic theory of evolution'* sticks to the same epistemological standpoint. Whatever the new insights in population dynamics, in processes of speciation (the formation of new species) and in animal geography, which the theory synthesizes within the neo-DARWINIAN account, confirms its paradigm (MAYR, 1942, SIMPSON, 1952, DOBZHANSKY 1951), under the restriction that only phenomena of 'microevolution', i.e., processes within the level of a

species, are considered. This excludes all the remarkable phenomena piling up from genera to the five kingdoms. The synthetic theory wishes to restrict research to areas where experiments are meaningful, however. It excludes phenomena beyond the life span of a species, and explanations of these phenomena, beyond the possibility of experimentation, can consequently only be deduced.

Yet even the old WEISMANN doctrine seemed to find its confirmation in the 'central dogma of molecular genetics'; today with two formulations. The 'hard' definition states: 'No transfer whatsoever of information from the phenes to the genes is possible.' The soft one admits: 'No chemically coded transfer can be expected'. This is again of interest, because within the same decades gene regulation, including a hierarchy of gene interactions were discovered, first by MONOD (e.g., 1971), and later confirmed by many others. But even results such as these did not change the paradigm. And this is still true for today's mainstream.

(15) *Similarities in molecular systematics* became a topic when chemistry made progress in biological issues. The idea was that basic structures may have special weight in taxonomy. Basic structures such as compounds however are, compared with the composition of anatomical features, relatively simple, i.e. poor in characters. In such cases one must refer to higher systems, as we have found: the 'position effect'.

To be sure about the homology of similar molecules, one is forced to restrict the investigation to closely related species and in general to the genealogical tree; established by comparative anatomy and classical systematics.

Methodically, several ways of computation are suggested for dealing with the parsimony of supplementary assumptions. This is reasonable; in particular, 'DOLLO parsimony', assuming the unlikeness that lost structures may reappear in the same way. In controversies of methodology, iterative processes are now recommended (HILLIS and MORITZ 1990); the recursive learning process of GOETHE and BOECKH has the chance to be reinvented.

(16) *The concept of systemic conditions* in evolution fell into oblivion, but never fully vanished. We noticed this view from GOETHE, VON BAER and DARWIN; it was not forgotten by 'old DARWINISTS' such as HAECKEL and PLATE. Concepts of this kind were still held by several paleontologists, systematists and comparative anatomists, all scholars dealing with the complexity of 'macroevolutionary' phenomena, from OSBORN (1934), SCHINDEWOLF (1936, 1950), SCHMALHAUSEN (1949), and BERTALANFFY (1968), to WADDINGTON (1957) and WHYTE (1964), to cite only a few.

I take the same standpoint because, in addition to giving a review of DARWIN's "large classes of facts" as they have accumulated to date, I can add a causal model. Not LAMARCKISM, but a systemic approach to evolution is the solution. Also, the synthetic theory is accepted as correct: it holds the truth, maybe nothing but the truth, but not the whole truth. It needs to be completed by a systems theory that allows to explain coherence and constraints, building planes and the whole order in living organisms, by feedback causal explanation between genes and phenes (RIEDL 1977, 1983, WAGNER 1983; 1985).

I foresee that among all gene interactions and pleiotropic factors that may develop by chance, those will be selected which, again by chance, code for phenes that are functionally interdependent. This process, boosted by speeding up adaptation, drains information, not from the environment, but from the Bauplan, back into the organization of the genome. But this is not a topic of this chapter any more, because the mainstream in biological thinking did not take notice: it either ignored feedback causality or confused my concept with LAMARCKISM.

IV. Substituting cognition by explanation?

(1) T*he question to be raised* is: What happens to the concept of evolution if the complex of information as assembled by the system of cognition of evolution is overrun by the current attempt to explain evolution? I say 'overrun', because what actually happened is the allocation of different weights. This was not done by epistemological considerations, it just happened, probably by convenience, and under the pressure of a remarkable growth and diversification of biological disciplines. It is, as usual, not much more but a part of the unfounded remnant, as in every scientific paradigm.

Basically the accumulated knowledge of evolution cannot be fully ignored. Even a chemist ignoring taxonomy will ask a biologist who is acquainted with what is sloppily called 'diversity', how the species is named from which he extracts a compound or a DNA sequence. And it is even more characteristic that he takes it for granted that his analysis makes sense because he assumes that it can be generalized within a frame of related species. What counts are nothing but different levels of impartiality or ease.

To take measure of the units of accomplished knowledge as well as of understanding (compare again Fig. 3), the units have to be specified. For the system of cognition I take our knowledge in comparative anatomy, taxonomy, and macroevolution; for the system of explanation the complex of understanding provided by the synthetic theory. It might seem more tricky to define comparable units of the accomplishments of two different fields of thinking. But we already found 'common features' in part 1. And particularly the fact that 'both are controlled' in the same way, namely (Fig. 1 and 2) by helicals of conjectures and experiences, will support this undertaking.

(2) *An underestimation of bare dimensions* drives thinking into a first simplification. Take the DNA code; which is translated by 'triplets' (three base pairs) into 21 amino acids, analogous to a Morse code's (say: point, hyphen and space) translation in 24 letters. The genome of a small fish contains already 3×10^8 of such triplets, corresponding to the 2.7×10^8 letters (and intervals) in the 24 volumes of the *Encyclopedia Britannica*.

Let us assume that this edition strangely is in danger of elimination, because one letter was found to be wrong: in a single sentence the word 'in' should read 'on'. The mistake must be found by trial and error; usually in a single position, as due to the genetic paradigm. By analogy, only a single alteration of an arbitrary letter can be allowed per new edition. Probability theory tells us that an average of about 10^8 (hundred million) new editions plus their eliminations are required to find and change this single 'i' in an 'o'. The allowance of ten times more alterations would find the error ten times earlier, but would introduce each time more errors to the text, which would deteriorate. Biological systems are able to overrule such errors. One mutant of a population of 10^8 reproducing specimens would eliminate such an error in every generation.

This was an unusually simple case. Assuming now the word 'and' in one sentence should read 'but'. We take again 10^8 editions, or specimens of a population, to correct each one of these letters. But in a first success we would receive only 'bnd', 'aut', or 'ant'. Would this make the edition any better? Obviously not. The three letters would have to change together. This is possible in principle, but the probability drops dramatically, corresponding to $10^{-3\times8}$, or 10^{-24}. In other words, a population of 10^8 specimens would have to wait 10^{16} generations. This is physically impossible. Life on earth exists only 3.5×10^9 years.

And a three-letter word is still a simple thing. If we consider a sentence or a whole paragraph to be corrected, which, in analogy to a context, is the normal case. The correction must be done in another way. Whole words, sentences and paragraphs would have to be combined, and also be replaceable. Contemporary genetics accepts the existence of larger units, such as operons and homeoboxes. But it distrusts this being a basic principle, because its paradigm has no explanation of how the parts of such units come together.

(3) *The architecture of complexity*, if failed to be noticed, drives into another dead-end road. Knowledge from comparative anatomy and systematics tells us that the manifoldness of features unfold a hierarchy of layers. And all these layers originate as insertions between two selective agents; their constituents (*causa materialis*) and an environment (*causa formalis*) in which they develop. Take a biological tissue; the constituting cells are holding the pre-selective decision whether, e.g., myosin can be produced (to form a muscle) or not. And the organ, the environment in which the tissue develops, decides post-selectively, i.e., after the attempt of such a formation, whether it needs a muscle, and if so, in which position.

Bats have hair, although they would do better with feathers, this is due to *causa materialis*, while dolphins could have a fur, but do better without it; *causa formalis*. No doubt, *causa efficiens* is involved. But the action, e.g., of a sieve is obviously poorly described by only the power necessary to shake it. Two-sided causation cannot be avoided.

In all complex systems we need to consider two-sided causes for the origin of all its hierarchical layers; their constitution as well as their functional environment (deciding, what they are good for). No doubt, the two are in permanent interrelation. The concept of environmental selection provided by the current explanation is too simple; as if the assembly line of a car factory would have to wait for the customers to select the fitter fittings. It obscures the much more complicated selective processes within the systems.

(4) *Baupläne, type, styles and designs* are concepts generally used for artifacts. Not only do they represent exactly the synthetic insights into the basic patterns of all complex systems; they are also representative of the optimized confirmation of gained knowledge, such as the major components of the Natural System.

Author's address

Rupert Riedl, Konrad Lorenz-Institute for Evolution and Cognition Research, A-3422 Altenberg/Donau, Austria.

We may recall that our knowledge of this system is based on a large number of confirmed perceptions: hierarchically organized, unconsciously formed but evident theories of structures and classes, composed of two million species and half a million higher taxa, each times ten to twenty special characters; known as the specific homologies in comparative anatomy, or as terms of the 'differential diagnoses' in taxonomy.

This vast hierarchic system of orderly patterns can, in its stability over billions of years, only be understood, if a comparable pattern is foreseen in the genotypes or archaeogenotypes. Nothing of this kind gets its appropriate place in the current paradigm of evolution. It suggests that all the units in the genome are about equal, and having equal chances for successful alteration. This is probably the worst handicap in explaining evolution today.

While the theories of sensory organs or of behavior are still numerous, forming symmetrical hierarchies confirming one another (RIEDL, 1985), evolutionary theory is empty in its center part, bridging the whole complexity by assuming that changes in the 'blueprint' (the genome) and selection, in whatever layer in the phene system, could explain the phenomena of evolution.

(5) *The riddle about the tautology* of DARWIN's theory of selection was brought up by POPPER. In his autobiography he rates DARWINISM only as "a metaphysical research program" (1974, cf. Autobiogr. 1994 p.244) and not as a testable scientific theory, mainly because it covers everything and cannot therefore be refuted. In his first DARWIN Lecture (1977) he refines his standpoint, supported by authorities such as FISHER, HALDANE, SIMPSON, and the statement by WADDINGTON that natural selection is a tautology, but one with enormous explanatory power. "Since the explanatory power of a tautology is obviously zero", POPPER concludes (1978, p.143), "something must be wrong here."

The solution he offers is that "not all phenomena of evolution are explained by natural selection alone" (p.145). It is a pity that even POPPER did not notice DARWIN's Pangenesis Theory. And although POPPER's example is poorly chosen, he is absolutely right. He is also correct in saying that "by some verbal maneuver one can get round any refutation of any theory."

But we have not only found "large classes of facts" (in the sense of DARWIN) that cannot be explained by environmental selection. I may add that those characters that allow to recognize a homolog, are all together recognizable only because they resisted this kind of selection. And we know homologies in the millions. It is surprising—and this is my point in this investigation—to what extent a lack of knowledge excludes understanding.

(6) *The final paradox* in confronting the path of cognition and the path of explanation of evolution, i.e. knowledge and understanding, is that the phenomenon of evolution has been discovered by the confirmation of an orderly pattern, far beyond any possibilities of chance effects, and that, in contrast, the theory of explanation operates with only two stochastic terms.

Now: Can cognition be substituted by explanation? Obviously yes; but, within the given situation, only through the loss of vast amounts of knowledge.

Acknowledgments

I am greatly indebted to Werner CALLEBAUT for the careful reading of the manuscript and for making most valuable suggestions and amendments.

Notes

1 The term *ratiomorphic* was first used by LORENZ (1978) for the assembly of the inherited problem solving mechanisms, because it acts reason-like (*vernünftähnlich*), but not at all rational.

2 *Hermeneutic*: from Hermes, a recursive process of mutual enlightenment (details in RIEDL 1985).

3 *Morphology*: since GOETHE (1795), the methodology of comparing Gestalt and to generalize the Typus; the cognitive basis for comparative anatomy, taxonomy and phylogeny (RIEDL 1977a, 1980a, and 1980b).

4 *Scientistic*: is often used to focus the current paradigm in inorganic sciences, assuming that a system can be sufficiently understood by dissecting it to its components.

References

Adanson, M. (1752) Histoire naturelle de Sénégal: Coquillages. Paris.

Apostel, L. (1959) Logika en geesteswetenschappen, De Tempel, Brugge.

Aristotle see Ross

Asma, S. T. (1996) Darwin's causal pluralism, Biology and Philosophy 11, 1–20.

Baer, C. E. von (1828) Über Entwickelungsgeschichte der Thiere. Beobachtung und Reflexion. Gebrüder Bornträger, Königsberg.

Berlin, B. et al. (1966). Folk taxonomies and biological classification. Science 154, 273–275.

Bertalanffy, L. von (1968) General Systems Theory: Foundations, Development, Application. Braziller, New York.

Bljacher L. J. a (1960) Vozniknovenie kletok v ontogeneze. Trudi Instituta istorii estestvoznanija i techniki 32, 3–57.

Boeckh, A. (1877, [2]1966) Enzyklopädie und Methodenlehre der philologischen Wissenschaften. I. Formale Theorie der philologischen Wissenschaften. Neuausgabe (von 1877). Wiss. Buchgesellschaft, Darmstadt.

Bonet, C. (1762) Considérations sur les corps organisés. In: Oeuvres d'histoire naturelle et de philosophie. Neuchatel, 1779, T. III, chap. VIII, sect. CXL, p 90.

Buffon, G. L. L. Comte de (1749–1789) Histoire naturelle générale et particulière, 36 Bd.

Campbell, D. T. (1974) Evolutionary epistemology. In: P. A. Schilpp, ed., The Philosophy of Karl Popper, vol. 1, Open Court, La Salle, Ill., 412–463.

Cuvier, G. (1828) Le Règne animal. Bd. 1. 2ième éd. Verdière et Lagrange, Paris (1ière éd. 1817).

Darwin, C. (1859): The origin of species by means of natural selection; or the preservation of favoured races in the struggle of life. John Murray, London.

Darwin, C. (1875) The variation of animals and plants under domestication. 2 Vol. John Murray, London.

Dilthey, W. (1883): Einleitung in die Geisteswissenschaften. Neuauflage, Bd. I–VII (71973). Vandenhoeck und Ruprecht, Göttingen.

Dobzhansky, T. ([3]1951) Genetics and the origin of species. Columbia Univ. Press, New York.

Elster, J. (1983) Explaining Technical Change. Cambridge University Press, Cambridge.

Florkin, M. (1966) A molecular approach to phylogeny. Elsevier, New York- Amsterdam.

Foucault, M. (1966, [2]1974) Die Ordnung der Dinge; eine Archäologie der Humanwissenschaften. Suhrkamp, Frankfurt.

Gegenbaur ([1]1858, [2]1870) Grundzüge der vergleichenden Anatomie. Leipzig.

Geoffroy, Saint Hilaire É (1818–1822) Philosophie anatomique, vol. 1: Des organes respiratoires, Méquignon-Marvis, Paris, 1818; 2ième vol.: Des monstruosités humaines, ibid., 1922. Réimpression anastatique, Culture et civilisation, Bruixelles.

Goethe, J. W. v. (1795): Morphologische Schriften. Böhlau, Weimar

Gould, S. J. (1984) Darwin nach Darwin. Frankfurt/M., Berlin, Wien.

Grünbaum, A. (1984) The Foundations of Psychoanalysis: A Philosophical Critique. University of California Press, Berkeley/Los Angeles.

Grünbaum, A. (1986) Précis of Grünbaum (1984) [with open peer commentary and the author's reply]. Behavioral and Brain Sciences 9, 217–284.

Gutting, G. (1984) Paradigms and hermeneutics: A dialogue on Kuhn, Rorty, and the social sciences. American Philosophical Quarterly 21, 1–15.

Hacking, I. (1983) Representing and Intervening: Introductory Topics in the Philosophy of Natural Science. Cambridge.University Press, Cambridge.

Haeckel, E. (1866) Generelle Morphologie der Organismen. 2 Bände. Reimer, Berlin.

Hassenstein, B. (1951) Goethes Morphologie als selbstkritische Wissenschaft und die heutige Gültigkeit ihrer Ergebnisse. Neue Folge d. Jahrb. d. Goethe-Gesellschaft 12, 333–357.

Hassenstein, B. (1958) Prinzipien der vergleichenden Anatomie bei Geoffroy Saint-Hilaire, Cuvier und Goethe. Act. Coll. int. Strasbourg Publ. Fac. lettr. 137, 155–168.

Hempel, C./Oppenheim, P. (1948): Studies in the logic of explanation. Philosophy of Science 15, 135–175

Hillis, D./Moritz, C. (1990) Molecular systematics. Sinauer Publ. Sünderland (Mass.).

Hume, D. (1748) Enquiry Concerning Human Understanding. London. Modern ed,: Penguin Classics, Harmondsworth.

Kaleri (1982) Zum methodologischen Streit zwischen Natur- und Geisteswissenschaften. Zeitschrift für allgemeine Wissenschaftstheorie 23: 223–239.

Lamarck, J. B. de (1809) Philosophie Zoologique. Dentu, Paris.

Lamarck, J. B. de (1963) Zoological Philosophy with Regard to the Natural History of Animals. Trans. H. Elliot. New York/London: Fafner

Linnaeus, C. v. (1735) Systema naturae. Haak, Leiden.

Lorenz, K. (1978) Behind the Mirror. Methuen, London and Harcourt-Brace-Jovanovic, New York. Original German edition: Lorenz K (1973) Die Rückseite des Spiegels. Versuch einer Naturgeschichte menschlichen Erkennens. Piper, München, Zürich.

Mayr, E. (1942) Systematics and the origin of species. Columbia Univ. Press, New York.

Mayr, E. (1965) Numerical phenetics and taxonomy theory. Syst. Zool. 14, 73–95.

Mayr, E. (1969) Principles of Systematic Zoology. McGraw-Hill Book Company, New York, St. Louis, San Francisco, Toronto, London, Sydney.

Monod, J. (1971) Zufall und Notwendigkeit. Philosophische Fragen der modernen Biologie. Piper, München-Zürich.

Oeser, E. (1976) Wissenschaft und Information. 3 Bände. Oldenburg, Wien-München.

Osborn, H. (1934) Aristogenesis, the creative principle in the origin of species. Amer. Naturalist 68, 193–235.

Owen, R. (1848) On the archetype and homologies of the vertebrate skeleton. Brit. Asoc. Rep. 1846, 169–340.

Popper, K. (1957) The Logic of Scientific Discovery. Hutchinson, London. Original German edition: Popper K (51973) Logik der Forschung. Mohr, Tübingen.

Popper, K. (1978) Natural selection and the emergence of mind. Dialectica 32, 3–4.

Popper, K. (1994) Ausgangspunkte. Hoffman und Campe, Hamburg

Putnam, H. (1990) Realism With a Human Face. Harvard University Press, Cambridge, Mass.

Remane, A. (1952, [2]1971) Grundlagen des natürlichen Systems, der vergleichenden Anatomie und Phylogenetik. Koeltz, Königstein/Taunus.

Riedl, R. (1975): Die Ordnung des Lebendigen. Systembedingungen der Evolution. Paul Parey, Hamburg-Berlin.

Riedl, R. (1977a) Order in Living Organisms. Wiley, London. Original German edition: Riedl R (1975) Die Ordnung des Lebendigen. Parey, Hamburg, Berlin.

Riedl, R. (1977b) A Systems-analytical Approach to Macroevolutionary Phenomena. Q. Rev. Biol. 52, 351–370.

Riedl, R. (1980a): Die Entwicklung des Begriffs vom taxonomischen Merkmal oder das Problem der Morphologie. Zool. Jb. Anat. 103, 155–168.

Riedl, R. (1980b): Homologien; ihre Gründe und Erkenntnisgründe. Verl. Dt. Zool. Ge., Gustav Fischer, Stuttgart, 164–176.

Riedl, R. (1983) The role of morphology in the theory of evolution. In: Grene M (Ed.) Dimensions of Darwinism. Cam-

bridge Univ. Press, Cambridge.
Riedl, R. (1984) Biology of Knowledge. The Evolutionary Basis of Reason. Wiley, Chichester-New York-Brisbane-Toronto-Singapore. Original german edition: Riedl R (1980) Biologie der Erkenntnis. Die stammesgeschichtlichen Grundlagen der Vernunft. Parey, Hamburg-Berlin.
Riedl, R. (1985) Die Spaltung des Weltbildes. Biologische Grundlagen des Erklärens und Verstehens. Parey, Hamburg-Berlin.
Riedl, R. (1987) Begriff und Welt. Biologische Grundlagen des Erkennens und Begreifens. Parey, Hamburg-Berlin.
Riedl, R. (1992) Wahrheit und Wahrscheinlichkeit. Biologische Grundlagen des Für-Wahr-Nehmens. Parey, Hamburg, Berlin.
Riedl, R. (1994) Mit dem Kopf durch die Wand. Biologische Grenzen des Denkens. Klett-Cotta, Stuttgart.
Riedl, R. (1995a): Deficiencies of adaptation in human reason. A constructivistic extension of evolutionary epistemology. Evolution and Cognition 1 (1) 27–37.
Riedl, R. (1995b): Goethe and the Path of Discovery. An Anniversary. Evolution and Cognition 1 (2) 114–120.
Ross, W. D. (1928) The Works of Aristotle. Clarendon, Oxford
Schindewolf, O. (1936) Paläontologie, Entwicklungslehre und Genetik. Borntraeger, Berlin.
Schindewolf, O. (1950) Grundfragen der Paläontologie. Schweizerbart, Stuttgart.
Schmalhausen, I. (1949) Factors of evolution: the theory of stabilizing selection. Blakiston, Philadelphia.
Simpson, G. (1952) The meaning of evolution. Yale Univ. Press, New Haven.
Simpson, G. (1961) Principles of animal taxonomy. Columbia Univ. Press, New York.
Sneath, P./Sokal, R. (1973) Numerical taxonomy; The principle and practice of numerical classification. Freeman, San Francisco.
Sokal, R./Sneath, P. (1963) Principles of numerical taxonomy. Freeman, San Francisco.
Starck, D. (1978) Vergleichende Anatomie der Wirbeltiere; auf evolutionsbiologischer Grundlage. Vol. I. Springer, Heidelberg-New York.
Stegmüller, W. (1974): Der sogenannte Zirkel des Verstehens. In: Hübner, K. u. Henne, A. (Hrsg.): Natur und Geschichte. Rowohlt, Hamburg, 21–46.
Stent, G. (1981) Cerebral hermeneutics. Journal of Social and Biological Structures 4, 107–124.
Thirring, W./Stöltzner M. (1994) Entstehung neuer Gesetze in der Evolution der Welt. In: Naturwissenschaften (81) p 243.
Toulmin, S. (1961) Foresight and Understanding: An Enquiry into the Aims of Science. Harper and Row, New York/Evanston.
Voigt, W. (1973) Homologie und Typus in der Biologie. Gustav Fischer, Jena.
Waddington, C. (1957) The Strategy of the Genes. Allen and Unwin, London.
Wagner, G. (1983): On the necessity of a systems theory of evolution and its population biologic foundation. Comments on Dr. Regelmann's article. Acta Biotheoretica 32. 223–226.
Wagner, G. P. (1985) Über populationsgenetische Grundlagen einer Systemtheorie der Evolution. In: Wagner, G., Ott, J. und Wuketits, F.M. (Hrsg.): Evolution, Ordnung und Erkenntnis. Paul Parey, Berlin-Hamburg, 97–111.
Wagner, G. P. (1988) The significance of developmental constraints for phenotypic evolution by natural selection. In: Jong, G de (ed.) Population genetics and evolution. Springer, Berlin-Heidelberg.
Wallace, A. R. (1889) Darwinism. Macmillan, London.
Whyte, L (1964) Internal factors in evolution. Acta Biotheoretica 16, 33-48.

Detlef Weinich

A Contribution to Constraints in the "Civilizing Process"

A Model taken from Norbert Elias

Abstract

This study is to be regarded as a contribution to interdisciplinary research and represents an attempt to clarify the question of whether and to what extent concepts that have been developed in the field of theoretical biology and which have a high degree of importance here can also be applied to sociological phenomena. In particular it is intended to examine the question of whether the civilizing process can be adequately treated using the evolutionary concept of 'Constraints'. In the clarification of this question, special reference is made to the 'theory of the civilizing process' by Norbert ELIAS, since here a highly respected scholar[1] has presented an important sociological theory. Moreover, there is such good scientific access to ELIAS because this author exemplifies his theses in historical terms and thus to a certain extent makes his explanations verifiable in scientific terms.[2] In the treatment of this topic, the central terms and theses of ELIAS will be presented from the considerable scope of his work, and then illustrated with the help of several selected historical case studies. Furthermore, reference will be made at the relevant points to parallels and analogies which the works of ELIAS have to other, predominantly system-theoretical concepts of evolution and which cause it to appear compatible to the latter.

Key words

Constraints, necessity and chance, order on order, civilizing process, immanent regularities, social figuration, interdependency, social institutions, royal

On the Person and the Works of Norbert Elias

On the person and works of Norbert ELIAS: He was born in the silesian capital Breslau (nowadays Wroclaw, Poland) in 1897 as the only child of the Jewish couple Hermann and Sophie ELIAS. The "weak child" (KORTE 1988, p67) did well at school, passed his final leaving examination in 1915 and matriculated at the university of his home town, initially to study philosophy and German language and literature. After an excursion into the medical faculty, ELIAS devoted himself fully to philosophy and became a Doctor of Philosophy in 1924. He did his oral examination in the subjects of philosophy, psychology, the history of art and chemistry (!). In his dissertation— subject: "Idee und Indi-viduum – Ein Beitrag zur Philosophie der Geschichte"—he treated topics which were later to dominate his main work.

ELIAS then started with his sociological interest at the University of Heidelberg, to which he came in 1924—with the intention of becoming a university lecturer. There he obtained a position as an assistant, and came into contact with the two opponents Alfred WEBER and Karl MANNHEIM. ELIAS reported that the more materialistic-orientated MANNHEIM gave him not only a new theory, but also a "specifically new experience of life" (KORTE 1988, p104), which then became the basis for his own works. ELIAS formulated these particular ideas on the civilizing process in his two-volume work with the title "The Civilizing Process. Sociogenetic and Psychogenetic Investigations".[3] At the time of its first publication in 1937[4] ELIAS had already left Germany because of the current political situation and had gone into exile in England.

Reference will essentially be made in the following to this epoch-making work, which—due to the war—fell into oblivion and, since its rediscovery in 1976, is today regarded as a "classic" (KORTE 1988, p33; similar: ARIÈS/DUBY 1991, p23)[5].

The Civilizing Process

The term 'civilization' makes some introductory remarks necessary: It is regarded as one of the most ambiguous terms in the whole of sociology, and it has completely different connotations in the various European languages[6]. Although ELIAS believes that the term 'civilization' expresses a specific Western experience ("this concept expresses the self-consciousness of the West" (Vol.I, p3), he also concedes that it includes such different phenomena as "level of technology", "type of manners", "development of scientific knowledge", "religious ideas and customs", "type of dwelling or the manner in which men and women live together", "the form of judicial punishment", "the way in which food is prepared" (Vol. I, p3). In his own investigations into the civilizing process, ELIAS puts his main interest in the change in the behaviour of people. He regards human affects—or, to be more precise, control over them—as an important yardstick for the degree of civilization.

It must also be emphasized that ELIAS' theory of the civilizing process is by no means '*the* academic opinion of sociology' on this topic. In this point, sociology presents itself in an unfortunate situation, which is characterized by a mass of different, in some cases mutually contradictory approaches. The reason for this situation is that until recent times there has been no consensus within the 'scientific community' about the question of *what* is to be regarded as 'social change', nor *what the causes* of this change are.[7]

Immanent Regularities of the Civilizing Process

Having now dealt with some preliminaries and thus the statements by ELIAS: the main focus of his argumentation is the thesis that it is not the "rational understanding" by individual persons or planning by larger groups of people that represent the "motor of the civilizing process" (Vol. I, p116), but on the contrary, that it is the 'immanent regularities' of the civilizing process itself (Vol. II, p152) that are finally responsible for the generation of new aspects in the civilizing process.

With this concept, ELIAS dissociates himself from two contradictory positions in the study of history: he denies that the civilizing process is a historical event which is unique, and which has occurred as such by chance and not otherwise, and about which forecasts are logically not possible because it is not based on any recognisable principle. ELIAS rejects such a conception just as he rejects deterministic views according to which the civilizing process is subject to a 'Law of succession', thereby creating a series of characteristic historical eras. The latter views in particular—Karl POPPER (1936; cf. 1965) called them 'historicistic'—have a long tradition and have played an important role in sociology right up to the present day.[8]

ELIAS characterises his own views on the subject of 'immanent regularities' by stating that "Taken as a whole this change is not 'rationally' planned; but neither is it a random coming and going of orderless patterns" (Vol. II, p230).

This brief statement and the subsequent explanations make it clear that with these 'immanent regularities', ELIAS is developing a concept whose basic approach sounds familiar to biologists: these are essentially ideas taken from systems theory, like those of Rupert RIEDL, who says that the course of evolution is controlled by chance and necessity (cf. 1982, pp221).

RIEDL discerns chance at the very lowest level of being in the form of the indeterminate character of microphysical processes. This microcosmic coincidence extends upwards to higher levels of organization, at each stage of evolution acquiring new, specific forms and thus takes part in the macroscopic perceptible events of everyday life. At biological level, for example, chance takes part in evolution in the form of mutation and recombination.

Necessity gains access to this process because as structures appear (by chance), the degrees of freedom for further developmental processes are reduced. Due to this restriction of the effect of chance, caused by 'systemic conditions', at all stages of evolution, not every conceivable mutation can take place at biological level, as its consequences would be so serious (as a disruption to the entire system), that the system in its entirety could no longer exist.[9] In this connection, RIEDL talks of an 'internal selection' (RIEDL 1976, p140) which starts as early as the embryonic stage of the individual, and thus becomes effective long before the classic, DARWINIAN selection by the environment (RIEDL 1976, pp157). This latter form of selection (RIEDL talks of it as a 'market selection') is only effective in about 5% of all cases and is thus considerably less important than 'internal selection' through 'internal factors', which eliminate 95% of the variations that are created by mutations. RIEDL therefore rightly finds that it is by no means DARWIN's 'struggle for existence' that is decisive for the survival of a phenotype, but the examination of functionality takes place at an earlier, more basic level (RIEDL 1976, p162). This view of the subordinated importance of the environment as an evolutionary factor is very aptly underlined by the evolution theorist Walter NAGL, who says that selection by the environment is merely a mechanism for rounding off the edges of the results of evolution (1993, p6).

This interplay between chance and necessity gives, in the last resort, a direction to the evolution process. The creative power of chance and the constraints that are caused by the products of its acting thus gives rise to a path along which it is only possible to return with the greatest of difficulty and through which completely 'new' phenomena appear. On account of this innovative power (resulting from the directed nature of evolution), RIEDL talks of a 'creative process' (RIEDL 1982, p225), which possesses a momentum of its own and which reaches up from physical evolution as far as the field of cognitive capabilities. This formal similarity which the evolution process thus shows to learning processes is the reason why RIEDL (following Konrad LORENZ) talks of evolution as a 'cognitive process' (RIEDL 1982, p223).[10]

Such a creative process driven by its own momentum—this is how ELIAS would be interpreted from this perspective—also exists in relation to the 'civilizing process'. Within this framework, those sequences of events and conditions occur which—although not predictable in principle—nevertheless possess a non-random character and of which ELIAS says that these are *not* 'orderless patterns'. Within the framework of the civilizing process—this is how ELIAS would be interpreted from the point of view of RIEDL—chance exists in the form of freedom of individual decision-making, while the channelling of chance is caused by those immanent necessities, which will be explained in the following, and which result from the collective actions of individuals.

The Principle of 'Social Figuration'

Another basic sociological premise with ELIAS is the idea that the individual and society are not entities that exist separately from each other. In this context, ELIAS stresses that all external relations of an individual *also* have a social aspect and thus that the individual and the collective areas determine and permeate each other (Vol.I, pp245)[11]: According to him, "Human beings only exist in pluralities, and it is the network of their interdependencies that binds them together" (Vol.I, p261). This is how those groups of interdependent human beings to which ELIAS assigns the term "figuration" (ibid.) are formed. They represent the basis for his concept of the civilizing process as a process of progressive "social figuration" of the individuals involved[12]: the civilizing process is thus characterized by the progressive interdependency of individuals, and it has its main roots in the growth in population[13] and in the limitation of physical space.

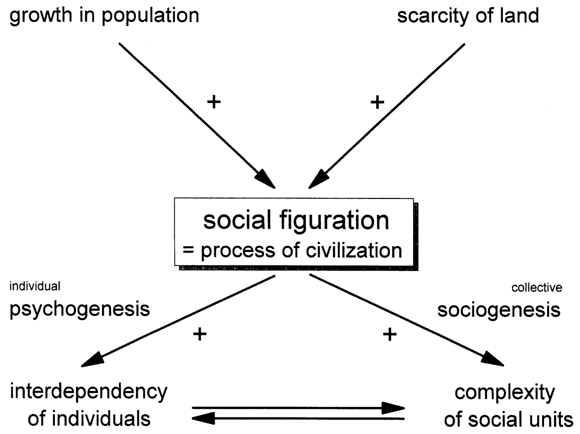

Figure 1: Social Figuration

On account of the inseparability of the society and the individual, ELIAS distinguishes two levels from an analytical point of view: a *'psychogenesis'* and a *'sociogenesis'.*

With his thinking in terms of social configuration, ELIAS also focuses on the dynamic character of society: dynamics—in agreement with the ELIAS follower Johan GOUDSBLOM (1984)—is a constitutional element of all social phenomena, and all 'nomothetic thinking'[14] is out of place, as it tries to reduce transformations to something which itself is supposedly unchangeable. The failure of this thinking, according to this author, comes from the notion that we cannot revert to some unchangeable principle, either in the description of individual changes or in the diagnosis and explanation of long-term processes (1984, p98).

Despite his passionate attacks against all concepts which postulate the existence of principles independent in time, it must be pointed out that with this concept of 'social figuration' ELIAS is also formulating such an abstract, valid, time-independent principle. However, this principle—as will be shown below—is so general and highly dependent on the epoch in question in terms of subject matter that it cannot be compared with the competing approaches of other authors[15]. The principle of 'social figuration' is, one would like to say, a 'systemic' condition, whose existence is located upstream of the social field.

Feudalization

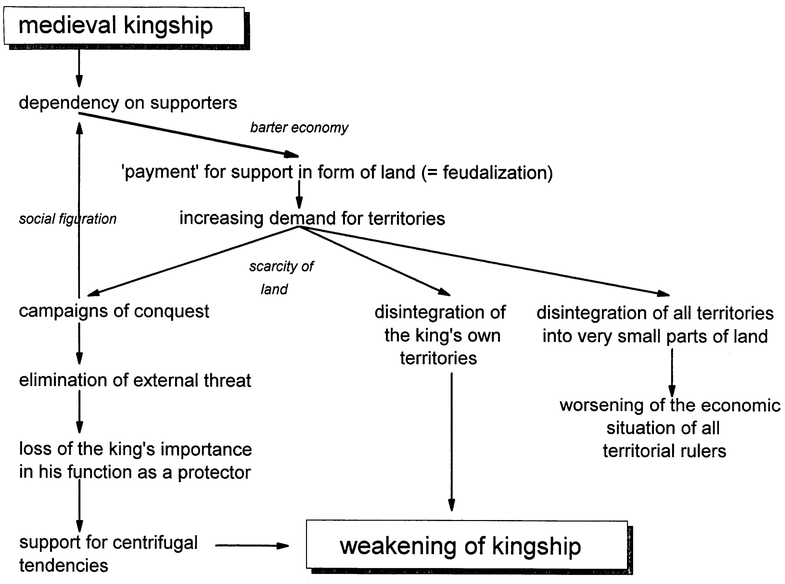

Figure 2: The process of feudalization.

In order to exemplify this concept, attention will now be drawn to the process of 'feudalization', which plays an important part in ELIAS' thinking: this also seems to be opportune as in this historical process the influence of the two central principles of ELIAS' theory of the civilizing process can be clearly demonstrated. These are the principle of history as a progressive 'social figuration' of those involved and those 'immanent regularities' which the civilizing process develops itself and from whose constraints the actors cannot withdraw.

One of the main characteristics of the early Middle Ages[16] is that the population density, which had fallen sharply in the time of the migration of peoples, gradually rose again (Vol. II, pp30). In this situation of a slowly growing population—and the related reduction in the availability of physical space—the 'social figuration' of the individual intensified continuously; for there was competition among the territorial rulers for land, which was becoming more and more scarce and which represented the economic basis for a territory.

Another characteristic feature of this time was a type of economy in which the goods produced were consumed locally, without long transport distances and intermediate trade channels, or in which they changed hands over by short exchange chains (Vol.II, p50). Such a so-called 'barter economy' was particularly characterised by money only circulating to a limited extent, and division of labour and social differentiation accordingly playing only a subordinate role.

In political terms, this phase was characterised by a fragmentation of the territorial areas. A large number of territorial rulers governed more or less small areas which they had received as fiefs from the king. The basis of this fief relationship was a mutual interdependency, which implied that the vassal was obliged to take up arms for his lord, while the king, with his royal protection, guaranteed the integrity of the possessions of his vassal against other territorial rulers (Vol.II, p61).

In order to meet these obligations towards his lord, the knight had to raise from his subjects levies, whose 'countervalue' consisted in him giving his subjects a promise to protect them from attacks by other territorial rulers. On account of the particular economic situation in the Middle Ages, these levies were in the form of the natural produce of the immediate surroundings of the knight's manor.

The following diagram is an attempt to illustrate this complex mixture of economic, social and political factors, as well as the resulting consequences for the kingdom:

To be able to conduct his campaigns, the king was dependent on followers, whose allegiance he could only be sure of—in a time of barter economy—by handing over land to his supporters as fiefs. However, this form of 'payment' had the side effect that in order to reward his followers, the king had to either conduct permanent campaigns of conquest or give away parts of his own territory, that is, his own power, as fiefs. As a consequence, the monarchy became weaker as the land to be conquered became scarcer and the reward of land as fiefs also took place at the expense of the actual royal territory, the 'Hausmacht'.

This decline in the importance of the monarchy was also accelerated as the king eliminated his enemies in successful campaigns, and thus one of his most important functions—the granting of protection for his subjects against external enemies—became dispensable.

This trend (systematic reduction of the power of the central ruler—the king—and simultaneous strengthening of the position of the territorial rulers) makes clear how a process gains its own momentum due to forces inherent in the system: this tendency towards weakening of the monarchy due to the splitting up of the available land into very small territories prevailed, and in historical terms this finally led to lasting change in the balance of power at the time, that is, to the disappearance of a strong monarchy. Another important aspect of this development is that a form of automatism which could not be stopped and which prevailed against the express wish of the rulers had been set in motion.

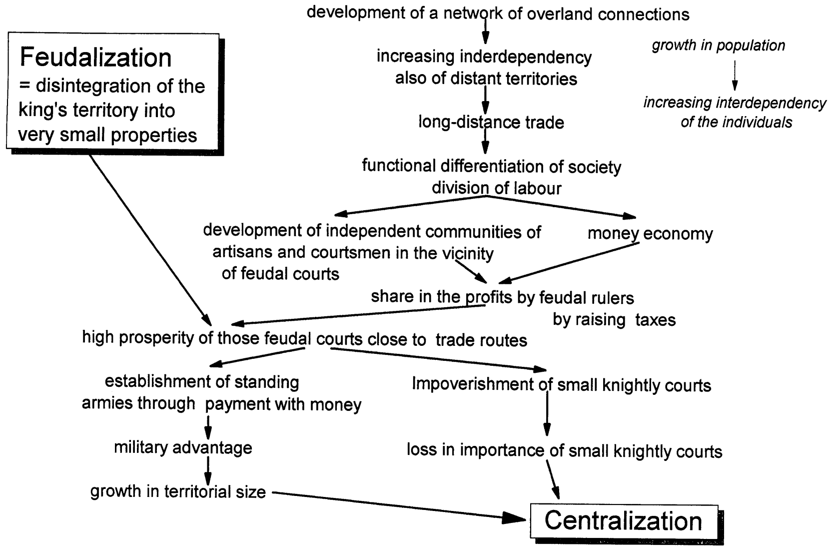

Figure 3: The prozess of centralization.

The Centralizing Process

This trend towards a particularization of power and the creation of small and very small territories did not, however, last long. On the contrary, it was brought to an end by a series of inherent constraints and was succeeded by a tendency towards enlarging the territories.

It is worth noting in this context that this process of 'centralizing'—which reached its climax in 'Absolutism'—actually had the same causes as the original tendency towards decentralizing, namely the progressive social figuration of single individuals. This trend, which began in the late Middle Ages and which countered the splitting up of territories, merely resulted from the general social conditions having changed compared with the early Middle Ages, and the principle effective up until then—progressive social figuration—now having a completely contrary effect. How is this to be understood?

While trade in times of low population density was of a largely domestic nature, and while the village around the lord's manor supported itself largely with its own products, this situation changed successively during the Middle Ages. The growth in population made it necessary to develop new areas of settlement, and this was implemented by clearing woodland, draining marshy land and similar cultivation measures. At the same time—for the first time in history—an extended network of roads was created inland to connect these centres of settlement with each other. This improved infrastructure, as well as the utilization of the horse as a means of transport for heavy loads and over greater distances in the 13th century (Vol.II, pp66) gave rise to a lively trade which differed from the original form of trade in that money acquired a greater degree of importance as a means of exchange due to the longer and longer trade chains, thus replacing barter trade (Vol.II, pp58).

The resulting money economy had a large number of far-reaching consequences for the feudal society and, in the final resort, is one of the major causes for the slowdown in the trend towards more and more splitting up into smaller and smaller territories.

This new situation thus promoted the internal differentiation of society (Vol.II, pp62): the division of labour was introduced, and in the wake of this communities of craftsmen and tradesmen were formed in France in the 12th century as the preliminary stages of urban settlements around the seats of the feudal lords. The loss of self-sufficiency in the barter economy led to greater dependence of the people on each other and to a higher form of dependence of the ruler on his subject than was the case in the older tribal society (Vol.II, pp80). All in all, this led to a higher degree of interdependence within society as a whole.

However, the fact that the rulers participated in this economy by raising taxes in the areas in which trade flourished on the basis of the money economy was important and decisive for the development of court centres as examined here. This led to the possession of money and thus to a new form of wealth in a society that up until then had considered only the possession of land to be the highest material good. However, this new form of wealth passed by those knights who—away from the great trade routes—ran a small estate in a small territory, raised levies from their serfs and thus led a modest, secluded life.

This trend, based on commercialization and monetarization, was in the final resort to the advantage of few powerful estate and feudal rulers, who were able to use financial means to achieve their political goals and thus continuously to increase their influence and wealth, by buying up more land and setting up 'standing armies'. The particular aspect of this development was that the payment of these 'soldiers' did not lead to the fragmentation of the ruler's own territory due to the distribution of land, but even contributed to the enlargement of the territorial power base of the warlords.

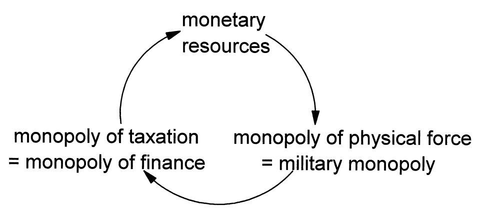

Figure 4: The prozess of monopolization.

The trend towards centralization thus introduced was additionally favoured by some changes in the general conditions: The development of long-range weapons led to the upgrading of the (untitled) infantry and made the (titled) armoured riders of the Middle Ages militarily meaningless (cf. P. MEYER 1977, pp76).

The feudal ruler who had the financial means to pay a large army of untitled warriors was thus not only tactically in a better position, he was also in a stronger social position vis-à-vis his vassals. The latter, as mounted fighters, had become dispensable as comrades-in-arms on account of their reduced military value and thus were only of slight value for the land they had been awarded.

This particular constellation of factors meant that towards the end of the Middle Ages al large number of 'free knights' had long since entered the sphere of influence of the great knightly feudal courts, had become militarily unimportant and economically impoverished (Vol.II, p72).

It must once more be mentioned that this entire historical development of the centralization of power was not intended by individual persons or planned by particular groups of persons (e.g., Vol.II, pp98 and 160). In each phase of developments, all of those involved were only looking for individual advantages that could be realised at short notice, and did not behave according to collective requirements or long-term objectives in any way. According to Arthur BOGNER (1991, p54), social processes, in ELIAS' view, are processes planned in the short term but not in the long term, as none of the actors is powerful or omniscient enough to be able to calculate and control their course in more than temporary terms. Although this statement suggests that BOGNER believes in the existence of the 'LAPLACE demon' (cf. LAPLACE 1814) and certainly still needs to be amended in the light of the 'Chaos theory' (cf. BRIGGS/PEAT 1990), the basic statement—that complex processes cannot be controlled for longer periods of time—cannot be contradicted from the view of the chaos theory. On the contrary, this shows that the inner conditions are important for the development of a system: with respect to the civilizing process, they exist in the form of inner constraints which result from the increasing social figuration of those involved and the specific inner condition of society in the Middle Ages. According to ELIAS , this is the reason why specific consequences appear to follow certain laws within the framework of a historical situation.

There are also obvious parallels to Karl POPPER (1965) here. He also emphasises that social processes can only be overlooked for short periods of time, and thus cannot be controlled to any great extent. However—and this is the main difference between the two authors—POPPER believes that this short-term calculability can be used to make developments controllable and thus to make the planning of social processes possible to a limited extent (he talks here of 'Stückwerk-Sozialtechnik').

In view of the historical facts, which speak in favour of ELIAS' theories, it may even seem legitimate to speak of a process that *constructs itself*, for it takes place even against the express wish of those involved, and follows only its own "relational dynamics" (Vol.II, p161): the decline of a group, or social class, namely that of knighthood, therefore took place within this historical development.

The theses of ELIAS, according to which non-intentionality is not an exception to the rule, but unplanned interaction between intended acts play a decisive role in the civilizing process[17], clearly contradict concepts in milieu theory which believe in the plannability of social conditions and attribute the failure to achieve certain goals to the wrong choice of means used to reach these goals.

These ideas are reflected at scientific level in the views of the physicochemist Manfred EIGEN, who aptly remarks that "everything that happens in our world resembles a big game in which nothing exists a priori but the rules" (1975, p11). These defined 'rules'—this is how ELIAS should be interpreted with respect to the civilizing process—represent the *increasing social figuration* and the *growing interdependence* among those involved. They are the reason for the social changes that took place in the Middle Ages.

The Royal Mechanism

ELIAS believes that the process of increasing centralization as an "inherent law" can be seen most clearly in France. ELIAS talks here of the so-called "royal mechanism", which was able to manifest itself most typically in France due to a number of peculiar cultural-geographic aspects in that country (cf. Vol.II, pp96).

Here too, ELIAS regards the increase in the population and the resulting scarcity of land suitable for cultivation as the main reason for this development. In the barter economy of the Middle Ages, land was the most important factor of production and therefore a highly desirable good, which led to keen competition (Vol.II, pp96). The result of this rivalry was a compulsion to expand, which was imposed on all those involved, if they did not want to perish. Thus not only the lowly knight, who fought with his peers to enlarge his territory, but also the great feudal rulers (who had been victorious in the competitions with their peers) were forced to fight to enlarge their territorial power bases.

This means that there was a hierarchy of competitions: the victor of the duel between two minor knights was able to increase his power, but he was soon faced with an opponent as powerful as he, and so there was another battle. Here, ELIAS describes a dynamism which did not permit any of those involved to quit the stage at one level, once events had got going.

ELIAS is able to exemplify this thesis excellently on the basis of the historical development in France. At the end of these competitions, which filled the whole of the Middle Ages, the victor was the absolutist ruler, who now enters a new arena of conflict, namely the conflict between states.

This process of centralization and monopolization, which ELIAS describes as not always direct and as falsified by a series of temporary contradictory tendencies (Vol.II, pp131), represents in its entirety a development towards social systems of greater complexity. Each competition gives rise to a new system which differs from the systems it emerged from in that it has a higher degree of interdependence among its units and thus a higher degree of integration. Thus towns and trade can only be formed after the territories have exceeded a certain size (Vol.II, pp129).

The 'Order-on-Order' Principle

It is worth noting that this process of centralization described by ELIAS is subject to analogous principles which, in a large number of points, play a part in biological evolution. For example, the idea of 'competition', which is particularly emphasised by ELIAS , is a classical DARWINIAN evolutionary factor which has its origin in the natural 'overproduction of descendants'. Although critical concepts of evolution (cf. GUTMANN 1989) stress the role of competition slightly differently, their validity as an important evolutionary factor is largely undisputed.

It is also significant that the process of territorial expansion described by ELIAS possesses some similarity with the principle formulated by the Austrian physicist and Nobel Prize winner (1933) Erwin SCHRÖDINGER as the "order-on-order principle" (1944). Its main statement, that order can only be created from another form of order, but never from nowhere, means that an evolving system can only develop further at the expense of other competing systems. The latter fall to a lower level of differentiation as they contribute to the differentiation of the 'higher' system.

This principle is regarded as valid for biological and prebiological systems, and it seems to reach up to higher levels of being ("Schichten des Seins", HARTMANN 1940). With the help of the statements by ELIAS, it also becomes clear that for its further development, a social system requires not only energy (which can be obtained from nature in the form of high-energy raw materials). On the contrary, an evolving social system takes away the development chances of a competing and in its complexity more or less comparable system, thus causing the downfall of the latter system.

The "monopoly mechanism" (Vol.II, pp104) is perfectly characterized by Rupert RIEDL, who states that "Civilizations generally eat civilizations" (1976, p271): he states that the victor gains influence and transposes his territory into a state of greater integration, while the defeated party loses importance and may even cede his whole territory to the victor.

What are the consequences, in social terms, of this decline, which takes place when one civilization is eaten up by another—as RIEDL states—and which, from an thermodynamic point of view, is to be regarded as a process of increasing entropy, or the loss of order? It is easy to see that a takeover of such a territory by a successful territorial ruler is preceded by a process of decline in the territory of the defeated party. Even if it is assumed that this decline is not an objective one, but is only a decline in relation to the neighbours, the weakening influence of a mediaeval knight should have been perceptibly painful for his subjects, for they were subjected to a great extent to the arbitrary action of the neighbour (Vol.I, pp191). Due to the weak 'molding of affects' of human beings in the feudal era and their high propensity to violence, this meant not only oppression in the sense of forcing them to pay levies: the rudimentary monetary trade made it considerably more advantageous for a victor to weaken the economic basis of his opponent by killing or maiming his subjects, or by destroying his means of production by burning his fields, filling in wells or cutting down trees (Vol.I, pp191).

The consequences in the field of jurisdiction were particularly drastic. The weakening of central power meant that law was no longer ensured in the interior. When Norbert ELIAS states that "Robbery, fighting, pillage, family feuds—all this played a hardly less important role in the life of the town population than in that of the warrior class itself " (Vol.I, p198), it becomes clear that lawbreaking in dealings among the people of a territory was very common at that time on account of the rudimentary 'tempering of affects' (cf. pp199), and it becomes clear that the weakening in the field of jurisdiction, or, to be more precise, the weakening of the organs responsible for maintaining law and order, inevitably led to considerable insecurity in daily life.

A consequence of this decline in the authority of a territorial ruler was that the inhabitants of this territory ran the risk of being victims of a crimes through which they suffered personal injury or material damage. In the final resort, this resulted in an increase in high risks and thus a reduction in life expectancy on account of the use of force or inadequate material supply[18].

As well as these quantitative consequences, the decline of a territorial ruler also had a number of other consequences that can be assigned to the cultural sphere. The entire complex of these phenomena could best be characterized as 'naturalization of society'[19]: For example, the increase in insecurity in matters of law and order meant that the individual was increasingly occupied with natural things and archaic living constraints and therefore restricted the production of goods to those required for daily life. Furthermore, enmity with neighbours and the related restriction of trade to the inhabitants' own—very small—territory[20] meant that only products that were necessary for physical survival were produced, because they could not be obtained elsewhere and because they only possessed any value in this agrarian society, a society that had returned to the barter type of economy. The related decline in the differentiation of the working process meant that the individual person was occupied to a considerably greater extent with carrying out general, non-specific activities and that special skills were of lesser value. This reduction in the degree of social interdependence had particularly serious effects in the artistic sector. Thus the disappearance of long-distance trade and remote markets meant that the demand for artistically produced objects, if not commissioned by local territorial rulers, dropped sharply, because these products had no exchange value whatsoever in a society characterized by barter and oriented towards securing physical survival.

Summing up, this development can be interpreted as follows: due to his enforced dedication to material values, the human being turns away from transcendental values and thus sacrifices specifically human fields of activity: he thus becomes 'naturalized'.

On the Role of Institutions

The above considerations of the consequences of the decline of a territorial realm for its inhabitants make one important aspect clear: within the framework of the civilizing process—to maintain a social community—the development and preservation of a whole series of controlling and coordinating authorities become necessary.

In this connection, ELIAS talks of " centralized monopoly institutions", which guarantee "an abundance of ... measures of co-ordination and regulation" (Vol.II, p163), such as

- the flourishing of the division of labour itself,
- the securing of routes and markets,
- the regulation of minting and monetary movements,
- the protection of peaceful production,

and finds that "from a certain degree of functional differentiation onward, the complex web of intertwining human activities simply cannot continue to grow or even to function without co-ordinating organs at a correspondingly high level of organization" (ibid.).

The contours, still blurred here, of the so-called 'social institutions'—such as economy, justice, politics—can be detected here; Arnold GEHLEN (1956, cf. similarly SCHELSKY 1952; PARSONS 1951) sees their particular achievement in the fact that they fulfil a double function, which the Austrian sociologist Wolfgang LIPP, who teaches in Würzburg, characterizes with the following words: they "contribute, at anthropological level, towards 'setting up' the human being, that is, securing him in part against his own 'risky', 'regression-susceptible' nature, in part against external disturbances—so that 'relief' (GEHLEN) takes place and thus existence becomes free for higher, culturally creative achievements; they have the effect, at a sociological level, of allowing society as a whole—as a mechanism of diverse part-systems, each of them specialised—to remain functionally together and to continue to exist" (LIPP 1989, p94).

Even though ELIAS recognises the importance of the above institutions, his statements would still have to be specified to the extent that it appears to be an essential condition of the civilizing process that the latter—on account of their 'immanent regularities'—produce the 'institutions' which secure the civilised achievement due to their special functions as controlling and coordinating authorities, and thus form a basis for further stages in the civilizing process.

From the perspective of systems theory, 'institutions' are the already mentioned fixtures which are so important within the framework of the evolution process because they 'channel random events' and thus make further developments possible. On the basis of this function, 'institutions' can thus be regarded as the 'constraints' which are effective at the level of developed social systems in the field of cultural evolution[21].

This term of 'constraints', which has only recently been introduced into the discussion as an evolutionary factor, comprises all of the internal factors which influence the further course of the evolution of a system by ruling out certain possibilities, thus showing a limiting effect. Although 'Constraints' go beyond the scope of DARWINIAN teachings about selection by the environment[22], they are increasingly accepted today as evolution factors by well-known exponents of DARWINIAN theory (cf. MAYNARD-SMITH 1985). The increase in popularity of 'constraints' is also an expression of the rediscovery of a phenomenon which was originally expressed by Rupert RIEDL and was introduced by him into German literature in the seventies (RIEDL 1975). In the language of RIEDL, 'constraints' are the so-called 'systemic conditions of evolution', which are to be understood as a consequence of the 'functional and genetic burden' of a system: the reduced options for different developments caused by the 'constraints' represent the disadvantage which a system has to accept—as RIEDL says—for gaining advantages in the speed of its development by systematising its functions (RIEDL 1976, p191).

It must be pointed out that in the course of the civilizing process the role of the 'constraints'—that is, of the institutions—consists in the securing of what has been achieved and thus making further social developments possible in the first place. To use a comparison from the field of architecture, they are the foundations of a building, and they assure that the statics of the next storey are guaranteed. However—and this must be emphasised once more and in agreement with ELIAS—these institutions, which have a limiting effect on possible future developments, are not the selective environmental factors as defined in DARWINIAN evolution theory, but they are products which the civilizing process itself has created as a 'creative process'.

The Opening-Up of the Monopoly

The importance of institutions as 'constraints in the civilizing process', which—once created—never disappear, and thus give a further direction to the civilizing process, is shown very clearly by the further historical development, which was characterized by the 'royal mechanism', (that is, the tendency towards monopolization), subsequently leading to the monopoly slipping out of the hands of the sole ruler and thus to an "opening-up of the monopoly" (Vol. II, p162).

Where does ELIAS see the causes for this new trend, which counteracted the process of monopolization and meant that the monopoly which was ideally concentrated in the hands of the absolutist ruler did not survive and finally vanished?

In this development too, ELIAS sees an 'immanent regularity': opening-up, he argues, occurs "when the area it controls ... begins to grow very large" (Vol. II, p162) and exceeds "an optimal size of possessions" (Vol.II, p109). The cause of this reversal of the original centralizing tendencies is that the monopolist is no longer in a position to consume all of the profits of his monopoly alone when the size of his area of influence has grown too large (Vol.II, p112). However, as the monopolist gives his subjects opportunities, his dependency on the latter increases, thus undermining the power of disposition of the monopolist in the long run. This starts up a process in which parts of the monopoly come into the power of disposition of those persons originally entrusted with administering the monopoly—for example civil servants—and is later extended to other groups in a society in which the division of labour is gaining importance (with a high level of interdependence between individuals!). This growing interdependence of individuals is the basis for this monopoly finally being withdrawn from the power of a sole ruler and being passed on into the hands of larger and larger collectives, until finally 'the people' disposes of these now 'public' monopolies in a democratic way.

From the point of view of institutional theory, this development means that the role of these institutions as non-personal controlling authorities increased as the monopoly was passed from the power of disposition of a central ruler "into the power, or at least the institutionally secured control, of broad social classes" (Vol. II, pp161). These tendencies were also promoted by an increasing division of labour and differentiation of society: this complexity of relations between individuals meant that the institutions with their controlling and coordinating functions became indispensable in the long run and their significance was not only restricted to critical phases, for example times of threat by external forces.

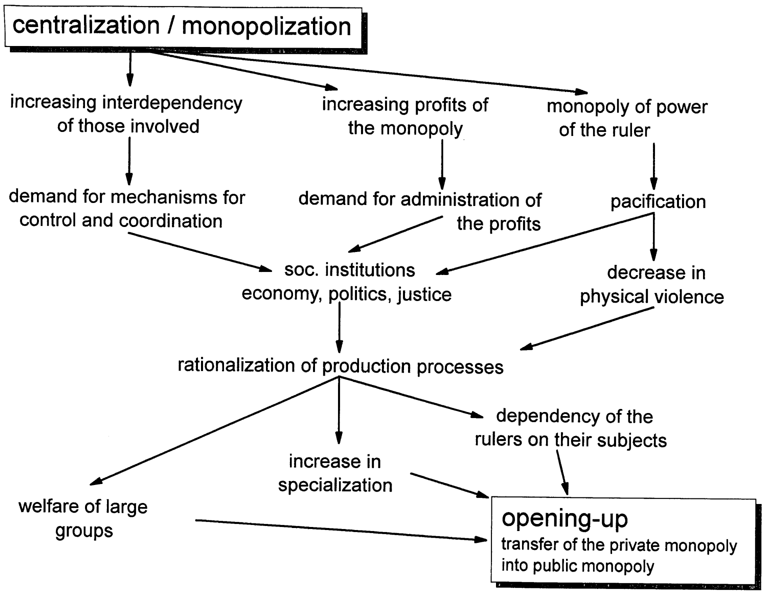

Figure 5: Opening-up of the monopoly.

With regard to the identification of institutions as the 'constraints' in the civilizing process it is particularly important in this development that the existence of these institutions, once established, is not endangered, by analogy with the 'constraints' of biological systems, but only modified. Thus the regulating and coordinating central organs, which had become indispensable due to a wider differentiation in society, changed their occupants and even their organization, but were not affected in their essence or even dissolved (Vol.II, p164).

Today, at the end of this process, the constitutional democratic state exists as a controlling authority of an enormous magnitude, which the Berlin philosopher of law Hasso HOFMANN (1995) regards as the most important achievement of European legal culture since the end of the 18th century and whose influence reaches into the private sphere of every single citizen. Hermann KORTE must be agreed with when he talks of the state as a coordinating organ and its bureaucracy as its regulating organs, and states that the existence of very modern, that is, highly functionalized and differentiated societies is inconceivable without state and bureaucracy (1988, p156).

Tempering of Affects

While only the concept of 'sociogenesis' has been discussed up until now, that is, the 'collective' aspect of the civilizing process, the 'individual' dimension, the so-called 'psychogenesis', will be explained in the following. This aspect is important because it illustrates the creation and importance of some other controlling and coordinating mechanisms, that is, the 'constraints of the civilizing process'[23].

At this individual level, ELIAS is concerned about a more precise definition of the role of the individual in the course of social processes. The idea that each era needed people with a personality structure corresponding to the spirit of that era is a necessary consequence of ELIAS' conviction that society and the individual are not entities existing separately from each other. It is Norbert ELIAS' special merit that he has defined more precisely this socially based psychological component of the civilizing process and has at the same time made important parts of Sigmund FREUD's psychoanalysis[24] accessible to sociologists[25].

Courtization of Warriors

In order to illustrate what ELIAS means by 'psychogenesis', the Middle Ages must be looked at again in more detail: as already mentioned, a social change in which the knights successively declined in importance took place in the Middle Ages. At the same time, a courtly-aristocratic upper class was established, taking the place of the knights and achieving the status of a social elite. The central figure in each of these courts was an influential, wealthy feudal ruler, who, thanks to his financial means, could afford to 'hold court', thus uniting a large number of persons within a relatively small area.

It was highly important for the large number of noblemen who had lost their possessions due to overpopulation and impoverishment to be present at such courts. This made adaptation to the customs of the court highly opportune. This gave rise to the so-called 'courtoisie', whose most general characteristic was that a code of behaviour was laid down essentially containing standards which imposed on the socially dependent members restraint towards persons of higher rank (Vol.II, p86).

This constraint to show restraint was also imposed on the courtier vis-à-vis his peers, for in this courtly environment—in which a large number of chains of action crossed each other, thus creating a maximum of social figuration—it was inevitable that the individual had to calculate the consequences of his actions and, if necessary, had to show greater restraint than was necessary in the simple world of the lowly knight (Vol.II, pp258).

Furthermore, life at court differed from that of the lowly but independent knight in one important point: in the life of the knight, the omnipresence of war also left its impression on everyday life, resulting in a way of life which by today's standards was very coarse and which was particularly characterized by a high propensity to aggression on the part of the knight (Vol.II, pp78)[26].

The situation was quite different at court: the monopoly of power exercised by the central ruler had a 'pacifying' effect on people and had consequences for the individual in two respects: the immediateness of great danger was absent, which assured a high level of personal safety, but on the other hand the courtier was subject to restrictions in dealing with his peers, since he could not resort to physical force to defend his interests. This meant that even at this level of interaction, restraint and self-restriction became an important behavioural feature.

Refinement of Customs

A form of 'automatism' was also initiated here: The 'courtois' habits were neither creations of rational planning by a ruler nor were they products of pure chance.

They also were a consequence of those inner constraints, the so-called 'immanent regularities', which are so typical of Elias' conception and which result from the fact that the formerly free knights, who had become courtiers at the end of the Middle Ages, were subject to two important influences: as well as the constraint to show restraint and moderation, these persons as members of an exposed group were, however, also exposed to the promotion ambitions of the members of lower social groups. They were forced to resist these ambitions if they did not want to lose the privileges they had already achieved at court. This pincer-like pressure gave rise to a code of behaviour which was passed on among courtiers, in some cases under the seal of secrecy, and which was preserved because it ensured that a social distance remained between them and the members of lower social groups.

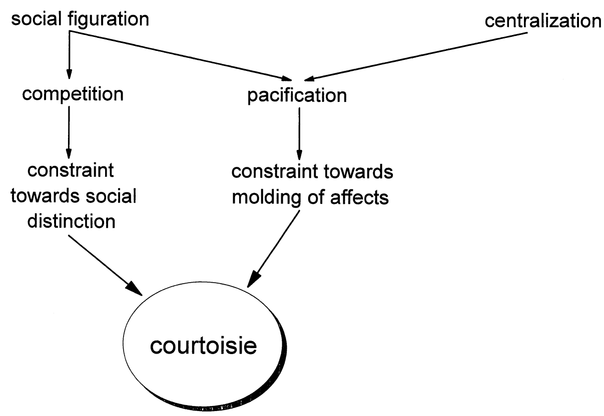

Figure 6: Tempering of Affects and Courtly Behaviour

The consequence of the increasing differentiation of society, on the one hand, and the attractiveness of courtly life on the other hand, was that these 'courtly forms of conduct' (Vol.II, pp66) also affected the members of the middle classes and later even became accepted in lower social groups. To ensure that 'courtly forms of conduct' did not lose their significance as a means of social distinction, and to preserve the original function of the courtoisie, customs and manners were permanently refined, modified and further developed. However, they were adapted just as quickly by the aspiring groups, thus making them useless as means of distinction as understood by their original creators.

This process is the motor for the never-ending change in habits in all levels of society. It finally ends where the contrasts in society become blurred as these classes become dependent on each other on account of the division of functions created within the framework of progressing social figuration (Vol.II, pp251). This condition, in the direction of which a first step was made with the collapse of the courtois society during the French Revolution, was finally achieved in industrial society. Here, where all social groups are closely intertwined and dependent on each other, it has become impossible to maintain privileges on the basis of social class.

However, there is no uniformity of customs and behaviour of all those involved, but modifications in behaviour can now develop freely within the limits of the general process of civilization (Vol.II, p297).

One reason for this is the decrease in external threat. This has decreased because the central power, which has been meanwhile become firmly established, has created 'pacified' areas by exercising its monopoly of power. Another reason is that after the rise of the bourgeoisie, there was no competing class left and thus no reason to distinguish oneself from parvenus by observing strict standards of behaviour. It is these changes in the general conditions that have made it possible for behaviour to free itself from strict patterns and to become more variable.

This variability, however, was not unlimited: just as the division of labour in society cast its shadow on the habits of the courtly aristocracy, this constraint to show more restraint and greater tempering of affects was not eliminated in a society shaped by the values of the working classes. The bourgeois professional and commercial functions took an even heavier toll from all those involved, and this toll was paid by the individual in the form of even greater control of his instinctual life (Vol.II, p307). This becomes particularly clear in the strict taboos relating to money and sexuality in the middle classes and which became generally applicable with the rise of the bourgeoisie. Furthermore, the containment of physical violence by no means meant that a repression-free state had been reached; instead it was replaced by economic violence and constraints and made 'tempering of affects' by the individual necessary (Vol.II, p235).

In biological terms, evolution has reached a new, higher level, which must be regarded as a 'typogenesis'. From this level, new variations can emerge in the form of 'adaptive radiation', which affect only minor characteristics, while the highly differentiated basic properties of large taxonomic groups remain unaffected.

The Role of the Family

Standards of behaviour and taboos therefore prevailed primarily through social constraint, which was apparent in that it was only possible for a person to rise to or remain in the upper class if this person submitted to the customary habits there. This external constraint was increasingly internalized by the individual and was transferred into the form of self-constraint acquired by self-conditioning. The consequence of this was that deviations from standards led to feelings of repugnance in the observer. The catalogue of standards was also internalized with respect to one's own behaviour: a deviant individual thus feels 'shame', which is in turn an expression of a deep psychological fear of being caught and the subsequent gestures of superiority by others, as well as social degradation by the group (Vol.II, pp292).

The adoption of these standards by the members of the new middle classes became the content of education for the children of these classes, and compliance with them was ensured by punishment or reward.

Through the rise of the bourgeoisie in the 19th century and the subsequent creation of a new upper class, the family increasingly became the place where this conditioning was primarily carried out (Vol.I, pp187). In sociological terms, this meant that the family obtained the status of a social institution whose function (among other things) consisted in the systematic moulding of the social behaviour of the child. This function is still applicable even today (cf. WURZBACHER/KIPP, 1968). In this connection, WURZBACHER and KIPP talk of the family as a very useful phenomenon (1968, p54) because of its manifold, mediating functions between the individual and society (1968, p48).

The family achieved this status because it met these requirements by society in two respects: firstly, the family was a place where spontaneous but undesirable impulsive acts could very effectively be suppressed by means of punishment, and secondly the children that were raised in the heart of the family and protected from the outside world were confronted very early with the customs of the bourgeoisie and thus obtained knowledge of highly civilized living conditions. The consequence of this was that the child—unaware of any other conditions—accepted the practised customs as natural, and adopted them without further contradiction. This function of the family was perfected as the public and private sectors of daily life were separated from each other and thus made the family the leading institution of the private sector (cf. ARIÈS/DUBY, 1991, p18).

The children brought up in this way required neither social respect nor pressure from their parents in their later lives to ensure compliance with social standards. On the contrary, this control took place in the form of 'self-restraint' by the individual, and the feelings of aversion, repugnance and shame thus generated are the basis of what every generation looks upon as 'natural' and 'self-evident'.

ELIAS' message is that the change in behaviour of the individual is thus realised through vertical constraints: first there is a constraint by the social elite vis-à-vis the members of the middle classes and then a constraint by the parents vis-à-vis their children. In both cases, the consequence is restraint, tempering of affects and drive abstention by the individual. An adolescent child thus goes through the same process as the middle classes have gone through, later in time and at a slower speed.

It must therefore be noted that in this historical situation of the rise of the bourgeoisie there was an urgent need for an institution that organized the drive abstention of the human being systematically and effectively. It is significant that no 'new' institution was developed in order to guarantee this social function, but that this function was transferred to the family, and thus to an already existing institution, but which, with respect to its traditional functions, was in a state that the biologist would describe as 'rudimentation'[27]: Jakobus WÖSSNER makes this situation clear when he states that since the Middle Ages the family has lost one function after the other: the family only achieved its character as an important social institution and the related present-day significance on account of the change in its function and its important role in the training of the social-cultural personality, according to this sociologist and social philosopher (1968, p312). The well-known family sociologist Rene KÖNIG also emphasises this change in the function of the modern family: in this connection he talks of a "disintegration of the modern family", which represents an "inevitable and irreversible" consequence (1946, p68) of the differential specialization processes taking place in the individual social sectors and which has led to a "functional reduction to purely family-related achievements" taking place with respect to the family (1946, pp68).

This process of the change in function of a social institution—as can be seen in the family in a particularly impressive form—has obvious parallels with the biological phenomenon of the 'change in function' which has only recently been discovered by Gerhard VOLLMER as an evolutionary factor overlooked up until now (1984, pp19): VOLLMER refers to an abundance of organs that have lost their original function and which today serve a completely different purpose for the organism compared to the original function. In this connection he mentions the ossicles of mammals (originally parts of the temporomandibular joint), the air bladder in fish (originally a primitive lung) or the teeth of vertebrates (former placoid scales). VOLLMER interprets these changes in function as follows: these organs originally possessed their 'later' important function only latently, in the form of a hardly noticeable 'double function'. As the organ served its 'old' function satisfactorily, this 'new' function was able to emerge slowly, it was evident as a special function after it had reached a certain stage of "functional maturity" (p25), and the use of this function was advantageous on account of changes in the environment.

The change in the function of the family into an institution of behaviour moulding for adolescents can be interpreted by analogy with this: this was originally a latent function of the family, which existed among numerous other manifest functions of the family, and matured as these manifest functions[28] became less important. Through this release of the family from its original tasks due to changes in the social environment, this new function became more important and gave a new function to the family as an institution.

This special feature of the process of institutionalization also expresses the legitimacy of these statements concerning the identification of social institutions as the 'constraints' of the civilizing process as essential conditions for the successful advance of all evolutionary processes. Institutions, once established, are not simply eliminated due to a change in the social environment, which made them useless, but in such a situation they undergo a change in meaning and function which formally preserves this institution but gives it a new status in the network of institutions.

Summing up, and taking account of the detailed documentation concerning the opening-up of the monopoly, institutions (like the 'constraints' of biological evolution) are not simply dissolved or eliminated when the general conditions change[29]. On the contrary, within the framework of the civilizing process they undergo a change in their validity, which can consist in either a change in occupancy with respect to control over an institution, or a change in function, which is characterized by an institution obtaining new functions and losing old ones.

Drive Abstention as a Characteristic of Civilization

As the social behaviour of the adolescent is formed to a decisive extent within the family, the importance of this social institution in relation to the psyche of the individual will now be considered in greater detail: As already mentioned, the drive abstention practised by the family and based on family conditioning is considerably more efficient compared with the historically older social conditioning by the upper classes. The reason for this is that the constraint internalized by the individual disappears from the consciousness of the individual and is thus no longer recognised as such. The resulting characteristic of 'naturalness', which is part of all customs and habits, is also the reason why children who resist such measures of affect moulding and do not 'adapt' are very quickly labelled as 'abnormal', 'perverse', or even 'sick' in a psychosomatic sense, and so are excluded from contact with children of the same age who comply with these standards. In other words, they exclude themselves from such contact.

These conditioning processes, however, are not without consequences for the child and thus for all members of society. To make this understandable, ELIAS falls back on basic statements of psychoanalysis by Sigmund FREUD: according to this theory, one consequence of family conditioning is that impulsive actions, which are expressions of the "Triebseele", are no longer controlled by the "ego" in the form of a conscious suppression, but that this function is taken over by the "super-ego" and is thus no longer subjected to a conscious assessment. Such a function of the "super-ego" consists in steering the aggressive energy (which is still present) from the outside to the inside, and thus against the "ego" (FREUD,1923, pp320). There it is regarded as "a sense of guilt" (1923, p304). The more aggression is reduced between individuals in the course of the civilizing process, the more it is bottled up within the individual. Social constraints, which are pushed into the subconscious by the "super-ego", where, however, they do not stay, are thus internalized. On the contrary, the suppressed tendencies and drives now rise up into the consciousness as the legitimation of libidinous wishes and give themselves a shape which is permitted by the "super-ego". FREUD describes this phenomenon as "rationalization" and points out that this is where the basis for sociogenic neuroses are to be found.

Furthermore, this self-constraint exercised by the "super-ego" results in fears in the subconscious, which, however, are no longer directed only against physical violence and existential threats, as was the case in the Middle Ages, when there was no real central power. On the contrary, there are now so-called 'inner fears', which are directed against one's own behaviour as well as against the behaviour of other people. In places where the human being violates standards through his own incorrect behaviour—and now fears disapproval by others, or even social degradation—these fears emerge in the form of the 'feelings of shame' already mentioned. Deviant behaviour by others evoke fears inside the observer which represent feelings of 'repugnance'. They can be given a deep psychological significance as fears relating to a challenging of standards and the related danger for one's own social status.

As the self-constraint of the individual is inseparably dependent on the behavioural patterns that are approved or disapproved in a society, and as these behavioural standards, on the other hand, are not natural but have been developed by society, this means that the inner fears which exist latently within the individual are not an expression of real, existing dangers but are to be regarded as a reflection of the particular state of the society itself (Vol.II, pp327).

The resulting fears, which necessarily must be described as 'sociogenic', are however only encountered from a certain stage of civilization onwards, namely whenever the physical threat of violence or hunger as dangers have been pushed into the background due to the increase in the division of labour in society and the existence of a power monopoly. Only then can the process begin which ELIAS describes as "rationalization" (Vol.II, pp276) and which, according to him, consists in "ego" and "id" being distinguished from each other to such an extent that a powerful "super-ego" can become established and can take over the control of behaviour.[30]

The final result of this conditioning based on self-restraint is thus a "division of the 'ego'" (Vol.I, p190), which ELIAS regards as being typical of people in a civilized society. There are pleasurable expressions of drive on the one hand, which, however, are held in check by prohibitions promising displeasure through shame and repugnance on the other hand. In this situation, a feeling of "discomfort in culture" (FREUD 1930) evolves; this is an indication that the price for cultural progress is paid by loss of happiness due to increasing feelings of guilt (1930, p260).

Situations of social upheaval show that these drive-related, pleasurable tendencies cannot be eliminated, as is frequently claimed in behaviourist theories (cf. H. ROTH, 1966) (Vol.I, p192). Affects which are thought to have been lost are released from self-constraint (which is a product of social constraint) and appear with incredible vehemence. These are not only coarse habits which can be seen. On the contrary, daily life is characterized by the threat or use of direct physical force, as is shown today in the course of the civil wars of the post-communist era in eastern Europe.

But even in 'quiet' times, these affects are not completely absent. People's preference for spectacles of an aggressive nature are an indication of the persistence of these affects. The civilizing process only leads to the indirect—visual—participation in an affect-related situation being regarded as pleasurable and thus as satisfying drives.

The 'Constraints' of the Civilizing Process

Summing up, the situation is thus as follows: besides the control at collective level by social institutions, which has already been explained above and is of increasing complexity, there is a further control system which is effective at the individual level of the civilizing process and which, together with the first control system secures the achievements of civilization: these are the authorities mentioned above which assure the affect control of the individual and which, in the course of their improvement, increasingly withdraw themselves from the awareness of the individual. Thus the central state power, which gained in importance during the Middle Ages and which forced compliance with certain behavioural standards, initially by physical force and later by more subtle methods, can be regarded as a more original form of controlling authority. In the course of the 'courtization of warriors', the consciousness of the individual became more important and the controlling authority was thus shifted from the social environment into the interior of the individual. As a consequence of this, the 'super-ego' dominates as a controlling authority as the external dangers and the existential threats lose significance through the civilizing process, and it controls the actions of the individual by giving him a feeling of what is 'natural'.

The 'constraints' which are subject to social change are thus largely identified: the civilizing process takes place between the two cornerstones of collective and individual control and is removed to a very great extent from the influence of the individual. While 'social institutions' carry out their functions like machines, as impersonal authorities of the constitutional state, and are not subject to direct intervention by the individual, such intervention is practically impossible with respect to the authorities of individual affect control. The reason for this is to be found in the particular nature of the 'super-ego': it controls the behaviour of the individual with the help of derivatives of fear—namely the feelings of shame and repugnance—and thus is no longer consciously registered by the individual.

Pacification

What role do these fields of 'psychogenesis' and 'sociogenesis', which so far have largely been considered as being isolated from one another, play together in the civilizing process?

The following conclusion shows that both fields are only different from an analytical point of view, but in fact cannot be separated from each other: the centralization of power could not have taken place without individuals whose affects are moulded to such an extent that they can live in larger communities. On the other hand, affect moulding of individuals would have been impossible if there had not been a central authority which established its monopoly of power with the help of strong social institutions.

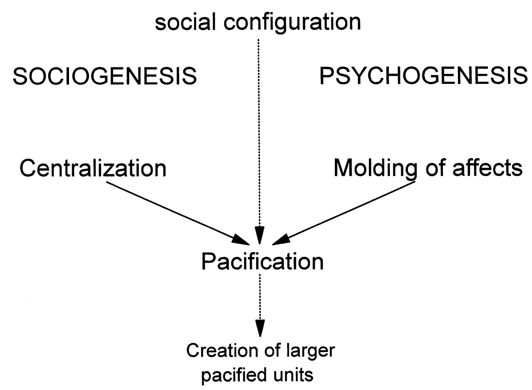

Figure 7: Pacification

Affect moulding and centralization can be regarded as the basis for the creation of pacified areas. Due to the unbroken trend towards the progressive scarcity of space and the related increase in 'social figuration', these led to larger and larger territories which were 'pacified' in their interiors due to the above-mentioned mechanisms.[31]

This makes the current references of ELIAS' theory of the civilizing process become clear: ELIAS believes that states are only temporary phenomena of such pacified areas and are progressively replaced by larger, supranational formations. When ELIAS states that he detects "the first outlines of a worldwide system of tensions composed by alliances and supra-state units of various kinds, the prelude of struggles embracing the whole globe, which are the precondition for a worldwide monopoly of physical force, for a single

Author's address

Detlef Weinich, Otto Stein-Straße, D–97076 Würzburg, Germany.

central political institution and thus for the pacification of the earth" (Vol.II, p332) this shows the whole dimension of his theory. Although these hopes must be judged with some scepticism in view of the increase in regional conflicts and the problems which the UNO has with them, ELIAS must be regarded as a visionary, for he formulated his ideas at a time in which terms like 'worldwide pacification' and 'global organizations' still seemed to be pure fabrication.

Summary

I should like to conclude my comments on Norbert ELIAS by stressing his message that social phenomena are never the result of planning by individuals or human collectives, but that on the contrary they developed on the basis of those 'immanent regularities' in the civilizing process, frequently against the express wish of the rulers or the ruling classes.

With the rejection of the (behaviourist) view that the process of civilization is something that can be instructed 'from the outside' and the emphasis on the 'inner' conditions it is subject to with respect to its direction and its 'objectives', it should also be clear that statements like the one made by the American sociobiologist E.O. WILSON, (who claimed that "contemporary human social behaviour ... comprise hypertrophic outgrows of simpler features of human nature joined together into an irregular mosaic", 1978, p94), do not take sufficient account of the complexity of the material. They ignore the fact that the civilizing process, like all evolutionary processes, has a 'creative element', and that they are an expression of what Julian HUXLEY so aptly characterised as a 'nothing-else-buttery' (cf. POPPER/LORENZ 1985, p37). The consequence of such a crude method is that specific phenomena of higher levels (of being) can only be inadequately perceived, and can seldom be satisfactorily explained.

I should like to sum up by saying that *biological* aspects of evolution theory can only be applied to higher-order phenomena (social or cultural evolution) if they take sufficient account of the process of evolution as a "process of creative development" (POLANYI 1985, p49). This conclusion, however, is also a clear rejection of the DARWINIAN paradigms about selection by the environment (which on closer inspection represents a variation of behaviourism transferred to phylogenetic processes) and makes it clear that evolutionary concepts taken from system theory or constructivism have to be applied here. Only with the help of such concepts—in which there is room for terms like 'constraints', 'emergence', 'self-organization' or 'autopoiesis' and as have been introduced by Rupert RIEDL (1976), Erich JANTSCH (1979), Humberto MATURANA (1982) or Wolfgang GUTMANN (1989)—is it possible to approach sociological or cultural phenomena by means of *biological* categories.

Notes

1 For the importance of ELIAS and his classification as a 'classic of sociology' cf. KORTE (1988, p33), KUZMICS/MÖRTH (1991, p7), FRÖHLICH (1991, p95), LEPENIES (1977, p29), BOGNER (1989, p13).

2 It is important and is to be regarded as an advantage compared with other theories that Norbert ELIAS gives certain historic examples and thus makes his explanations more clear. This makes a big difference to the explanations of those scholars who talk about metaphysical laws in cyclic processes. One of these researchers is Oswald SPENGLER (1922), who talks about the 'natural ageing' of civilizations.

3 Although this book made ELIAS known to a larger group of people, he refers to his own investigations, which he first published in his postdoctoral thesis required for qualification as a university lecturer. The title is: "Die höfische Gesellschaft. Ein Beitrag zur Soziologie des Hofes, der höfischen Gesellschaft und des absoluten Königtums". A new edition of this was published in 1969.

4 It appeared as a pre-print by the publishing house C. Schulze & Co.GmbH in Gräfenhainichen. The very first publication appeared at 'Haus zum Falken', a publishing house in Basle (Switzerland) which published books which were not allowed to be published in Germany at that time.

5 The following references to ELIAS will only mention the number of the volume and the page.

6 The difference between "Zivilisation" and "Kultur" and the more technical touch of "Zivilisation" is a peculiarity of the German language. In France and England the word 'civilization' corresponds in its meaning to the German word 'Kultur'.

7 c.f. Hams Peter DREITZEL (1967, p17) and Horst REIMANN (1979, p88)

8 This applies to the theories of Henri SAINT SIMON(1760-1825) and Auguste COMTE (1798-1857) and their famous 'rule of the three stages': It speaks about a continuity of a theological, a metaphysic and a scientific era.

9 cf. GUTMANN (1989) and his approach to the organism as a 'hydraulic system'. This view shows the constriction of this room for further development very clearly.

10 LORENZ spoke about 'adaptation' as an 'essential cognitive process' (1983, p58).

11 This view of the mutual conditionality of individual and collective phenomena is approved by numerous authors and makes the theory of the civilizing process particularly compatible with the basic theory of human ethology. This is expressed in the statement by Konrad LORENZ, according to which the human spirit has to be regarded as a "social effect" (1983, p69). From this point of view, the individual is not a real human being. LORENZ goes on to say that culture is the individual concrete realization of common spiritual interest (1983, p70).

12 This means a turning away from the traditional sociological formation of concepts. Hermann KORTE looks upon it as the most outstanding characteristic of ELIAS that he makes no difference between individual and society. He also notes a break with the sociology of Max WEBER: This appears very clearly with respect to the 'ideal types' of Max WEBER, which are abstracts of a possible reality. They are imaginary constructions invented by sociologists who felt forced to invent them in order to put things in order in an ambiguous environment (KORTE 1988, p162). Instead of this procedure, ELIAS does not create any 'sociological basic terms to tidy up the supposed untidiness of social conditions; but he formulates basic problems of the people and of the society which these people build together' (1988, p164).
This view includes also a rejection of ideas like those which are formed by the authors of the 'Structural-Functional Theory'. Talcott PARSONS (one of the most important representatives of this theory) looks at society as a solid phenomenon and he postulates the existence of abstract, eternal valid characteristics of society. From this point of view, people are only the holders of established posts, and exist in order to make social principles come true (cf. Vol. I, pp225). ELIAS also rejects individual-psychological ideas which regard the individual as "homo clausus", and postulates the existence of human constants which exist independent of the world outside (Vol.I, pp252).

13 ELIAS talks about the development of the population as an important factor which should never be neglected, but he rejects the idea that the development of the population is a "first cause" (Vol.II, p32).

14 cf. in detail BOGNER (1991)

15 cf. the idea of Marxism. It says that class differences are decisive and history is a sequence of class struggles (MARX/ENGELS 1974, p462).

16 Although he often declares that history is a dynamic process where every phenomenon is a result of preceding events, and that there is therefore no zero point in history, many of his explanations start in the Middle Ages.

17 For a detailed description of this aspect of ELIAS see: BOGNER (1989).

18 An increase in dangerous diseases (followed by death) has also to be considered in this context. The social historian Arthur E. IMHOF showed that there was a reappearance of forgotten epidemics in times of post-war destruction—like typhus fever which was transmitted by lice (IMHOF 1988, p56: with a study on the different life expectancy of ethnic groups and people with differing degrees of civilization, pp79)

19 For a detailed view on these tendencies (especially on the 'renaturalization' of man in post-modern society) see WEINICH (1995).

20 Such a situation appeared in late antiquity, when there was a decrease in population (due to the restlessness caused by the migration of peoples) which made the markets collapse and led to a decline in long-distance trade and the money economy (Vol.II, p34). This decline in the money economy meant that an important criterion for the division of labour was lost and thus the whole character of society was changed by the increase in importance of the barter economy. The urban sector of society becomes smaller, for products were only produced where they were needed. The chains through which the products were traded became shorter and shorter, and the level of interdependence in society even fell until the typical self-sufficiency of individual estates and villages in the feudal age was achieved.

21 The term 'cultural evolution' is used here with the same meaning as famous biologists use it. They regard 'cultural evolution' as a phenomenon which includes all levels of being which lie beyond biological evolution and in which information is not transferred genetically (cf. BONNER 1983).

22 In this assessment of the DARWINIAN point of view, I refer to the statements by the well-known German evolutionist Günther OSCHE: he said that natural selection is the only directing force in evolution, and that it brings a direction to the raw material which is provided by mutation and recombination (OSCHE 1972, p43).

23 ELIAS shows the changes in individual behaviour since the Middle Ages through a analysis of books on manners and behaviour. The paper "De civilitate morum puerilium" written by Erasmus von Rotterdam in the year 1530 is very important in this context.

24 For a detailed view about the importance of FREUD for ELIAS cf.: GOUDSBLOM, 1984b.

25 In 1933, within the framework of his critical analysis of Marxism, FREUD formulated the thesis that in the course of the process of civilization there is a change in the drives and emotions of individuals, as well as the development of economic necessities. He also stated that it should be one of the main goals of social research to discover these fundamentals (FREUD, 1933, p605-6).

26 Here, the question must be asked whether ELIAS' view of Middle Age man as a blood-thirsty creature is not somewhat exaggerated. Irenäus EIBL-EIBELSFELD, in the course of his investigations into 'primitive peoples' was able to show that there are a lot of aggression-inhibiting mechanisms which are at least partly the biological heritage of mankind. Furthermore, even at very 'low' cultural levels, there are mechanisms for avoiding conflicts and curbing aggression (EIBL-EIBELSFELDT 1975).

27 This biological specialist term means, that an organ has lost its initial function and is in the process of regression (cf. SCHEMMEL 1984).

28 The terms of 'manifest' and 'latent function' were introduced to sociological terminology by R. K. MERTON. They are used to distinguish between planned and non-planned functions in social structures (1949, p49).

29 Günther OSCHE characterises this situation biologically with an apt aphorism by stating that "evolution ... cannot hang up a notice on the wall saying 'Closed for renovation'" (OSCHE 1966).

30 In this point, the views of ELIAS converge in a remarkable way with the principles of another important sociological theory. These are the notions of Theodor ADORNO and Max HORKHEIMER, two representatives of the "Frankfurter Schule". In their "Dialektik der Aufklärung" (1947) they also stress the importance of the creation of an 'internal' controlling authority" (cf. also BOGNER 1989, p70) in the course of the process of civilization. This organ, which controls the "inner" as well as the "outer" world is considered by the two authors as the "ego", which modifies the drives, on the one hand, and learns how to instrumentalize the regulated nature of the outer world on the other hand.

31 'Pacification' and 'freedom from violence' also play an important part in the theories of Max WEBER. Although it is impossible to introduce the very comprehensive work of the 'classic of sociology' adequately, it must noted that WEBER also assigns these two factors an important role in the process of 'occidental rationalization': 'the formal rationality of trade'—which is an important element in this process—is highly dependent on the long-term establishment of a monopoly of power by the state, thus eliminating violence as a means in economic competition (cf. 1972, p32, also: p519). Arthur BOGNER (1989) systematically compared ELIAS and WEBER, and sees important parallels between the concept of 'rationalization' and the 'theory of the civilizing process'. He believes that both theories are useful for mutually correcting the bottlenecks and gaps in their approaches (1989, p193).

References

Ariès, P./Duby, R. (ed) (1991) Geschichte des privaten Lebens, vol. 3: Von der Renaissance zur Aufklärung. In: Ariès, P./ Chartier, R. (eds.) Geschichte des privaten Lebens. S. Fischer, Frankfurt.

Bogner, A. (1989) Zivilisation und Rationalisierung. Die Zivilisationstheorien Max Webers, Norbert Elias und der Frankfurter Schule. Westdeutscher Verlag, Opladen.

Bogner, A. (1991) Die Theorie des Zivilisationsprozesses als Modernisierungstheorie. In: Kuzmics, H./Mörth, I. (eds) (1991) pp33–58.

Bonner, J.T. (1983) Kulturevolution bei Tieren. Parey, Berlin-Hamburg.

Briggs, J./Peat, F.D. (1990) Die Entdeckung des Chaos. Eine Reise durch die Chaos-Theorie. Hanser, München.

Comte, A. (1956) Rede über den Geist des Positivismus. Felix Meiner, Hamburg.

Dreitzel, H.P. (ed) (1967) Sozialer Wandel. Zivilisation und Fortschritt als Kategorien der soziologischen Theorie. Luchterhand, Neuwied und Berlin.

Eibl-Eibesfeldt, I. (1975, [2]1984) Krieg und Frieden aus der Sicht der Verhaltensforschung. Piper, München.

Eigen, M./WInler, R. (1975) Das Spiel. Naturgesetze steuern den Zufall. Piper, München-Zürich.

Elias, N. (1939) Über den Prozeß der Zivilisation. Soziogenetische und psychogenetische Untersuchungen, Vol. 1 ([1]1976) Wandlungen des Verhaltens in den weltlichen Oberschichten des Abendlandes, Vol. 2 ([15]1990) Wandlungen der Gesellschaft. Entwurf zu einer Theorie der Zivilisation. Suhrkamp, Frankfurt.
English translation: The Civilizing Process. Sociogenetic and Psychogenetic Investigations, Vol. 1: The History of Manners (1978), Vol. 2: State Formation and Civilization (1982). Blackwell, Oxford.

Elias, N. (1969) Die höfische Gesellschaft. Untersuchungen zur Soziologie des Königtums und der höfischen Aristokratie. With an Introduction: Soziologie und Geschichtswissenschaft; Luchterhand, Neuwied-Berlin. Suhrkamp, Frankfurt.

Elias, N./Lepenies, W., (eds) (1977) Zwei Reden anläßlich der Verleihung des Theodor W. Adorno-Preises. Suhrkamp, Frankfurt.

Freud, S. (1969) Studienausgabe, 10 vols. ed. by Mitscherlich, A./A. Richards/J. Strachey. Fischer, Frankfurt.

Freud, S. (1923) Das Ich und das Es. In: vol. 3 of the Studienausgabe 1969ff, pp273–330.

Freud, S. (1930) Das Unbehagen in der Kultur. In: vol. 9 of the Studienausgabe 1969ff, pp191–286.

Freud, S. (1933) Neue Folge der Vorlesungen zur Einführung in die Psychoanalyse. In: vol.1 of the Studienausgabe 1969ff. pp448–608.

Frohlich, G. (1991) Inseln zuverlässigen Wissens im Ozean menschlichen Nichtwissens. Zur Theorie der Wissenschaften bei Norbert Elias. In: Kuzmics, H./Mörth, I. (eds) (1991) pp95–111.

Gehlen, A. (1956; [5]1986) Urmensch und Spätkultur. Philosophische Ergebnisse und Aussagen. Aula, Wiesbaden.

Gleichmann, P./Goudsblom, J./Korte, H. (eds) (1984) Macht und Zivilisation. Materialien zu Norbert Elias' Zivilisationstheorie 2. Suhrkamp, Frankfurt.

Goudsblom, J. (1984) Die Erforschung von Zivilisationsprozessen. In: Gleichmann, P./Goudsblom, J./Korte, H. (eds.) (1984) pp83–104.

Goudsblom, J. (1984b) Zum Hintergrund der Zivilisationstheorie von Norbert Elias: Das Verhältnis zu Huizinga, Weber und Freud. In: Gleichmann, P./Goudsblom, J./Korte, H. (eds.) (1984) pp129–47.

Gutmann, W. F. (1989) Die Evolution hydraulischer Konstruktionen. Organismische Wandlung statt altdarwinistischer Anpassung. Kramer, Frankfurt.

Hartmann, N. (1940, [3]1994) Der Aufbau der realen Welt. Grundriß der allgemeinen Kategorienlehre. deGruyter, Berlin.

Hofmann, H. (1995) Geschichtlichkeit und Universalitätsanspruch des Rechtsstaates. Der Staat. Zeitschrift für Staatslehre, öffentliches Recht und Verfassungsgeschichte 34: 1–32.

Horkheimer, M./Adorno, T.W. (1947) Dialektik der Aufklärung. Philosophische Fragmente. Querido, Amsterdam.

Imhof, A.E. (1988) Reife des Lebens. Gedanken eines Historikers zum längeren Dasein. Beck, München.

Jantsch, E. (1979, [4]1988) Die Selbstorganisation des Universums. Vom Urknall zum menschlichen Geist. Carl Hanser, München.

König, R., (1946, [2]1974) Materialien zur Soziologie der Familie. Kiepenheuer & Witsch, Köln.

Korte, H. (1988) Über Norbert Elias. Das Werden eines Menschenwissenschaftlers. Suhrkamp, Frankfurt.

Kuzmics, H./Mörth, I. (eds) (1991) Der unendliche Prozeß der Zivilisation. Zur Kultursoziologie der Moderne nach Norbert *ELIAS*. Campus, Frankfurt-New York.

Lepenies, W. (1977) Ein Außenseiter, voll unbefangener Einsicht. In: Elias, N./Lepenies, W., (eds.). pp9–34.

Lipp, W. (1989) Entstitutionalisierung. Wie erfaßt man sozialen Verfall? In: Papalekas, CH. (ed) Kulturelle Integration und Kulturkonflikt in der technischen Zivilisation. Campus, Frankfurt, pp86–114.

Lorenz, K. (1983) Der Abbau des Menschlichen. Piper, München-Zürich.

Marx, K./Engels, F. (1974) Manifest der Kommunistischen Partei, MEW 4. Dietz, Berlin.

Maynard-Smith, J. (1985) Developmental constraints and evolution. The quarterly review of Biology 60: 265–287.

Maturana, H. R. (1982) Erkennen: Die Organisation und Verkörperung der Wirklichkeit. Vieweg, Braunschweig-Wiesbaden.

Merton, R. K. (1949, 1995) Soziologische Theorie und soziale Struktur. de Gruyter, Berlin.

Meyer, P. (1977) Kriegs- und Militärsoziologie. Goldmann, München.

Nagl, W. (1993) Grenzen unseres Wissens am Beispiel der

Evolutionstheorie. Ethik und Sozialwissenschaften 4, Heft 1:3–16.

Osche, G. (1972, 1978) Evolution. Grundlagen—Erkenntnisse—Entwicklungen der Abstammungslehre. Herder, Freiburg.

Osche, G. (1978) Die Sonderstellung des Menschen in biologischer Sicht: Biologische und kulturelle Evolution. In: Siewing, R. (ed) (1978) pp379–95.

Parsons, T. (1951) The Social System. The Free Press, New York.

Polanyi, M. (1985) Implizites Wissen. Suhrkamp, Frankfurt.

Popper, K. (1965) Das Elend des Historizismus. Mohr, Tübingen.

Popper, K./Lorenz, K. (1985) Die Zukunft ist offen. Das Altenberger Gespräch. Piper, München.

Reimann, H./Giesen, B./Goetze, D./Kiefer, K./Meyer, P./Mühlfeld, C./Schmid, M., (1977, [2]1981) Basale Soziologie: Hauptprobleme. Westdeutscher Verlag, Opladen.

Riedl, R. (1975) Die Ordnung des Lebendigen. Systembedingungen der Evolution. Parey, Hamburg-Berlin.

Riedl, R. (1976, [4]1985) Die Strategie der Genesis. Piper, München-Zürich.

Riedl, R. (1982, [2]1984) Evolution und Erkenntnis. Antwort auf Fragen aus unserer Zeit. Piper, München.

Roth, H. (1952) Begaben und Begaben. Über das Problem der Umwelt in der Begabungsentfaltung. Die Sammlung. Zeitschrift für Kultur und Erziehung; 7.Jg 1952: 395–407.

Schelsky, H. (1952) Über die Stabilität von Institutionen, besonders Verfassungen. Jahrbuch für Sozialwissenschaft 3: 1–21.

Schemmel, C. (ed) (1984) Regressive Evolution und Phylogenese. In: Fortschritte in der zoologischen Systematik und Evolutionsforschung 3. Parey, Hamburg—Berlin.

Schrödinger, E. (1944) What is life? The physical aspect of the living cell. Cambridge Univ. Press, Cambridge.

Siewing, R. (ed) (1978) Evolution. Gustav Fischer, Stuttgart-New York.

Spengler, O. (1922) Der Untergang des Abendlandes. Umrisse einer Morphologie der Weltgeschichte, 2 Bde. Beck, München.

Vollmer, G. (1984) Die Unvollständigkeit der Evolutionstheorie. In: VOLLMER, G. (1988) Bd.2, pp1–38.

Vollmer, G. (1988) Was können wir wissen? Band 1: Die Natur der Erkenntnis; Band 2: Die Erkenntnis der Natur. Hirzel, Stuttgart.

Weber, M. ([5]1972) Wirtschaft und Gesellschaft: Grundriß der verstehenden Soziologie. Mohr (Siebeck), Tübingen.

Weinich, D. (1995) Zurück zur Natur. Fortschritt oder Niedergang? Studien zum Zivilisationsprozeß, besonders in interdisziplinärer Betrachtensweise. Universität Würzburg [Diss. masch.].

Wilson, E.O. (1978) On Human Nature. Harvard Univ. Press, Cambridge.

Wössner, J. (1968) Kirche-Familie-Sozialisation. In: WURZBACHER (1968) pp308–52.

Wurzbacher (ed) (1968) Die Familie als Sozialisationsfaktor. Der Mensch als soziales und personales Wesen, vol. 3. Enke, Stuttgart.

Wurzbacher, G./Kipp, H. (1968) Das Verhältnis von Familie und öffentlichem Raum. In: Wurzbacher (ed) (1968). pp1–62.

Werner Callebaut

Thomas Kuhn as an Evolutionary Naturalist

Abstract

Prima facie, KUHN was clearly a philosophical naturalist. In a far more limited sense, he was also an evolutionary epistemologist of sorts, but only at the level of scientific change. Contemporary naturalism, including its evolutionary variety, is distinguished from its main rivals, foundationalist epistemology and the epistemological relativism associated with the SSK ('Sociology of Scientific Knowledge Claims') account in the sociology of science. Some common denominator seems to be emerging despite the vertiginous pluralism that characterizes neonaturalism. KUHN's comparison of the conceptual transposition of his own alternate view of scientific advance to the DARWINIAN revolution is discussed, and his thesis that what matters most in this comparison is the replacement of teleological thinking by a causal account is identified as naturalistic rather than evolutionary. Several other features of his account of science are reviewed in the light of more genuinely evolutionary approaches. To conclude, it is suggested that the integration of insights from the study of biological and psychological development into evolutionary theory sheds new light on the question of the unsatisfactory nature of KUHN's stage model of scientific change.

Key words

Evolution, Evolutionary Epistemology, historicism,

Introduction

Thomas KUHN (1922–1996) is the single most influential author in science studies to date. Combined sales of the second American edition of *The Structure of Scientific Revolutions* (KUHN 1970a)[1] and its twenty or so translations must reach about a million. Undoubtedly, KUHN's impact is now waning. This becomes clear, for instance, if one compares the rather ritual references to KUHN in the recently issued *Handbook of Science and Technology Studies* (JASANOFF et al. 1995) to the more substantial way his work was used in the older 'yellow bible' of science studies, *Science, Technology and Society* (SPIEGEL-RÖSING/PRICE 1977). Actually, in his "Reflections on Receiving the John Desmond BERNAL Award", KUHN already complained: "If *Structure* is a contribution to the sociology of science"—the fastest-growing among the science studies disciplines—"then the core of that contribution has been missed or else denied by many of those who trace their own work to it." For his concerns, he insisted—in marked contrast to the latter-day apostles of our presumed 'post-epistemological era'—, had always "also and inseparably been cognitive or epistemic" (KUHN 1983, p28; see also LEFÈVRE 1995).

Nevertheless, it seems fair to regard KUHN, negatively, as the main challenger of the so-called 'received view' of science (HACKING 1981), and, positively, as the premier metascientist (cf. HOYNINGEN-HUENE 1992) because of his original and relatively fruitful way of combining philosophical, historical, and sociological approaches to the scientific enterprise.

In an obituary in *The New York Times*, his theory of the structure of scientific revolutions was called "a profoundly influential landmark of 20th-century intellectual history", which "punctured the widely held notion that scientific change was a strictly rational process" (VAN GELDER 1996, pB7). 'Rationalism' may be opposed to 'naturalism' in this respect. There can be no doubt that KUHN was a philosophical naturalist of sorts, although his work is not usually presented in this light, being linked inextricably to the historical school in science studies if not to *historicism*, which pictures the scientific enterprise in terms of paradigms or other "closed circles" (Peter MUNZ) and is usually taken to be at odds with naturalism (but see GIERE 1988 and STENGERS 1993).

"My view of science is fundamentally evolutionary"

It is usually taken for granted that KUHN's theory of science was concerned with one central topic, *scientific development* in the generic sense of *change over time* (HOYNINGEN-HUENE 1993, p3–7). In addition, KUHN was certainly instrumental in making the nature of scientific change one of the most important problems, if not the core problem, in current philosophy of science.[2] Whether KUHN was an evolutionary epistemologist in any specific sense of that expression is much more controversial.

But hasn't KUHN (1970b, 264) written himself that "it must … be clear that my view of scientific development is fundamentally evolutionary"? And isn't it true that he is sometimes presented, especially in the older literature, as favoring a sort of 'random mutation—natural selection' scheme to explain scientific change, being bracketed together with Sir Karl POPPER in this respect (e.g., BLACHOWICZ 1971, p178)?

Alternatives in science

As an aside, let me note that when rereading such literature one is actually struck by the fact that it is not so much the (pseudo)problem of 'random' or 'blind' versus 'guided' or 'intelligent' or 'foresighted' variation that occupied several of the early commentators of KUHN's work, but the very idea that there could be something *contingent* about scientific change at all.[3] (I say "pseudo problem" because, pace BLACHOWICZ [1971], random variation is not even a requirement in the neo-DARWINIAN biological exemplar that orients evolutionary epistemology [see, e.g., CALLEBAUT 1993, p378–380]; and Donald CAMPBELL's legendary insistence on the 'blindness' of variation and selective retention was always intended as a foundational or, if you prefer, skeptical point that clearly said nothing about the actual power—or limitations, for that matter—of human heuristics [e.g., CALLEBAUT 1993, ch. 7; CAMPBELL 1997].)

One feature of KUHN's account of science that met with fierce resistance in the older literature is that rather than picturing scientific change as somehow *deterministic*, as in some of the older, "logocentric" accounts (AXTELL 1993), KUHN's account left room for *alternative developments*. For instance, "ordinary inductionists and KUHNIANS have a curious feature in common", WATANABE (1975, p114) observed: "Given a field of experience, they both believe that there are many alternative theories possible to cover the same field of experience." But, he went on to object, "those who face the actual history of science with intellectual honesty discover that in reality the one theory which has been adopted by history is practically the only theory possible."

In the same vein, BLACHOWICZ (1971, p178) invoked systems theory as a general criticism of the evolutionary model "insofar as it demonstrates the necessity of supplementing [the] mechanism [of relatively random change] with the non-random influences exercised by the internal organization of a system on its own development." The irony of this line of criticism of the KUHNIAN account of science is, of course, that KUHNIAN paradigms may be seen as precisely the kind of vehicles that, in human organizations, *constrain* the system's pathways in accord with the bounded rationality of us humans (see, e.g., ANDERSEN 1991).

Revolutionary change

Back to our main question, Did KUHN articulate an evolutionary epistemology? It is well known that in speaking of scientific *revolutions* and invoking an analogy between dramatic scientific change and political upheavals (cf. WERTHEIM 1974), KUHN unequivocally distanced himself from the assumption of 'cumulative' development that guided much older historiography of science: "Scientific development depends in part on a process of non-incremental or revolutionary change. Some revolutions are large, like those associated with the names of COPERNICUS, NEWTON, or DARWIN, but most are much smaller, like the discovery of oxygen or the planet Uranus." (KUHN 1977, xvii)

The trouble is that current evolutionary accounts of scientific change, although not committed to incremental growth at all (similarity being a non-transitive relation, evolutionary novelty, *even if it is of the 'saltational' type*, can be taken into account here just like in the biological case: CALLEBAUT 1993, p330; CALLEBAUT/VAN MEER 1997; cf. MASLOW 1978), are often conflated with the older growth model that *is* committed in this way (cf. RICHARDS 1987, p561–563 and passim).[4]

To complicate matters even more, authors who contrast KUHN's insistence on the major importance of occasional scientific revolutions with a properly evolutionary approach (e.g., EFRON/FISCH 1991) fail to notice that POPPER is often regarded as an evolutionary epistemologist, although he strongly opposed gradually developing "normal science" and insisted that all good science must be "revolutionary" (POPPER 1970).

Already at this preliminary stage of our investigation, it turns out, then, that simple 'yes' or 'no' answers to the question, 'Was KUHN an evolutionary epistemologist?' are misleading.

I provisionally conclude that the case for viewing KUHN as a naturalized epistemologist and philosopher of science is more compelling than the case for his account of scientific change being a brand of evolutionary epistemology in the sense of the Altenberg group's 'second EE' or BRADIE's (1986) EET ("Evolutionary Epistemology of Theories").

Structure of the paper

The structure of the remainder of this paper is as follows. First, I review the prima facie evidence for regarding KUHN as a naturalist in matters philo-sophical.

To prepare the ground for a better look at his philosophical stance, I then attempt to pinpoint what singles out contemporary philosophical naturalism, including its evolutionary variety, from its main rivals, viz., on the one hand, the foundationalist program of traditional epistemology, and, on the other, the wholesale epistemological relativism associated with the SSK ('sociology of Scientific Knowledge Claims') account that has come to dominate the sociology of science, at least in Europe. Given the vertiginous pluralism that characterizes neonaturalism (which, in addition to epistemology and the philosophy of science, influences much of the current debate in moral philosophy and the philosophy of mind as well), this will prove no easy task; yet some common denominator seems to be emerging.

I go on by discussing the passage toward the end of *The Structure of Scientific Revolutions* in which KUHN compares the "conceptual transposition" that his own "alternate view of scientific advance" suggests to that implied by the DARWINIAN revolution in biology in the nineteenth century. I identify his thesis that what matters most in this comparison is the abolition of teleological thinking and its replacement by a causal account as properly naturalistic rather than evolutionary.

I then consider several other features of his account of science in the light of more genuinely evolutionary approaches.

By way of conclusion, I suggest that current attempts to integrate insights from the study of biological and psychological development into evolutionary theory may shed new light on the question of the unsatisfactory nature of KUHN's stage model of scientific change.

"A role for history"—KUHN as a philosophical naturalist

In asking rhetorically, "How could history of science fail to be a source of phenomena to which theories about knowledge may legitimately be asked to apply?", KUHN (1970a, 9) was clearly prompting a *naturalistic turn* in epistemology and in the multidisciplinary endeavor that has come to be called 'science studies' (GIERE 1985, 1988; CALLEBAUT 1993, p13, 41–43, 73; KITCHER 1993, p6, 87n, 220).[5] In the now famous first sentence of *The Structure of Scientific Revolutions*, he expressed the hope that "history, if viewed as a repository for more than anecdote or chronology, could produce a decisive transformation in the image of science by which we are now possessed" (KUHN 1970a, 1).

Ronald GIERE, for one, reads KUHN's claim as an exhortation to develop the philosophy of science as *testable theory*, the relevant data for which are to be secured by history and sociology (GIERE 1985, p331–333; 1988, p32–34; cf. HOOKER 1987, 1995). Let me remark as an aside that this interpretation suggests a division of labor in which the history and sociology of science become subservient to the philosophy of science.[6] This may not bother the philosopher GIERE too much, but it is certainly something historians and sociologists find difficult to swallow.[7] Getting rid of this sort of asymmetries would be one good reason to develop a truly *unified* study of science. To be fair, I should add that among current philosophers of science GIERE is a prominent advocate of a "unified theory of science" (see most notably GIERE 1984).

As KUHN saw clearly, the "decisive transformation in the image of science" he envisaged presupposed that "elementary logical or methodological distinctions, which would … be *prior* to the analysis of scientific knowledge", now become "integral parts of a traditional set of *substantive answers* to the very questions upon which they have been deployed" (KUHN 1970a, 9; italics mine; cf., in the same vein, LUHMANN 1990, 15–16). The "entire arsenal of dichotomies" that traditional epistemologists invoke to suggest that such a shift cannot be properly accomplished—'history is descriptive while epistemology is normative', 'the context of discovery must be clearly distinguished from the context of justification', etc.—did not impress KUHN. Nor did the implied circularity of this shift—the main reason why traditional epistemologists reject the naturalistic stratagem—seem to bother him at all; for he continued: "That circularity does not at all invalidate [standard philosophical distinctions]. But it does make them parts of a theory and, by doing so, subjects them to the same scrutiny regularly applied to theories in other fields. If they are to have more than pure abstraction as their content, then that content must be discovered by observing them in application to the data they are meant to elucidate." (KUHN 1970a, 9)

Replacing the aprioristic, analytical type of philosophy of the logical empiricists and critical rationalists by an account of science that, in being testable, is scientific itself, points to yet another tenet of naturalism: that the hallmark of a naturalized theory of science is its *reflexivity*. What this boils down to is that, in marked contrast to a transcendental theory, a naturalized theory of knowledge allows the results of inquiry to impinge on the conditions of knowledge (LUHMANN 1990, 15–16; CALLEBAUT 1993, 1997). As a result, a naturalistic theory can only be *foundational* in a dynamic sense (WIMSATT 1996).

At this juncture we are approaching the limits of KUHN's own peculiar brand of naturalism, which has little to say about the issue of reflexivity (cf. GIERE in CALLEBAUT 1993, p114–115). What is more, by insisting on what *distinguishes* the history of science from science itself (most notably in KUHN 1980), KUHN departs from naturalism's stress on the *continuity* of nature and human existence (CALLEBAUT 1995a, 1997; KHALIL 1996), which seems to me to make his position (qua naturalism) inconsequential in favor of a historicist emphasis on *verstehen* (cf. Thomas NICKLES in CALLEBAUT 1993, p453–454).

Science as a natural human endeavor

A naturalistic account of science views contemporary (techno)science as a human practice that has become an integral part of our society (LUHMANN 1990) and culture (PICKERING 1992)—so integral indeed that it may now mortgage the very future of life on our planet (cf. RIEDL/DELPOS 1996).

Secularization notwithstanding, it continues to be important to drive home the point that the 'knowledge game' is a quintessentially *human*, not a Promethean or Godlike affair.[8] For, as Paul FEYERABEND (1974) has shown brilliantly in an almost forgotten review of POPPER's 'Three Worlds' ontology, PLATONISM continues to cast its shadow on twentieth-century analytical philosophy in the guise of nonmaterialist ontologies, be they 'interactionist' or other (NOTTURNO 1985; in a different vein, cf. also MATSUNO/SALTHE 1995).

In retrospect, it is fair to say that one of KUHN's lasting merits is his call for an *anthropological turn* in science studies (cf. Bruno LATOUR in CALLEBAUT 1993, p110, 114): the urge to focus, as *objectively* as possible, on the very (individual and, in KUHN's own case, predominantly collective) *subjectivity*—and, I would add, essential *bounded rationality* (Herbert SIMON)—of science so abhorred by FREGE's intellectual heirs, the positivists and the POPPERIANS.[9]

Evolutionary hedgehogs

The naturalistic approach also departs dramatically from the modern BACONIAN and especially CARTESIAN traditions in epistemology in taking a wholesale justification for science and its method to be "an unreasonable demand and a misplaced effort" (NAGEL 1956, p15).[10] Naturalism certainly rejects the view that genuine knowledge must be inferred from the 'self-luminous' and 'self-evident' principles of a first philosophy. William WIMSATT (1996, p1) has characterized our predicament pointedly in the following words: "Foundationalists since DESCARTES have sought certainty, clarity, and economy from generative deductive structures rooted in certain foundations. Whether they sought to anchor ethics, science, or epistemology, and to anchor it in experience, reason, or some combination of the two, all took the same road. Similar images have driven scientific methodologies in the more 'exact' sciences since EUCLID. This search for certainty through deduction has skirted the fact that conceptual systems are constructed by fallible people, and imbedded in a context of institutions, disciplinary practices, and socialized personal aims. The internal validity of practices within a given perspective must be reconceptualized to recognize these richer relationships."

Let me note here that KUHN, for one, was rather critical of the prospect for the sort of integration envisaged by WIMSATT, mainly because he sensed the methodological obstacles: "The development of scientific institutions, values, methods, and world views is clearly in itself a worthy subject for historical research. Experience suggests, however, that it is by no means so nearly coextensive with the study of scientific development as its practitioners have ordinarily supposed. The relationship between the metascientific environment, on the one hand, and the development of particular scientific theories and experiments, on the other, has proved to be indirect, obscure, and controversial," (KUHN 1977, p33)

At any rate, and independently of the previous considerations, the naturalistic approach leaves no room for *extrascientific criteria* as a "warrant for the epistemic role it accords the methods of science" (ROSENBERG 1996, p1), or, alternatively, the sociocultural mechanisms that eventually succeed in producing "good models of reality" if we are lucky (cf.

Ryan TWENEY in CALLEBAUT 1993, p345; CALLEBAUT 1995c).[11] For the warrant for a proposition, NAGEL insisted in his presidential address to the American Philosophical Association in 1955, "does not derive from a faith in the uniformity of nature, or any other principle with cosmic scope" (NAGEL 1956, p15). (A contrario, it might be worthwhile for evolutionary epistemologists who think that human understanding must be necessarily 'limited' because of our 'mesocosmic' predicament to pay heed to this insight as well!)

Continued NAGEL: "The warrant derives exclusively from the specific evidence available, and from the contingent historical fact that the special ways employed in obtaining and appraising the evidence have generally been effective in yielding reliable knowledge." (ibid.). In a similar way, GIERE (1988, p7) argues that "theories come to be accepted (or not) through a natural process involving both individual judgment and social interaction. No appeal to supposed rational principles of theory choice is involved." The "rational principles" GIERE has in mind are the aprioristic principles of rational choice theory. *Whatever 'rationality' is involved in the heuristics scientists use in their decision making must be the result of past individual and collective human experience*, i.e., a posteriori (ontogenetically or phylogenetically).

As EFRON/FISCH (1991, p191) point out, traditional foundationalism and relativism share the presupposition that only *extrascientific* principles can justify our belief that "scientific advance is rational and orderly": philosophical principles or criteria of rationality in the first case, social and political negotiation in the second. Evolutionary naturalism can escape both dead ends—unfathomable first principles here, epistemological disaster (wholesale relativism) there, in favor of an intelligent "dynamical foundationalism" (a label I borrow from WIMSATT 1996) that takes cognizance of relevant human evolutionary origins.

Referring to QUINE, one could say that there is "encouragement in DARWIN" (cf. KITCHER 1993, p300) to transcend *Subjektphilosophie*, i.e., generally speaking, any epistemological investigation that begins (à la DESCARTES or HUME) with a solitary individual facing his or her own immediate, conscious experience. For, "thanks to DARWIN and a century of further research in evolutionary theory and genetics, we now know that no humans ever faced the world with only their subjectively accessible experience to guide them. Our perceptual and other cognitive capacities, which are poorly reflected in our subjective experiences, are very well adapted to the environment in which we evolved. There is no longer any need to wonder why we are not systematically deceived by our environment, or why we associate colors and motions with material objects, or why we perceive things as being in a local EUCLIDIAN space-time, or why we possess an empathetic understanding of our fellow humans. If we had not evolved these capacities, we would not be here." (GIERE 1984, p9)

To my knowledge, KUHN never endorsed Evolutionary Epistemology in this sense, although he was probably influenced by QUINE more deeply than is usually realized (see CAMPBELL in CALLEBAUT 1993).

EFRON/FISCH summarize the evolutionary-naturalist consensus that one sees emerging from recent discussions in the following words (they are specifically addressing GIERE's variety of evolutionary naturalism, viz., 'constructive realism', but their characterization applies to the view of CAMPBELL [1988], HOOKER [1987, 1995], WIMSATT [1996, 1997], and others as well): "The generalizations of cognitive science are *scientific* generalizations about human information processing. Because human information processing skills are part of our evolutionary inheritance, they are effectively stable and universal.… By understanding the nature of these skills, which govern scientific endeavors just as they do all other human information processing, we can come to some objective and, more important, *scientific* understanding of how particular scientists work and of the mechanism of scientific development in general. … Thus … the generalizations of cognitive science provide a firm and objective basis for generalizations about how science is done." (EFRON/FISCH 1991:189–190)

NAGEL, who ventured to write that "there is no one 'big thing' which, if known, would make everything else coherent, and unlock the mystery of creation", could not anticipate the special role DARWINIAN evolutionary theory has come to play in latter-day evolutionary naturalism (ROSENBERG 1996, p3). DENNETT's (1995) vindication of an all-encompassing DARWINIAN philosophy may be the best example of a contemporary naturalist having become a "hedgehog" (NAGEL invoked Isaiah BERLIN's simile of the fox and the hedgehog). The fascination that DARWINISM has come to have for latter-day naturalists even incites ROSENBERG (1996,4) to view DARWINISM as one of the three tenets of contemporary naturalism, along with the repudiation of 'first philosophy' and "scientism", viz., that the sciences are to be the guide to epistemology and metaphysics.

I conclude this section by noting that KUHN's view of science departs significantly from latter-day evolutionary naturalism in at least two important respects: (i) As an anti-foundationalist, he seems to have consistently eschewed the tendency to find "encouragement in DARWIN" that is endemic in evolutionary epistemology in the sense of bioepistemology, BRADIE's (1986) EEM ("Evolutionary Epistemology of Mechanisms"), or evolutionary psychology. (ii) As an anti-realist with neo-KANTIAN inklings (GUTTING 1980; HOYNINGEN-HUENE 1993)—viz., arguing that paradigms define their own "worlds"—, he actually did not resonate with the kind of expectations that make realist evolutionary naturalists tick (CALLEBAUT 1993, 1995a,b; KITCHER 1993; THOMPSON 1995; ROSENBERG 1996; contrast PUTNAM 1990).

The evolutionary analogy—pushed too far, or not far enough?

"Need there be any goal for science?"

Mario BUNGE (1962/1979) offered a classical statement of the historical naturalization of western science in terms of (i) a progressive *restriction of causation to natural science*, beginning with certain pre-Socratic philosophers of nature, (ii) the further restriction of all varieties of natural causation to *efficient* causation ('the only legitimate sort of cause is an efficient cause; ends are not efficient causes; so explanation by appeal to ends is illegitimate'—cf. LENNOX 1983), (iii) the endeavor to reduce efficient causes to *physical* ones (*mechanism*), (iv) the requirement of *testing* causal hypotheses by means of repeated observations and, whenever possible, through reproduction in controllable experiments, and (v) *parsimony*, understood as both caution in the assignment of causes and minimization of the number of 'ultimate' natural causes.

We have seen before that KUHN put great emphasis on the need to articulate a testable account of science. Toward the end of *The Structure of Scientific Revolutions*, in a section called "Progress through Revolutions", KUHN also addressed BUNGE's other requirements, tackling (ii) directly, and, by drawing a parallel between his account and the DARWINIAN revolution in 19th-century biology, indirectly pointing to (i), (iii), and (v) as well.

We are all "deeply accustomed to seeing science as the one enterprise that draws constantly nearer to some goal set by nature in advance"—truth—, he claims, and suggests it is time to replace this teleological view by a causal account: "Can we not account for both science's existence and its success in terms of evolution from the community's state of knowledge at any given time? Does it really help to imagine that there is some one full, objective, true account of nature and that the proper measure of scientific achievement is the extent to which it brings us closer to that ultimate goal? If we can learn to substitute evolution-from-what-we-do-know for evolution-toward-what-we-wish-to-know, a number of vexing problems may vanish in the process." (KUHN 1970a, p171)

KUHN's statement of the "transposition" that he "recommended" was crystal-clear. Yet, if one considers the hopeless blurring of the distinction between *reasons* (the stuff of foundational epistemology) and *causes* (the stuff of naturalized epistemology) that characterized much of the debate that followed the publication of KUHN's major book (e.g., SPIEGEL-RÖSING/PRICE 1977), one is drawn to the conclusion that his point remained largely unnoticed.

KUHN granted that he could not yet "specify in any detail" the consequences of his alternate view of scientific advance, but found the parallel with Darwinism "particularly helpful because in both cases the main obstacle to transposition is the same", viz., *the abandonment of (PLATONIST) teleology*:[12] "All the well-known pre-DARWINIAN evolutionary theories — those of LAMARCK, CHAMBERS, SPENCER, and the German *Naturphilosophen* — had taken evolution to be a goal-directed process. The "idea" of man and of the contemporary flora and fauna was thought to have been present from the first creation of life, perhaps in the mind of God. That idea or plan had provided the direction and the guiding force to the entire evolutionary process. Each new stage of evolutionary development was a more perfect realization of a plan that had been present from the start." (KUHN 1970a, p171–172)

The *Origin of Species* replaced goals (set either by God or nature) by the operation of natural selection, operating in the given environment and with the actual organisms presently at hand. KUHN invoked the famous examples of the eye and hand—organs "whose design had previously provided powerful arguments for the existence of a supreme artificer and an advance plan"—to compare them to the resolution of revolutions as "the selection by conflict within the scientific community of the fittest way to practice future science". He specified that "the analogy that relates the evolution of organisms to the evolution of scientific ideas can easily be pushed too far. But with respect to the issues of this closing section … it is very nearly perfect" (KUHN 1970a, 172):

"The net result of a sequence of such revolutionary selections, separated by periods of normal research, is the wonderfully adapted set of instruments we call modern scientific knowledge. Successive stages in that developmental process are marked by an increase in articulation and specialization. And the entire process may have occurred, as we now suppose biological evolution did, without benefit of a set goal, a permanent fixed scientific truth, of which each science in the development of scientific knowledge is a better exemplar." (KUHN 1970a, 172–173)

I quoted extensively from this passage because it contains KUHN's most explicit and elaborate discussion of what prima facie seems to amount to a (cultural-)evolutionary underpinning of his account of scientific change (cf. CAMPBELL 1997). One factor complicating a proper assessment is KUHN's evident (relative) unfamiliarity with evolutionary biology present *and* past; for example, he does not seem to have envisaged the very possibility of nonadaptive evolution, although its use in an analogy could have strengthened his case against realism tremendously. This points to a more general weakness of his account: although it claims to address biology in addition to physics, the biological sciences are "extremely underrepresented in his examples" (HOYNINGEN-HUENE 1993: 4–5), and his own suggestion that the DARWINIAN revolution fits his account of revolutionary science has been refuted in painstaking detail by John GREENE (1971) and others as well. At any rate, our conclusion must be sobering: What KUHN actually does here is vindicate naturalism (in the sense specified by BUNGE), but not evolutionary epistemology!

Modeling science on an inadequate understanding of evolution

Rather than embarking on a speculation here as to the precise meaning and possible implications of KUHN's evolutionary analogy that would remain inconclusive anyway, I am tempted to endorse HULL's (1986, p643) general conclusion vis-a-vis the evolutionary analyses of KUHN, TOULMIN, and others, viz.: "The major reason for evolutionary analyses of conceptual change being so unsatisfactory is that they are modeled on an inadequate understanding of biological evolution". I mention some of the major flaws in KUHN's account, as identified by recent evolutionary epistemology:

■ Thus KUHN does not seem to be aware that by deploying his stage model as a model of evolution (phylogeny), he falls victim to the *confusion of evolutionary (or "variational") explanation at the level of populations or species and developmental (ontogenetic) explanation at the level of individual organisms* such as scientists. Richard LEWONTIN, Elliot SOBER and others have offered careful analyses of these two distinct types of explanation, which are both perfectly legitimate when applied within their respective domains (CALLEBAUT 1993, p144–147, 334).

■ A consistent use of the evolutionary rather than the developmental analogy could also have prevented KUHN from mapping tradition (the conservative moment of scientific change) one-to-one on the 'normal science' stage and innovation one-to-one on the 'revolutionary' stage, in favor of a much more realistic *'coexistence' model of tradition and innovation* (as he actually imagined himself in KUHN 1977). By way of example: KUHN's attention to the importance of (generally unconscious) changes in the meaning of scientific terms that can signal profound shifts in worldviews (cf. MARGOLIS 1993) has undermined the traditional goal of univocality and precision for scientific language. This possibility recedes yet further if one believes that meanings do not simply change, but, in a certain sense, accumulate—"carry[ing] the mark of the historical (sedimented) circumstances of their origin and use in ever new ways." (EDIE 1976, p154–158).

■ The essential irreversibility of evolution *('DOLLO's law') cannot be captur*ed by means of KUHN's (crucial) metaphor of the Gestalt switch, for Gestalt switches are time-indifferent (WATANABE 1975; CALLEBAUT 1990). In light of this, one can also better appreciate RICHARDS' (1987, p561) claim that "[the static] model of the origin and course can be detected in transmogrified form in Thomas KUHN's Gestalt model …, which assumes that in a moment of insight the transformed vision of an inspired genius may establish the framework and fundamental premises of a science, the details of which may be left to the normal plodding of disciples."

Gestalt switches seem more PLATONISTIC than DARWINIAN indeed!

■ Finally, attention to the details of the evolutionary analogy (and the ecological setting in which evolution takes place: see already BLACKBURN 1973) would have allowed KUHN to better delineate his units of analysis, viz. individual scientists on the one hand, and paradigms and communities on the other. As GIERE (1989, p8) notes, "what is not debatable is that for any evolutionary model to work, individual scientists must at least figure prominently among the units of analysis. It is variations among individuals that make an evolutionary model work." Hence it makes sense to conceive of a complementary relationship between evolutionary models of science as a whole and cognitive models of individual scientists (e.g., GIERE 1992; THAGARD 1988, 1992). As to communities and paradigms, neither form discrete natural units. Moreover, as HULL (1986, p293–294) has shown, "contrary to KUHN, the boundaries of scientific communities and paradigms do not always coincide perfectly. A scientist can be sociologically part of a research group even though he does not share its paradigm, and many scientists who share the same paradigm are not part of the same scientific community."

Better delineated units would in turn have allowed devoting more attention to the extreme variability of both conceptual and social evolution. There is far less consensus in science than KUHN used to suggest.

Marrying self-organization and evolution

The confusion between evolution and development which KUHN was prone to was tantamount in biology before the triumph of WEISMANNISM and still occurs frequently in the social sciences today. In this context, it may be interesting to point out that KUHN's "normal science" may be likened to PIAGET's "assimilation", and that both may be viewed as forms of *internal selection*; whereas periods of "revolutionary science", during which science is more vulnerable to *external selection*, may be profitably compared to "accommodation" (CAMPBELL in CALLEBAUT 1993, p296). This rapprochement points to the possibility of delineating 'development' and 'evolution' in a proper way. Here recent attempts to integrate insights from the study of biological and cognitive development into the study of evolution become relevant (CALLEBAUT 1993, ch. 9; STOTZ 1996).

For instance, it is populations or species that evolve, not individuals. In the logic of variational explanation, evolutionary change is the result of the replacement of individuals within a population and thus does not require that the individuals who compose the population change themselves. Adding *developmental constraints* to this picture—opening up the black box that the Modern Synthesis turned the organism into, so to speak—may complicate our extended picture of evolution considerably (GLASSMAN/WIMSATT 1984; WIMSATT 1986; CALLEBAUT/STOTZ 1997). It may also provide an important source model for analogies between biological evolution and scientific change that, among other things, may strengthen KUHN's model.

The key notion in WIMSATT's approach is *generative entrenchment*. Intuitively, the idea is "to focus upon relationships among parts of a developing structure and its environment, and then to notice that things which play a generative role, wherever they are located, are gene-like in many of their characteristics" (WIMSATT 1996, p1). Rather than "preestablished genes" using generated structures as survival means (e.g., DAWKINSIAN "vehicles"), the unit of analysis becomes a generative structure, and the focus of attention the developmental consequences of pivotal elements, which are *contextually* defined as 'gene-like'. Both heritability and fitness are properties of the relationship between organism and environment, and development is viewed as intrinsically relational. "These reinterpretations point the way to non-reductionist extensions of evolutionary ideas into the social and cultural realms" (WIMSATT 1996).

In the context of scientific change, "the survival of a variant under the most diverse and adverse conditions is mirrored … by the survival of those discoveries and concepts that find usefulness in the greatest variety of further applications—of those conceptual schemes that withstand the constant check against experience" (HOLTON 1953, p97–98)

WIMSATT and his associates have developed both analytical models and more realistic simulations that predict which parts of a generative structure are likely to be preserved, or, alternatively, to change, independently of the nature of the entities involved (they might be elements in a generated network of inferences, laws or consequences in a scientific theory, technological parts or procedures, rituals or institutions in an economy or normative system, meanings, memes, structures or behavioral traits in a developing phenotype, or what-have-you.

When generative entrenchment is applied to KUHN's account of scientific change, it allows to explain why scientific revolutions must remain rare: the more fundamental the change, (i) the less likely it will work, and (ii) the broader its effects; so (iii) the more work it will make for others, (iv) who therefore resist it ac-

Author's address

Werner Callebaut, Konrad Lorenz Institut for Evolution and Cognition Research, A–3422 Altenberg/Donau, Austria, and Department of Sciences, Limburgs Universitair Centrum, B–3590 Diepenbeek, Belgium; Department of Philosophy, Maastricht University, P.O. Box 616, 6200 MD Maastricht, the Netherlands.
Emails: callebaut@kla.univie.ac.at; callebaut@philosophy.unimaas.nl.

tively (WIMSATT 1986b, 1996). (iii) and (iv) are institutional, social, or social-psychological, and notoriously difficult to quantify, one reason being that weights tend to shift considerable according to one's perspective as to where (s)he will end up after the revolution (WERTHEIM 1974; BARNARD 1979; SCRIVEN 1979).

But (i) and (ii) are "robust structural and causal features of our world and of our schemes—inescapable features of material and abstract generative structures" (WIMSATT 1996).

This concludes our brief review of KUHN's importance as an evolutionary naturalist (see also CALLEBAUT/STOTZ 1997).

Notes

1 A reprint, enlarged with a Postscript written in 1969, of the original 1962 publication in vol. 2 of the International Encyclopedia of Unified Sciences.

2 How radically KUHN's account actually departed from the 'received view' is a topic of current debate. Thus it is interesting to note that Rudolf CARNAP, in his correspondence with KUHN, expressed large agreement with the latter's view of scientific change, and wrote that theory change is better viewed as "the improvement of an instrument" than as a search for "the ideal system". CARNAP also agreed that scientific revolutions bring with them proposals for a new conceptual and linguistic framework (AXTELL 1993, p121).

3 The idea of a degree of freedom for alternative developments in science was also the very starting point of the 'finalization' thesis of the so-called STARNBERG group (BÖHME/VAN DEN DAELE/KROHN 1973), which at one time caused quite a stir in German public life and the media (PFETSCH 1979). Put succinctly, the STARNBERG group argued that scientific change typically follows a KUHNIAN stage or phase pattern allowing for "an orientation of theory development to external goals" in the post-paradigmatic stage. "The reason for this", the STARNBERGERS believed, "is that in this phase there exists no internal logic which selects and regulates the direction and the problems of scientific development" (W. VAN DEN DAELE 1977, quoted in PFETSCH 1979, p118).

4 The presumed incommensurability of paradigms shouldn't pose a fundamental problem to the evolutionary approach either (CALLEBAUT 1995b; cf. also KITCHER 1983; MILLER 1991; MALONEY 1993; and SANKEY 1993).

5 In CALLEBAUT (1993) I called science studies "an emerging field", having in mind the rather specific technical meaning given to this term by Dudley SHAPERE in his papers "Scientific Theories and Their Domains" (1974) and "Remarks on the Concepts of Domain and Field" (1977), reprinted in Shapere (1984). This reflected an optimistic belief I shared with many sympathetic observers at the time: that under the beneficial influence of new currents in the sociology of science such as social constructivism and, maybe even more so, "the exciting perspectives on science now emerging from Paris", viz. Michel CALLON and Bruno LATOUR's actor-network theory (COZZENS/GIERYN 1990, 1–2), the previous "marriage of convenience" of the history and philosophy (GIERE 1973, BURIAN 1987) would turn into a "family romance" (MANIER's 1980) between, on the one hand, the more theoretical approaches to science studies—philosophy of science, cognitive studies of science (e.g., GIERE 1992), and now also the economics of science (MIROWSKI/SENT 1997)—, and, on the other, the case study approach typical of the new history and sociology of science on the other.
Alas, it seems to me that of late, the sociology of science has been showing symptoms of (exaggerated) self-sufficiency; to become convinced of this one only has to check the proportion of 'external' as opposed to 'internal' references in the index of a major collection such as JASANOFF et al. (1995). And as the most recent volumes of the proceedings of the Philosophy of Science Association document, the philosophy of science community, which has been on the defensive ever since KUHN launched his "academic challenge to positivist epistemology" (ELZINGA/JAMISON 1995, p573), reacts to (what it perceives as) its rejection by a comparable, and equally regrettable, retrenchment.

6 My point presupposes that in science in general and in scientific studies of science ('metascience') in particular, it is, on average, theory-driven rather than data-driven research that will turn out to be most rewarding intellectually (and, conceivably, also practically).
Although this conviction would seem to be consistent with KUHN's overall view, it may, in my own case, betray a realist inclination KUHNIANS cannot be expected to share. Cf. HARRÉ (1986, p29): "How little physics would have advanced had physicists from Bradwardine onwards eschewed philosophical analysis of concepts or refrained from ontological speculation! The real community of scientists does not work within the [positivist or empiricist] myth, though it often uses it as a rhetoric."
It is ironic indeed that an empiricist methodology of the most naive "We're-only-trying-to-let-the-facts-speak-for-themselves" variety seems to underlie many of the current case studies in STS (Science and Technology Studies)!

7 See, e.g., Latour in CALLEBAUT (1993, p76, p111–114 and passim).

8 As one major evolutionary epistemologist, who otherwise strongly opposes essentialist distinctions, ventures to put it, "Success in the knowledge game is hardly an incidental feature of Homo sapiens. It is our chief adaptation. It is the only thing in the struggle for existence that we do better than any other species." (HULL 1988, p26). I have explored the metaphor of science as a game in some detail in CALLEBAUT (1995a).

9 This is actually less paradoxical than it may seem. What I have in mind here is first and foremost KUHN's (would-be) 'no nonsense', descriptive account of science as opposed to traditional normative accounts that, from the Vienna Circle to FEYERABEND, were ultimately more concerned with emulating Enlightenment ideals than with descriptive adequacy (CALLEBAUT 1995b). I specify, 'would be', because one ought to sharply distinguish here between KUHN's naturalistic program and the soundness of his actual account (at the general level and/or at the level of his case studies), which is taken to be highly problematic by philosophers, historians and sociologists of science alike (e.g., KOURANY 1979; BARNES/EDGE 1982, p4–8; LAUDAN et al. 1988, p5–7); cf. AXTELL (1993) on what he considers Kuhn's best empirical work, viz. KUHN (1978).
At a deeper philosophical level, the evolutionary epistemologist who is also a realist—currently the majority viewpoint—may argue that the naturalization of the epistemic subject (individual or collective, as in the case of KUHNIAN

paradigms-as-scientific-communities) situates the subject spatiotemporally in the (one) external world, thus 'objectifying' it to a certain extent (cf. Donald CAMPBELL's 1969/1988 "phenomenology of the other one"). The KUHNIAN's neo-KANTIAN, constructivist creed that 'epistemic subjects create their own worlds' of course precludes any sort of rapprochement of object and subject through the latter's 'objectification'.
This is but a special case of the general problem that genetically object-sided and genetically subject-sided moments are inextricably united in "the network of similarity and dissimilarity relations coconstitutive of a given phenomenal world", to paraphrase HOYNINGEN-HUENE's (1993, p270–271) words. As a result, an "undistorted view of the purely object-sided, or absolute reality or of the world-in-itself" is not to be had by us humans: "The concrete properties of the world-in-itself are, rather, inaccessible; though we feel the effects of these properties in the resistance the world offers to our epistemic efforts, we aren't in a position to grasp this resistance as it is in itself."

10 As ROSENBERG (1996) reminds us, Ernest NAGEL is an important pioneer of naturalism in his own right, although he is seldom acknowledged as such—Willard Van Orman QUINE being the philosopher who has taken most of the credit for the current revival of naturalism.

11 On the conventional interpretation of KUHN's theory of science (GUTTING 1980a, 1), the latter is about the authority of science (science being "the only generally recognized cognitive authority in the world today"). Rather than being a matter of "a rule-governed method of inquiry" whereby reliable knowledge is obtained, "science's authority ultimately resides … in the scientific community that obtains the results". As I have argued elsewhere (CALLEBAUT 1995b, 7f.), rather than implying the presumed 'end of rationality', such a sociologization of epistemology (cf. FULLER 1988) implies the relocation of scientific rationality—which is now being defined instrumentally and procedurally rather than teleologically and substantively—within the structure of scientific communities.

12 For an elaboration of the distinction between PLATONIST and ARISTOTELIAN varieties of teleology, see LENNOX (1992, 1994). As to LENNOX' suggestion that DARWIN was, after all, a teleologist, contrast GHISELIN (1969, 1994).

References

Andersen, E. S. (1991) Techno-economic paradigms as typical interfaces between producers and consumers. Journal of Evolutionary Economics 1:114–144.

Axtell, G. S. (1993) In the tracks of the historicist movement: Re-assessing the Carnap—Kuhn connection. Studies in the History and Philosophy of Science 24:119–146.

Barnard, F. M. (1979) The evaluation of revolutions: A comment on Michael Scriven's paper. In: Johnson, H. J./Leach, J. J./Muehlmann, R. G. (eds.) Revolutions, Systems, and Theories, D. Reidel, Dordrecht, 11–19.

Barnes, B./Edge, D. (1982) General introduction. In: Barnes, B./Bloor, D. (eds.), Science in Context: Readings in the Sociology of Science. Open University Press, Milton Keynes, 1–12.

Blackburn, T. R. (1973) Information and the ecology of scholars. Science 181: 1141–1146.

Blachowicz, J. A. (1971) Systems theory and evolutionary models of the development of science. Philosophy of Science 38:178–199.

Böhme, G./van den Daele, W./Krohn, W. (1973) Die Finalisierung der Wissenschaft. Zeitschrift für Soziologie 2:128–144.

Bradie, M. (1986) Assessing evolutionary epistemology. Biology and Philosophy 1:401–450.

Bunge. M. (11962/31979) Causality and Modern Science. Dover, New York.

Burian, R. M. (1987) More than a marriage of convenience: On the inextricability of history and philosophy of science. Philosophy of Science 44:1–42.

Callebaut, W. (1990) Temps et épistémologie: Remarques sur une variable absente en philosophie des sciences. In: Thill, G./Feltz, B./Lambert, D. (éds.), Actes du Congrès international de l'Association Internationale de Cybernétique. Association Internationale de Cybernétique, Namur.

Callebaut, W. (1993) Taking the Naturalistic Turn, Or How Real Philosophy of Science is Done. University of Chicago Press, Chicago.

Callebaut, W. (1995a) The future of naturalistic philosophy of science. Ludus Vitalis 3(5):19–52.

Callebaut, W. (1995b) Science dynamics: The difficult birth of a metascience. In: Stachowiak, H. (Hrsg.) Pragmatik: Handbuch des pragmatischen Denkens, Band V. Felix Meiner, Hamburg, 1–29.

Callebaut, W. (1995c) The normative issue in naturalistic philosophy of science. Theoria 10:101–116.

Callebaut, W. (1997 forthcoming) Neonaturalism in contemporary theories of science. In: Kampis, G./Nowotny, H. (eds.), Proceedings of the Workshop on Current Trends in the History, Philosophy and Social Studies of Science, Collegium Budapest, 6 November 1995, Institute for Advanced Studies, Budapest.

Callebaut, W./Stotz, K. (1997 forthcoming) Position paper for the 2nd Altenberg Workshop in Theoretical Biology, "Evolutionary Naturalism 1997: Bioepistemology and the Challenge of Development and Sociality". Konrad Lorenz Institut, Altenberg/Donau.

Callebaut, W./Van Meer, E. (1997 forthcoming) What can evolutionary epistemology contribute to our understanding of technological chnage? In: Ziman, J. (ed) Proceedings of the 1997 Epistemological Group Workshop on Technological Evolution. Cambridge Univ. Press: Cambridge.

Campbell, D. T. (1988) Methodology and Epistemology for Social Science (Overman, E. S., ed.). University of Chicago Press, Chicago.

Campbell, D. T. (1997 forthcoming) From Evolutionary Epistemology via Selection Theory to a Sociology of Scientific Validity (Frankel, B./Heyes, C. M., eds.). Evolution and Cognition 3(1).

Cozzens, S. E./Gieryn, T. F. (1990) Introduction: Putting science back in society. In: Cozzens, S. E./Gieryin, T. F. (eds.), Theories of Science in Society. Indiana University Press, Bloomington/Indianapolis, 1–14.

Dennett, D. C. (1995) Darwin's Dangerous Idea. Harper and Row, New York.

Edie, J. M. (1976) Speaking and Meaning. Indiana University Press, Bloomington.

Elzinga, A./Jamison, A. (1995) Changing policy agendas in science and technology. In Jasanoff et al., 572–597.

Feyerabend, P. (1974) Popper's Objective Knowledge. Inquiry 17:475–507.

Fuller, S. (1988) Social Epistemology. Indiana University Press, Bloomington.

Ghiselin, M. T. (1969) The Triumph of the Darwinian Method. University of California Press, Berkeley/Los Angeles.

Ghiselin, M. T. (1994) Darwin's language may seem teleological, but his thinking is another matter. Biology and Philosophy 9:489–492.

Giere, R. N. (1973) History and philosophy of science: Intimate relationship or marriage of convenience? British Journal for the Philosophy of Science 24:282–297.

Giere, R. N. (1984) Toward a unified theory of science. In Cushing, J. T./Delaney, C. F./Gutting, G. M. (eds.), Science and Reality, Notre Dame University Press, Notre Dame, IN, 5–31.

Giere, R. N. (1985) Philosophy of science naturalized. Philosophy of Science 52:331–356.

Giere, R. N. (1988) Explaining Science: A Cognitive Approach. University of Chicago Press, Chicago.

Giere, R. N. (1989). The units of analysis of science studies. In: Fuller, S. (ed.) The Cognitive Turn, Kluwer, Dordrecht, 3–11.

Giere, R. N. (ed.) (1992) Cognitive Models of Science. Minnesota Studies in the Philosophy of Science, vol. 15. University of Minnesota Press, Minneapolis.

Glassmann, R. B./Wimsatt, W. C. (1984) Evolutionary advantages and limitations of early plasticity.In: Almli, R./Finger, S. (eds.), Early Brain Damage, vol. I. Academic Press, New York, 35–58.

Greene, J. C. (1981) Science, Ideology and World View: Essays in the History of Evolutionary Ideas. University of California Press, Berkeley/Los Angeles.

Griesemer, J. R./Wimsatt, W. C. (1988) Picturing Weismannism: A case study in conceptual evolution. In: Ruse, M. (ed.), What Philosophy of Biology Is. Martinus Nijhoff, Dordrecht, 75–137.

Gutting, G. (ed.) (1980) Paradigms and Revolutions: Applications and Appraisals of Thomas Kuhn's Philosophy of Science. University of Notre Dame Press, Notre Dame, IN.

Hacking, I. (1981) Introduction. In: Hacking, I. (ed.), Scientific Revolutions. Oxford University Press, Oxford.

Harré, R. (1986) Varieties of Realism: A Rationale for the Natural Sciences. Basil Blackwell, Oxford.

Holton, G. (1953) On the duality and growth of physical science. American Scientist 41:89–97.

Hoyningen-Huene, P. (1992) The interrelations between the philosophy, history and sociology of science in Thomas Kuhn's theory of scientific development. British Journal for the Philosophy of Science 43:487–501.

Hoyningen-Huene, P. (1993) Reconstructing Scientific Revolutions: Thomas S. Kuhn's Philosophy of Science. University of Chicago Press, Chicago.

Hooker, C. A. (1987) A Realistic Theory of Science. State University of New York Press, Albany, NY.

Hooker, C. A. (1995) Reason, Regulation and Realism: Toward a Naturalistic, Regulatory Systems Theory of Reason. State University of New York Press, Albany, N.Y.

Hull, D. L. (1986) Conceptual evolution and the eye of the octopus. In: Barcan Marcus, R. et al. (eds.), Logic, Methoology and Philosophy of Science, vol. 7, Elsevier, Amsterdam, 643–665. Reprinted in Hull 1989: 221–262.

Hull, D. L. (1988) Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science. University of Chicago Press, Chicago.

Hull, D. L. (1989) The Metaphysics of Evolution. State University of New York Press, Albany.

Jasanoff, S./Markle, G. E./Petersen, J. C./Pinch, T. (eds.) (1995) Handbook of Science and Technology Studies. SAGE, London.

Khalil, E. (1996) Economic action, naturalism, and purposefulness. Research in the History of Economic Thought and Methodology 14:119–140.

Kitcher, P. (1983) Implications of incommensurability. In: PSA 1982, vol. 2. Philosophy of Science Association, East Lansing, MI, 689–703.

Kitcher, P. (1993) The Advancement of Science: Science Without Legend, Objectivity Without Illusions. Oxford University Press, New York/Oxford.

Kourany, J. (1979) The nonhistorical basis of Kuhn's theory of science. Nature and System 1:46–59.

Kuhn, T. S. (1970a) The Structure of Scientific Revolutions. 2nd ed. University of Chicago Press, Chicago.

Kuhn, T. S. (1970b) Reflections on my critics. In: Lakatos, I./Musgrave, A. (eds.) Criticism and the Growth of Knowledge, Cambridge University Press, Cambridge, 231–278.

Kuhn, T. S. (1977) The Essential Tension: Selected Studies in Scientific Tradition and Change. University of Chicago Press, Chicago.

Kuhn, T. S. (1978) Black-Body Theory and the Quantum Discontinuity, 1894–1912. University of Chicago Press, Chicago.

Kuhn, T. S. (1980) The halt and the blind. British Journal for the Philosophy of Science 31:181–192.

Kuhn, T. S. (1983) Reflections on receiving the John Desmond Bernal Award. 4S Review: Journal of the Society for Social Studies of Science 1:26–30.

Laudan, R./Laudan, L./Donovan, A. (1988) Testing theories of scientific change. In: Donovan, A./Laudan, L./Laudan, R. (eds.) Scrutinizing Science: Empirical Studies of Scientific Change. Kluwer, Dordrecht, 3–44.

Lefèvre, W. (1995) Material and social conditions in an historical epistemology of scientific thinking. Preprint no. 14, Max-Planck-Institut für Wissenschaftsgeschichte, Berlin (11 pp.).

Lennox, J. G. (1992) Teleology. In: Keller, E. F./Lloyd, E. A. (eds.) Keywords in Evolutionary Biology, Harvard University Press, Cambridge, MA, 323–333.

Lennox, J. G. (1994) Teleology by another name: A reply to Ghiselin. Biology and Philosophy 9:493–495.

Luhmann, L. (1990) Die Wissenschaft der Gesellschaft. Suhrkamp. Frankfurt am Main.

Maloney, M. E. (1993) Kuhn reconstructed: Incommensurability without relativism. Studies in the History and Philosophy of Science 24:69–93.

Manier, E. (1980) History, philosophy and sociology of science: A family romance. Studies in the History and Philosophy of Science 11:1–24.

Margolis, H. (1993) Paradigms and Barriers: How Habits of Mind Govern Scientific Beliefs. University of Chicago Press, Chicago.

Maslow, S. Yu. (1978) Macroevolution as deduction process. Synthese 39:417–434.

Matsuno, K./Salthe, S. N. (1995) Global idealism/local materialism. Biology and Philosophy 10:309–337.

Miller, A. I. (1991) Have incommensurability and causal theory of reference anything to do with actual science?—incommensurability, no; causal theory, yes. International Studies in the Philosophy of Science 5:97–108.

Mirowski, P. E/Sent, E.-M. (1996) Announcement, Conference on the Need for a New Economics of Science, University of Notre Dame, IN, March 13–16, 1997.

Nagel, E. (1956) Logic Without Metaphysics. Free Press, Glencoe, IL.

Notturno, M. A. (1985) Objectivity, Rationality and the Third Realm: Justification and the Grounds of Psychologism. A Study of Frege and Popper. Martinus Nijhoff, Dordrecht.

Pfetsch, F. R. (1979) The 'finalization' debate in Germany: Some comments and explanations. Social Studies of Science 9:115–124.

Pickering, A. (ed.) (1992) Science as Practice and Culture. University of Chicago Press, Chicago.

Popper, K. R. (1970) Normal science and its dangers. In: Lakatos, I./Musgrave, A. (eds.) Criticism and the Growth of Knowledge, Cambridge University Press, Cambridge, 51–58.

Putnam, H. (1990) Realism With a Human Face (Conant, J., ed.). Harvard University Press, Cambridge, MA.

Richards, R. J. (1987) Darwin and the Emergence of Evolutionary Theories of Mind and Behavior. University of Chicago Press, Chicago.

Riedl, R./Delpos, M. (eds) (1996) Die Ursachen des Wachstums. Orac, Wien.

Rosenberg, A. (1996) A field guide to recent species of naturalism. British Journal for the Philosophy of Science 47:1–29.

Sankey, H. (1993) Kuhn's changing concept of incommensurability. British Journal for the Philosophy of Science 44:759–774.

Scriven, M. (1979) The evaluation of revolutions. In: Johnson, H. J./Leach, J. J./Muehlmann, R. G. (eds.) Revolutions, Systems, and Theories, D. Reidel, Dordrecht, 1–9.

Shapere, D. (1984) Reason and the Search for Knowledge: Investigations in the Philosophy of Science. D. Reidel, Dordrecht.

Stengers, I. (1993) L'invention des sciences modernes. La Découverte, Paris.

Stotz, K. (1996) The psychology of knowledge in the context of evolutionary theory: Reflections on the link between cognition and sociability. Evolution and Cognition 2:16–36.

Thagard, P. (1988) Computational Philosophy of Science. MIT Press, Cambridge, MA.

Thagard, P. (1992) Conceptual Revolutions. Princeton University Press, Princeton, NJ.

Thompson, P. (1995) Evolutionary epistemology and scientific realism. Journal of Social and Evolutionary Systems 18:165–191.

Van Gelder, L. (1996) Thomas Kuhn, 73; devised science paradigm (obituary). New York Times, June 19, p. B7.

Watanabe, S. (1975) Needed: A historico-dynamical view of theory change. Synthese 32:113–134.

Wertheim, W. F. (1974) Evolution and Revolution: The Rising Waves of Emancipation. Penguin Books, Harmondsworth.

Wimsatt, W. C. (1986a) Developmental constraints, generative entrenchment, and the innate-acquired distinction. In: Bechtel, W. (ed.), Integrating Scientific Disciplines. Martinus Nijhoff, Dordrecht, 185–208.

Wimsatt, W. C. (1986b) Generative entrenchment, scientific change, and the analytic-synthetic distinction: A developmental model of scientific evolution. Unpublished ms., University of Chicago.

Wimsatt, W. C. (1996) The evolution of generative structures: The new contextualism and a new dynamical foundationalism. Unpublished ms., University of Chicago.

Wimsatt, W. C. (1997 forthcoming) Approximations to Reality: A Realist Philosophy of Science for Limited Beings. Harvard University Press, Cambridge, MA.

Gary T. Dempsey[1]

Hayek's Evolutionary Epistemology, Artificial Intelligence, and the Question of Free Will

Abstract

This paper examines the Evolutionary Epistemology of the Austrian economist Friedrich HAYEK. I argue that HAYEK embraces a connectionist theory of mind that exhibits the trial-and-error strategy increasingly employed by many artificial intelligence researchers. I also maintain that HAYEK recognizes that his epistemology undermines the idea of free will because it implies that the mind's operation is determined by evolutionary interaction of the matter that comprises ourselves and the world around us. I point out, however, that HAYEK responds to this implied determinism by explaining that it can have no practical impact on our day-to-day lives because, as he demonstrates, the complexity of the mind's evolution prevents us from ever knowing how we are determined to behave. Instead, we can only know our mind at the instant we experience it.

Key words

Connectionism, emergent system, long-term potentiation, nonmonotonic reasoning, physiological memory, self-organizing maps, spontaneous order.

In the field of economics, Friedrich HAYEK has been long recognized for his contributions to the discipline, and in fact was awarded the Nobel Prize in 1974 for his pioneering work in the theory of money and economic fluctuations. Similar attention, however, has not been paid to other facets of his work. This is unfortunate considering HAYEK's important epistemological insights.

Throughout *The Sensory Order* (1952) and other writings, HAYEK makes it clear that the apparatus that allows us to know the world—the mind—is itself subject to evolution; that is, it is a 'work in progress' prone to modification by experience. The mind, he explains, is "incessantly changing" (1984, p243) and its contents constitute an adaptive "capacity to respond to [its] environment with a pattern of actions that helps [it] to persist" (1973, p18). This view puts HAYEK squarely in the camp of the evolutionary epistemologists. Indeed, like Donald CAMPBELL (1960, 1974), Karl POPPER (1972, 1984, 1987), Konrad LORENZ (1977, 1982), and other expositors of the evolutionary model of human knowledge, HAYEK maintains that knowledge is the product of trial-and-error learning and that our minds are characterized by gains in adaptive advantage due to the selective retention of useful representations of the physical world.

It is not surprising, then, that scholars have concluded that HAYEK's epistemology has an essentially "evolutionary character" (KUKATHAS 1989, p49), or that HAYEK takes the "evolutionist standpoint" (GRAY 1986, p11) or the "evolutionary perspective" (VANBERG 1994, p96) in his epistemology. In the following essay, I attempt to piece together HAYEK's epistemology and to explore the particulars of its evolutionary quality. I begin with a brief overview of HAYEK's connectionist theory of mind, followed by an account of his evolutionary epistemology. This latter section consolidates HAYEK's diverse thoughts on evolution to bring them into sharper focus for our discussion and shows how his view anticipates a number of conceptual developments in artificial intelligence. Finally, I explain how HAYEK's epistemology undermines the idea of free will and how he responds to this claim.

The Connectionist Mind

At bottom, HAYEK is a materialist. For him, there is no mind-body split. Instead, all our thoughts, memories, and ambitions result from the operation of matter. Indeed, for HAYEK, "the assertion that...mental phenomena are 'nothing but' certain complexes of physical events [is] probably defensible" (1989, p88). Or more assertively, "mind is...the order prevailing in a particular part of the physical universe—that part of it which is ourselves" (1952, p178).

HAYEK's materialism begins with the recognition that the locus of the mind—the human brain—is made up of a vast weave of fibrous cells called neurons; the cerebral cortex being the most dense with more than ten-thousand million. Each of these neurons, in turn, can be functionally connected to neighboring neurons via junctions called synapses; the potential number and complexity of connective patterns that can be built up between them is therefore practically unlimited. It is out of this universe of possibility, says HAYEK, that the order we call 'the mind' emerges.

With respect to the formation of the mind, HAYEK contends that the sensory experiences the brain processes are not unitary, but entail a *collection* of impulses. That is to say, like a suitcase filled with an assortment of shirts, pants, shoes, belts, socks, etc., sensory experiences are made up of many impulses corresponding to various aspects of the observed object or event. What is more, these impulses emanate not from one, but from many neighboring receptors in the sensory organ, and they occur in conjunction with still other impulses associated with participation in a specific kinesthetic activity—such as touching, looking, or listening. This package of impulses then courses through our nervous system and, through what HAYEK calls "physiological memory" (ibid., p53), forges connective pathways or "links" (ibid., p104) between neurons. Such connections are formed, says HAYEK, because the electrochemical impulses triggered by sensory stimuli change the "threshold of excitation" (1978c, p40) of affected neurons so that future impulses are 'positively weighted' or flow more easily through those already in "a state of preparedness to act" (ibid.). This view is not without some basis in modern neuroscience. Neuroscientists maintain that sensory experiences, especially recurrent or traumatic experiences, generate connections between neurons. What occurs is a physiological process called long-term potentiation, or LTP (BAUDRY and DAVIS, 1996). The LTP process involves changing the efficiency of synaptic transmissions along pathways that connect neurons—in other words, certain electrochemical signals travel more easily along LTP pathways. According to this theory, the connective pathways between neurons possess a class of postsynaptic amino acid receptors known as NMDAs. NMDA receptors are activated each time the pathway is confronted with an electrochemical impulse so that the receptivity of neurons with worn NMDAs is enhanced over time.

As there are a multiplicity of impulses associated with each sensory experience we encounter, impulses from different sensory experiences may employ one or more of the same neural pathways. There will occur, in other words, an *overlapping* of physiological memory. This overlapping is perhaps the single most important concept of HAYEK's theory of the mind for it leads to what he calls "simultaneous classification" (1952, pp180–181). Simultaneous classification is the idea that sensory experiences are *at the same time* related to all sorts of other sensory experiences via shared neural pathways. These shared pathways have the effect of grouping together or categorizing sensory experiences along the lines of a neural commonalty. Returning to our suitcase metaphor for sensory experiences, it would be as if one containing shirts, pants, and shoes was linked to all the others with shirts, and *at the same time* linked to the ones with pants and, *still further*, linked to all the ones with shoes. The concept of simultaneous classification, in other words, means that at any given moment a sensory experience will be a member of *more than one class of events*, related through physiological memory to many other sensory experiences.

HAYEK also contends that there are connections that 'negatively weight' or inhibit the flow of impulses (1952, pp67–68). This not only defines a second way that the brain's electrochemical signals are channeled, but compounds the complexity of neural patterns by introducing the possibility that different impulses can create connections that oppose or counteract each other.

Under this connectionist[2] model of brain functioning, the "possibilities of classification of...different individual impulses and groups of impulses...are practically unlimited [and] adequate for building up an extremely complex system of relations among millions of impulses" (ibid., p71). As HAYEK sees it, initial sensory impulses destined for the brain "pass in a great variety of directions...merely diffus[ing] and dissipat[ing] themselves in our neural fibers" (ibid., p120). An afferent impulse arriving for the first time, in other words, will "not yet occupy a definite position in the order of such impulses," or have a "distinct functional significance" (ibid., p103). "But since every occurrence of a combination of such impulses will contribute to the gradual formation of a network of connexions of ever-increasing density, every neuron will gradually acquire a more and more clearly defined place in the comprehensive system of such connexions" (ibid., p103). It is out of this 'thickness' of connections that we are able to detect patterns and come to know the world. In fact, says HAYEK, one of the things that distinguishes an adult mind from an infant mind is that an infant has a "much thinner net of ordering relations" (1978c, p44). Thus our experience is "richer than theirs as a consequence of our mind being equipped, not with relations which are more abstract, but with a greater number of abstract relations" (1989, p66).

This intimate relationship between neural connections and sensory experiences leads HAYEK to what is called the correspondence theory of perception—the idea that the physical workings of the brain come to map things out in the world. As HAYEK puts it, the "mental order involves...a gradual approximation to the order which in the external world exists between the stimuli evoking the impulses which 'represent' them in the central nervous system" (1952, p107). HAYEK, however, is careful to point out that our representations are not in some manner originally attached to, or an original attribute of, the individual physiological impulses or stimuli. Rather, the process of physiological memory *creates* the distinctions in question. Indeed, the representations are "determined by the system of connections by which the impulses can be transmitted from neuron to neuron" (ibid., p53). This may also be expressed as the specific character of a particular representation is "neither due to the attributes of the stimulus which caused it, nor to the attributes of the impulse, but [is] determined by the position in the structure of the nervous system of the fiber which carries the impulse" (ibid., p12). In other words, a given sensory impulse does not in and of itself designate specific mental representations. Rather, a mental representation is designated by the order of *all the connections* established in the mind.

HAYEK's connectionism, therefore, leads him to assert that there is no basis to believe that the representation of physical reality that the mind makes possible is a complete representation of the world *Ding an sich*. Rather, each mind functions through a recognition of what is similar to that mind at the expense of what is particular to an item. "What we perceive of the external world," explains HAYEK, "are never all of the properties which a particular object can be said to possess objectively, not even only some of the properties which these objects do in fact possess physically, but always only certain 'aspects,' relations to other kinds of objects which we assign to all elements of the classes in which we place the perceived objects. This may often comprise relations which objectively do not at all belong to the particular object but which we merely ascribe to it as a member of the class in which we place it as a result of some accidental collection of circumstances from the past." (ibid., p143).

In other words, HAYEK's connectionist mind is not a strict catalogue of empirical data, but an extracted collection of similarities or analogies. As Anna GALEOTTI correctly summarizes his view, the mind does not know specific things, but kinds (1987, p170).[4]

But how is it that sensory impulses come to contribute to the gradual formation of an order of connections, especially one that is capable of distinguishing things in life's storm of sensory events? Or, as HAYEK puts it, "the question which thus arises for us is how it is possible to construct from the known elements of the neural system a structure which would be capable of performing such discrimination in its responses to stimuli as we know our mind in fact to perform" (1952, p47). According to HAYEK, the answer to this question has to do with the process of evolution. This is a predictable starting point for HAYEK given that he asserts that "wherever we look, we discover evolutionary processes leading to...increasing complexity," and moreover, "we understand now that all enduring structures above the level of the simplest atoms, and up to the brain and society, are the results of, and can be explained only in terms of, processes of selective evolution" (1989, p92).

The Evolutionary Mind

Once a 'thick' net of ordering connections is established in the mind, says HAYEK, a range of possible neural routing patterns is engendered. Simultaneous classification, in other words, results in "a process of channeling, or switching, or 'gating' of the nervous impulses" (1967, p51). Yet HAYEK is emphatic that this 'lock-and-dam' system of neural connections does not in and of itself specify the neural routing patterns that will be employed by the mind. Instead, neural connections constitute "dispositions" (1978c, 40) and only through competition among many different neural dispositions and combinations of dispositions will distinctly functional patterns be discovered.

HAYEK thus embraces the view that the physiological apparatus that enables us to know the world is itself subject to the pressures of the natural selection process.[5] As he explains it, the brain "first develops new potentialities for actions and that only afterwards does experience select...those which are useful as adaptations to typical characteristics of its environment" (ibid., p42). In other words, the mind "simultaneously plays with a great many action patterns of which some are confirmed and retained as conducive to [its] preservation" (ibid., p43). The neural patterns produced in the structure of the nervous system "will first appear experimentally and then either be retained or abandoned" (ibid.). Since the "chance of persistence" of the mind is evidently increased if it possesses the capacity of "retaining a 'memory' of the connexions between events" that are capable of "correct[ly] anticipat[ing] future events" (1952, p129), there will emerge from natural selection among the brain's changing repertoire of neural dispositions and combinations of dispositions, patterns that conform to the requirements of survival.

This evolutionary model of the mind's operation is analogous to the one employed in Oliver Selfridge's artificial intelligence computer program, *Pandemonium*, or 'many demons.' That program contained numerous semi-independent sub-programs, or demons. When problems were encountered, all the demons would compete, and after a brief struggle, the winner would get to try to solve the problem. If it failed, others would try until a demon was found that allowed the overall program to continue to operate. Later *Pandemonium*-type programs involved random connections between demons so that they could build on each other and experiment with more sophisticated solutions. The longer the demons continued to function, the stronger the bond or connective confidence between them would grow. According to Daniel DENNETT (1995), director of the Center for Cognitive Science at Tufts University, this selection process may appear disorganized and inefficient with "all these different demons working on their own little projects...building things and then...tear[ing] them apart (ibid., p183). But, "its also a great way of getting something really good built—to have lots of building going on in a semicontrolled way and then have a competition to see which one makes it through to the finals" (ibid.).

Yet we should realize that such a process of natural selection is useful insofar as the relevant circumstances for survival are *unknown*. Indeed, it would be pointless to employ the natural selection process if it could be determined beforehand what a successful outcome would be. In this sense, natural selection is practical precisely because viable results cannot be precalculated. Instead, they are "discovered" (1984a, p255) through a trial-and-error procedure whereby unsuccessful solutions (or neural dispositions) are eliminated. What remains after the procedure, according to HAYEK, is a form of "knowledge" (1984a, p257), a kind of residue of information on how to survive. This "knowledge," however, is of a negative or POPPERIAN (1963) sort; that is, learning to meet the requirements of survival through natural selection does *not* consist of "verifying" solutions, but of "falsifying" unfit alternatives.

It is also important to remember that the "knowledge" generated by this natural selection process cannot be called intentional. Although it may be highly conducive to survival, it does not itself have that aim. Rather, it is passively acquired through the ordeal of trial and error. It is on this point that HAYEK's epistemology most clearly mirrors Samuel Pufendorf's algedonic notion of "implicit obligation" (BUCKLE 1991, pp63–64). According to PUFENDORF, an early natural rights theorist, we can speak of two kinds of obligation: "explicit obligation," where the need for an action is secured through its self-evident beneficence, and "implicit obligation," where the need for an action is *not* apparent, but secured through the guiding pressures of harmful actions. With respect to HAYEK's theory of the mind, this means that since it is constantly exposed to newly arriving sensory signals and its "persistence...will...be increased if it...happens to respond appropriately to harmful and beneficial influences" (1952, p129), its "actions will appear self-adaptive and purposive" (ibid., p122). The crucial point here, however, is that the "knowledge" that enables the mind to persist is "not built up by [itself], but that it is by a selection among mechanisms producing different patterns that the system...is built up" (1978c, pp42–43). In other words, says HAYEK, the mind's evolution is "blind" (1988, p15). It depends *not* upon premeditated objectives or foresight, but upon a "process of exploration" (1984a, p263) or a "discovery procedure" (ibid., p255) that gropes through the space of what is possible and happens upon routines that fit the requirements of survival. Thus understood, the natural selection process is a nonteleological explanation of the mind's manifestly practical achievements.

HAYEK's recognition that each person's mind is made up of a blind accumulation of useful responses to the demands of survival is critical to his evolutionary epistemology; it allows him to dispense with three properties commonly misattributed to the evolutionary process: optimalism, progressivism, and sequentialism.

Optimalism

HAYEK points out that although blind selection makes possible the learning necessary for survival, it does not mean that permanent or absolute solutions will be discovered. It simply means that of the solutions available at a specific moment in time, the instrumental one(s) will survive, and at the same time, that any solution may diminish survivability in another context or future scenario. In other words, it is a fundamental mistake to view natural selection as a discovery process that necessarily supplies the optimal answers *through time*. Any evolutionary adaptation selects for some particular attribute over others, thus entailing a trade-off or opportunity cost. They are trade-offs because they close-off future lines of development and lock-in traits that may be maladaptive in future environments. Accordingly, it would be wrong to conclude, starting from evolutionary premises, that whatever solution has evolved is "always or necessarily conducive to the survival or increase in the populations following them" (1988, p20).

Progressivism

HAYEK maintains that we cannot describe evolution as a phenomenon forever progressing forward and upward. Indeed, 'progress' in evolutionary terms merely means adaptation to a changing environmental context and what that entails—success and failure, persistence and elimination. That solutions become more complex and better adjusted to generate survival, explains HAYEK, happens not because they are approaching a superior end state, but because those prospered that happened to change in ways that made them increasingly adaptive. Indeed, "all evolution…is a process of continuous adaptation to unforeseeable events, to contingent circumstances which could not have been forecast" (ibid., p25). As in the case of biological evolution, natural selection "describes a kind of process (or mechanism) which is independent of the particular circumstances in which it has taken place on Earth, which is equally applicable to a course of events in very different circumstances, and which might result in the production of an entirely different set of organisms" (1967c, p32). Thus, "evolution [is] not linear, but result[s] from continual trial and error, constant 'experimentation' in arenas wherein different orders contend" (1988, p20).

Sequentialism

HAYEK rejects the idea that evolution must follow a set sequence of phases. Indeed, "although…the original meaning of the term 'evolution' refers to such an 'unwinding' of potentialities already contained in the germ, the process by which the biological…theory of evolution accounts for the appearance of different complex structures does not imply such a succession of particular steps" (1973, p24). That is to say, the concept of evolution does not denote "necessary sequences of predetermined 'stages' or 'phases,' through which the development of an organism…must pass" (ibid.), and evolution does not know anything like "'laws of evolution' or 'inevitable laws of historical development' in the sense of laws governing necessary stages or phases through which the products of evolution must pass" (1988, p26). "One of the main sources of this misunderstanding," says HAYEK, "results from confusing two wholly different processes which biologists distinguish as *ontogenetic* and *phylogenetic*" (ibid.). Ontogenesis has to do with the predetermined development of an individual thing, something set by the inherent mechanisms built into it, such as the way DNA determines our physical make-up. By contrast, phylogenesis deals with the evolutionary movements of whole populations, such as the way giraffes acquired long necks as a result of an extended shortage in ground vegetation. Non-biologists tend to make the error that phylogenesis operates in the same closed way as does ontogenesis. It does not. Phylogenesis is not preprogrammed, but dependent upon the changing requirements of survival. Indeed, as in the case of the mind, a phylogenetic system is conditional; that is, it is reflective of the demands of the environment *at that point in time*.

The most significant implication of the blindness of natural selection, however, is that survival is based on chance, not foredesign. On this point, HAYEK explicitly recognizes the practical value of error toleration. Indeed, as has been pointed out by complexity theorists Gregoire NICOLIS and Ilya PRIGOGINE (1989) in their discussion of the adaptability of ant colonies: "A permanent structure in an unpredictable environment may well compromise the ability of the colony and bring it to a suboptimal regime. A possible reaction toward such an environment is thus to maintain a high rate of exploration and the ability to rapidly develop temporary structures suitable for taking advantage of any favorable occasion that might arise. In other words, it would seem that randomness presents an adaptive value in the organization of the society" (ibid., p293). Similarly, for HAYEK, random "mutations" (1973, p9) and "historical accidents" (1988, p20) are the raw material of the mind's evolution, and it is a capacity to generate and accumulate variations that makes possible the adaptive learning and innovations that are necessary to accommodate the open-ended problem of survival.

It is on this point that HAYEK's epistemology most clearly moves in the direction of Marvin MINSKY's (1995) approach to machine thinking. According to MINSKY, co-founder of MIT's Artificial Intelligence Laboratory, the "wrongheadedness" (ibid., p163) of early artificial intelligence research was its emphasis on preprogramming the best strategies for dealing with particular situations. But what MINSKY and others realized is that one does not understand something unless one understands it in "several different ways" (ibid.) at the same time. That is to say, "if you understand something in just one way, and the world changes a little bit and that way no longer works, you're stuck, you have nowhere to go. But if you have three or four ways of representing the thing, then it would be very hard to find an environmental change that would knock them all up" (ibid.). To then rely one kind of strategy is to invite cognitive paralysis. The 'trick' to productively interacting with the world, therefore, is to "accumulate different viewpoints" (ibid.) so that there are alternatives standing by when one fails.

This essentially is HAYEK's view. He recognizes that "the immediate effects of…conflicting experiences will be to introduce inconsistent elements into the model of the external world; and such inconsistencies can be eliminated only if what formerly were treated as elements of the same class are treated as elements of different classes" (1952, p169). In other words, neural patterns based on past neural connections do "not always work," and such events force the mind to adapt its approach to incorporate novel experiences (ibid., p168). According to HAYEK, such adaptation can occur because the mind has "a large repertoire of…patterns…provid[ing] the master moulds (templates, schemata, or *Schablonen*) in terms of which will be perceived many other complex phenomena in addition to those from which the patterns are derived" (1967b, p51). In short, we are never of only one mind. We have numerous 'recruits' ready to perform. Accordingly, the mind assimilates inconsistent impulses when it eliminates ineffective 'recruits' and conforms to ones that are more instrumentally fit. The mind is thus "autoepistemic," learning through the default logic of the natural selection process or what artificial intelligence researchers call "nonmonotonic reasoning" (ANTONIOU 1996).

Robert DEVRIES (1994) provides an instructive metaphor for this default process. Suppose that someone with little formal education, but plenty of curiosity, began to contemplate the linear order of sentences. This hypothetical person develops a list of various rules to explain how sentences can be transformed into questions. He then notices the following regularities:

John has called his sister.
Peter can buy a bicycle.
People won't die.
and
Has John called his sister?
Can Peter buy a bicycle?
Won't people die?

What does our aspiring linguist do? He uses the following rule: statements can be transformed into questions by reversing the order of the first two words of a sentence. But then our would-be linguist encounters some new experiences that upset his rule.

The big house is cheap.
People without lungs will die.
and
Is the big house cheap?
Will people without lungs die?

If his rule were applied to the above questions they would have the following syntactic form:

Big the house is cheap?
Without people lungs will die?

As these questions are meaningless, our linguist's first rule—although previously effective—no longer accounts for his experiences; that is, the first and second words of a sentence are shown *not* to be the fundamental elements of the grammatical world. When this refutation occurs, our linguist is compelled to employ an alternate rule to adapt to his new experiences. He may then default to a second rule, one that states that sentences can be transformed into questions by reversing the order of the object and the finite verb. This second rule has greater adaptive advantage; not only does it explain his new experiences, *but it would have explained his old experiences had he employed it before*. In short, the second rule is more instrumentally fit than the first.

Similarly, HAYEK conceives of the mind's successful neural patterns and combinations of patterns as "rules" (1967a, p67). But unlike our metaphor, he does not use this term to mean explicit instructions. For HAYEK, neural rules comprise instructions of an *unarticulated* sort. They consist, instead, of an implicit "capacity" (1984a, p257) to effectively survive in a given environment; a "knowing how" rather than a "knowing that," to borrow from Gilbert RYLE (1945–6). The rules HAYEK speaks of thus do not imply an awareness of purpose on the part of the mind but merely that it embodies "regularities of conduct" (1967a, p67) conducive to its maintenance—presumably because those that operate in certain ways have a better chance for survival than those that do not. In other words, minds are successful because they adapt to facts that are *not* known, and this success is brought about by the discovery of neural patterns that are not designed, but followed in action. Or, to put it differently, the mind does not consist primarily of insight into the relationship between resources and objectives, but of the blind selection of rules that enable it to endure. Indeed, says HAYEK, trial-and-error learning "is a process not primarily of reasoning but of...the development of practices which have prevailed because they were successful...[and] the result of this development will...not be articulated knowledge but a knowledge which...cannot [be] stat[ed] in words but merely...honor[ed] in practice" (1973, p18). Our evolutionary mind, therefore, must be conceived of as having a meta-structure, of "being guided by rules of which we are not conscious but which in their joint influence enable us to exercise extremely complicated skills without having any idea of the particular sequence of movements involved" (1978c, p38).

This ignorance of the particular sequence of movements involved in the mind's evolution exemplifies what HAYEK calls the "primacy of the abstract" (ibid., p35). The primacy that HAYEK is concerned with is chronological. He contends that the concrete particulars we experience "are the product of abstractions which the mind must possess in order that it should be able to experience particular sensations, perceptions, or images" (ibid., pp36–37). Thus, by abstraction HAYEK does not mean something complicated, but "presuppositions" (1989, p63). That is to say, the mind depends upon, or is secondary to, the unintended discovery of regularities of conduct that are obeyed in practice, but which are not deliberately devised. "The formation of abstractions," he says, "ought to be regarded not as actions of the human mind but rather as something which happens to the mind" (1978c, p43); i.e., the formation of a rule seems "never to be the outcome of a conscious process, not something at which the mind can deliberately aim, but always a discovery of something which already guides its operation" (ibid.).

HAYEK insists, however, that just because the evolutionary mind "does not so much make rules as *consist* of rules," (1973, p18) it does not follow that its operation should be characterized as 'sub' conscious. HAYEK puts this point clearly when he explains that it "is generally taken for granted that in some sense conscious experience constitutes the 'highest' level in the hierarchy of mental events, and that what is not conscious has remained 'subconscious' because it has not yet risen to that level" (1978c, p45). And, HAYEK does not doubt that many mental processes through which stimuli evoke actions do not become conscious because they proceed literally on too low a level, "but this is no justification for assuming that all the [cognitive] events determining action to which no distinct conscious experience corresponds are in this sense subconscious" (ibid.). Instead, HAYEK maintains that if his conception is correct, then processes of "which we are not even aware determine the sensory qualities which we consciously experience, [and] this would mean that of much that happens in our mind we are not aware, not because it proceeds at too low a level but because it proceeds at too high a level. It would seem more appropriate to call such processes not 'subconscious', but 'superconscious', because they govern the conscious process without appearing in them. This would mean that what we consciously experience is only part, or the result, of processes of which we cannot be conscious." (ibid.).

Thus, to paraphrase Carl Gustav JUNG (CALVIN 1996), just as we are not able to see the stars that are above during the day because the sun is too bright, unconscious cognitive activity is going on all the time—we simply cannot discern it through the medium of consciousness. If, on the other hand, we could know our mind's evolutionary processes, we could know the rules upon which our connectionist mind is based. But this, he notes, is impossible; we cannot self-consciously know the evolutionary activity to which all our conscious thoughts necessarily refer. In order to describe such knowledge, we would need to know how it is conditioned and determined. But in order to describe *this* knowledge, we would need to possess additional knowledge on how *it* is conditioned and determined, and so on *ad infinitum*. Such a mind would soon find itself locked in a perpetual cycle of introspective analysis analogous to what computer scientists call an 'infinite loop' error. "The whole idea of the mind explaining itself is [thus] a logical contradiction" (1952, p192).[6]

Instead, HAYEK conceives of the mind as an emergent system or "spontaneous order" (1978a, p250), a concept that forms the basis of his argument against the "constructivistic" (1978b, p3) fallacy that evolutionary phenomena, like "life, mind, and society" (1973, p41), must be centrally ordered. This fallacy is a variation of William PALEY's "argument from design" (POPPER 1987, p13), an argument that holds that if you find a complex structure, like a watch, you cannot doubt that it was designed by a watchmaker.

So when you consider a complex structure, like the mind, you are bound to conclude that it must be designed by a directing consciousness.

HAYEK vigorously disputes the argument from design. Throughout his writings, he makes the distinction "between an order which is brought about by the direction of a central organ...and the formation of an order determined by the regularity of actions toward each other of the elements of a structure" (1967a, p73). The former is a designed order, the later is a "spontaneous order." HAYEK further uses Michael POLANYI's notion of "monocentric" and "polycentric" (ibid.) orders to clarify this distinction: A monocentric order is organized by a directing core, a polycentric order, on the other hand, emerges out of "the relation and mutual adjustments to each other of the elements of which it consists" (ibid.). A polycentric order, in other words, is a "self-organizing" (1984a, p259) system; a system that "dispenses with the necessity of first communicating all the information on which its several elements act to a common centre" (1967a, p74) and operates, instead, through the trial-and-error interaction of many parts. In the case of the mind, this process gradually builds up a mental "geography" (1952, p109) or "sensory order" (1952, *passim*) suitable for effectively navigating the external world, or as artificial intelligence theorist Teuvo KOHONEN (1995) might put it, the mind is a "self-organizing map" derived from the "unsupervised learning" of natural selection.

Yet this is not to say that the mind is merely the sum of its parts. HAYEK rejects the idea of "one-directional laws of cause and effect" and supports the idea of "downward causation" (1989, p93). Downward causation is the argument that the mind operates like a feedback loop in which the whole is constrained by the micro-level activity, and the micro-level activity is constrained by the whole. Or, to put it another way, the local interactions of the parts give rise to a collective pattern or global dynamic and, in turn, this global dynamic sets the context in which the local interactions occur. It is important to remember, however, that this is not an argument against materialism; downward causation does not break the causal chain, but simply turns the chain inward on itself.

Before we turn our attention to how this account of the mind's operation undermines the idea of free will, two additional points bear noting. First, HAYEK makes it clear that the discovery of neural rules is *not* going on in only one mind, but in everyone's mind, and that the discoveries made in one mind can "infect" (1967b, p47) other minds through speech and example. As such, he argues that humans are intelligent, in part, because neural rules can be accumulated and transmitted from person to person, generation to generation. "What we call the mind," says HAYEK, "is not something that the individual was born with...but something his genetic equipment helps him acquire, as he grows up...by absorbing the results of a tradition that is not genetically transmitted" (1988, p22).[3] In other words, language, morals, law, etc., are not discovered *ex nihilo* by each mind, but simply constitute an epidemic of "imitation" (ibid., 24), of successful neural rules combining and spreading through populations. Under this view, "learning how to behave is more the source than the result of insight, reason, and understanding" (ibid., p21) and "it may well be asked whether an individual who did not have the opportunity to tap such a cultural tradition could be said even to have a mind" (ibid., p24).

Second, we must consider what Elliott SOBER (1994) calls the "problem of the units of selection" (ibid., pxii). According to SOBER, the question we should ask in cases of evolutionary phenomena is: "What kinds of objects should we regard as the relevant beneficiaries" of natural selection (ibid.)? For example, did opposable thumbs evolve because they helped guarantee the transmission of the genes of the carrier organism, or because they helped individual creatures to survive, or because they helped the whole species avoid extinction (ibid.)? With regard to HAYEK's evolutionary epistemology, our attention has focused primarily on the trial-and-error discovery of useful neural rules. But what is the relevant unit of selection? I.e., what is it that is really being selected? Is it the particular neural pattern, or the individual whose mind makes the useful discovery, or the cultural group whose survival is improved by the spread of useful knowledge? Moreover, should our approach to this question be "variational" or "developmental" (SOBER 1993)? A variational account refers to how, given a population of differentiated units, only those with certain qualities will survive. A developmental account refers to how individual units react to change and that only those adapting to the new situation will survive. The success of neural rules conforms to the variational account, while the success of the overall order of the mind conforms to the developmental account. At the same time, the success of a particular individual depends upon their cognitive attributes (a variational account), and the success of a culture depends upon its capacity to adapt to changing conditions (a developmental account). According to HAYEK, these morphological issues are not easily unraveled because the mind—besides being subject to pressures of genetic and biological selection—is embedded in three other levels of selection: neural, individual, and cultural (HERRMANN-PILLATH 1994).

The Implication for Free Will and HAYEK's Response

HAYEK's view that the mind is an emergent system or "spontaneous order" holds a significant implication for the age-old controversy about free will—defined as a will that is *not* the exclusive and necessary result of the interaction of physical material. As far as we have seen, the mind consists of matter and its relations, and since everything can be realized in these materialist terms, there is simply no room for freedom of will. Indeed, it is another way of saying that our choices, judgments, and decisions are *determined* by the operation of the material that constitutes ourselves and the world, or as Oxford scholar John GRAY summarizes HAYEK's view, "our ideas are merely the visible exfoliation of spontaneous forces" (1986, p30). But if this account is correct, why should we do anything purposeful at all? Doesn't HAYEK's materialism destroy the idea of goal-oriented action?

Not so fast, responds HAYEK; we can never introspectively predict how our mind is to be determined. Instead, "we can know [our mind] only through directly experiencing it" (1952, p194). With regard to the issue of purposeful action, then, HAYEK makes it clear that his materialism makes no *practical* difference in our daily lives; we must still conduct ourselves as if we are free because we can never know how we are meant to behave. Indeed, "we may...well be able to establish that every single action of a human being is the necessary result of the inherited structure of his body (particularly of its nervous system) and of all the external influences which have acted upon it since birth. We might be able to go further and assert that if the most important of these factors were in a particular case very much the same as with most other individuals, a particular class of influences will have a certain kind of effect. But this would be an empirical generalization based on a *ceteris paribus* assumption which we could not verify in the particular instance. The chief fact would continue to be, in spite of our knowledge of the principle on which the human mind works, that we should not be able to state the full set of particular facts which brought it about that the individual did a particular thing at a particular time." (1989, pp86–87).

HAYEK thus salvages the idea of purposeful action from the grips of materialism by maintaining that we cannot avoid acting as if we are free because we are never in a position to know how we are determined to behave. In other words, HAYEK does not assert that our will is free, but that *we are incapable of knowing how to behave like our will is unfree*.

In order to gain a fuller understanding of this argument, we must begin with the recognition that HAYEK is a materialist without being a reductionist. Or as he puts it, "those whom it pleases may express this by saying that in some ultimate sense mental phenomena are 'nothing but' physical processes; this, however, does not alter the fact that in discussing mental processes we will never be able to dispense with the use of mental terms [for] we shall never be able to explain [them] in terms of physical laws" (1952, p191). Our minds, he contends, "must remain irreducible entities" (ibid.).

A primary obstacle to reduction, says HAYEK, stems from the mind's interconnectivity. This occurs because the mind's elements—sensory experiences—are linked to one another in such a way that they actually determine what the others are through their interconnections. The mind, in other words, is a quality of arrangements; "its actions are determined by the relation and mutual adjustment to each other of the [multiple] elements of which it consists" (1967a, p73) and the multitude of connections "proceeding at any one moment, can mutually influence each other" (1952, p112). Thus, adding or removing even one sensory experience will change all the others in some subtle way.

The practical implication of such interconnectivity is that a sensory experience cannot be analyzed without reference to the *other* sensory experiences that a mind has encountered; that is, in order to describe a sensory experience *all the way through*, one must describe its relations to other experiences, which are, in turn, are related to still other experiences, and so on in an infinite regress. Logically, any attempt to describe precisely a sensory experience would have to take into account the complete order that emerges from a person's previous sensory experiences. As a result, the mind cannot be broken down into linear, A causes B terminology and reassembled into an explanation of the whole. No sensory experience is autonomous. Rather, all sensory experiences are embedded in complex relations with other sensory experiences. The relations change the experience so that it constitutes more than itself. It "resonates" with what Jaques DERRIDA (1976) might call "traces" of something "other." Consequently, where one experience ends and another begins is undecidable. There are only sensory experiences *in relations to* other sensory experiences; there essence lies in their relations to the others and their effects on the same.

Another obstacle to reduction, says HAYEK, has to do with the mind's dynamic quality. This occurs because the order of the mind is constantly being updated; that is, when the mind encounters a new bit of sensory data, it is itself altered by that data—it recontextualizes. The mind, in other words, successively publishes revised editions that incorporate the immediately preceding sensory experience. What results, to paraphrase HAYEK's student Ludwig LACHMANN (GARRISON 1987), is a "kaleidic" process in which the order that we call the mind is continually cascading into new and novel patterns.

Given this, the mind is not a closed system. A closed system is like a finite collection of musical notes, where the possible patterns that can be played today are identical to the possible patterns that can be played next week, next year, or next century. But what happens when the unity of the system is broken and a new note is introduced? The whole nature of possible permutations changes. No possible permutation of the former set of notes can replicate a sequence containing the new note. The introduction of a new note, therefore, dramatically changes the possible outcome of *all* future scenarios.

Similarly, the introduction of a new sensory experience alters the mind's possible future scenarios. That is, each new sensory experience one witnesses will be interpreted within the context of an updated network of neural connections, one that incorporates the immediately preceding sensory information. As a consequence, each contemplation is unique or, as Heraclitus might have put it, you cannot step into the same stream of thought twice. The order of the connections in the mind, explains HAYEK, "is modified by every new action exercised upon it by the external world, and since the stimuli acting on it do not operate by themselves but always in conjunction with the process called forth by the preexisting excitatory state, it is obvious that the response to a given combination of stimuli on two different occasions is not likely to be exactly the same. Because it is the whole history of the organism which will determine its action, new factors will contribute to this determination on the latter occasion which were not present in the first. We shall find not only that the same set of external stimuli will not always produce the same responses, but also that altogether new responses will occur." (1952, p123).

What this suggests is that even if we could know the precise order and intensity of new experiences, this would not enable us to explain why a mind responds the way it does. The reason for this is the actual impossibility of ascertaining the particular circumstances which, in the course of a lifetime of experiences, have decided the emergence and trajectory of the complex order that we call the mind. In other words, the mind is biographical, and its manifestation is dependent upon a staggeringly long and statistically unrepeatable sequence of variables and intensities. Indeed, to paraphrase paleobiologist Stephen J. GOULD (1989), wind back the tape of the mind to its early days; let it play again from an identical starting point, and the chance becomes vanishingly small that anything like the identical mind will grace the replay (ibid. p14). Subsequently, a more appropriate question to ask than Thomas NAGEL's (1974) famous "what is it like to be a bat?" is "what is it like to be another person?" Since each mind is historically fingerprinted, this cannot be known. An identical sensory experience would require "an identical history"—a requirement that ultimately "precludes the possibility that at any moment the maps [or minds] of two individuals should be completely identical" (1952, p110). Thus, although people can refer to the same sensory experience, it neither follows that it has the same location or intensity in their evolutionary mind, nor that all the connections that extend from it are the same. Each experience is, in this sense, 'private'—just as there are no two identical snowflakes, there are no two identical sensory experiences of a snowflake.

In conclusion, it should not be difficult now to recognize that although HAYEK rejects the idea of free will, he accepts the idea of a subjective will; that is, a willfulness *unique* to each individual. It should also not be difficult to recognize the predictive limitations applying to explanations of such a will. In fact, HAYEK rejects the possibility of "specific prediction" in the case of the individual will and finds that such a goal is "completely unjustified" (1989, p88). He maintains, rather, that specific prediction of the will could "be achieved only if we were able to substitute for a description of events in...mental terms a description in physical terms which included an exhaustive enumeration of all

Author's address

Gary T. Dempsey, Cato Institute, 1000 Massachusetts Avenue, N.W., Washington, D.C. 20001, U.S.A.
Email: gdempsey@cato.org

the physical circumstances which constitute a necessary and sufficient condition of the...mental phenomena in question" (ibid.). But, as has been argued, viewing the mind as a "spontaneous order" creates an "impossibility of ascertaining all the particular data required to derive detailed conclusions" (ibid., p86). As a result, says HAYEK, "the individual personality [will] remain for us as much a unique and unaccountable phenomenon...but whose specific actions we [can] generally not predict or control, because we [can] not obtain the information on all the particular facts which determined it" (ibid., pp86–87). In other words, even though we may know the general principle by which the emergent system we call the mind is causally determined by evolutionary processes, this does not mean that a particular human action can ever be introspectively recognized as the necessary result of a particular set of facts. Indeed, HAYEK maintains that we are in no better position to predict the specific future motions of our mind than we are "able to predict the shape and movement of [a] wave that will form on the [surface of the] ocean at a particular place and moment in time" (1984, p243). Returning to the topic of artificial intelligence, this raises an important closing observation. If the same quality of irreducibility applies to intelligent machines, then they too will face limits to introspection. But more significantly, it will also mean that we will be incapable of recursively describing their evolved will; that is, *we won't be able to tell from their operation the precise sequence and relation of events that contributed to their specific manifestation*. As a result, if HAYEK's epistemological insights hold for artificially intelligent machines, we can already recognize an imminent limitation on our ability to predict and/or plan their behavior. We shall see.

Notes

1 The author recently received his M.A. in political theory from the College of William and Mary and has been published in *The Southern Journal of Philosophy*. He wishes to thank Michael GIBERSON and an anonymous referee for their helpful comments and criticisms, to which the standard disclaimer applies. He also wishes to especially thank Manfred WIMMER for his assistance and indefatigable patience.

2 This is a term that Barry SMITH (1996) also applies to HAYEK's theory of mind. What I (and Smith) mean to suggest by using this term is that there is an affinity between HAYEK's views and those of the research field pioneered by Warren MCCULLOCH and Walter PITTS (1943) that today encompasses artificial neural networks. With respect to the development of this field, it should be noted that one of the first connectionist computer models of the brain, Frank ROSENBLATT's "Perceptron," was directly inspired by the writings of HAYEK and psychobiologist Donald O. HEBB (ROSENBLATT 1958).

3 This 'memetic' view not only provides a mechanism for minds to limit/correct catastrophic errors in society, but for society to limit/correct catastrophic errors in the mind. In other words, the system that we call the mind can act to 'regulate' the system of society and the system of society can act to 'regulate' the system that we call the mind. Moreover, when these systems are taken as one, this larger system can act in a 'self-regulating' way.

4 For further discussion of HAYEK's theory of mind, see WEIMER (1982), JONKER (1991), HERRMANN-PILLATH (1992), STREIT (1993), and BIRNER (1995).

5 This view should not sound unusual to readers acquainted with Gerald M. EDELMAN's writings, and it bears noting that EDELMAN and HAYEK were familiar with each other's work. In fact, EDELMAN cites HAYEK in his 1987 book, *Neural Darwinism*, and HAYEK cites Edelman in his 1988 book, *The Fatal Conceit*.

6 Elsewhere, HAYEK employs Kurt GÖDEL's incompleteness theorem to make the point that we cannot know our own mind because the 'set' that we call the mind cannot logically contain itself (1967b, p62).

References

Antoniou, G. (1996) Nonmonotonic Reasoning. MIT Press: Cambridge.

Baudry, M. and Davis, J. (1996), Long-Term Potentiation, Vol. 3. MIT Press: Cambridge.

Birner, J. (1995) The Surprising Place of Cognitive Psychology in the Work of F. A. Hayek, unpublished paper. University of Maastricht, Department of Economics.

Buckle, S. (1991) Natural Law and the Theory of Property: Grotius to Hume. Clarendon Press: Oxford.

Calvin, W. (1996) The Cerebral Code. MIT Press: Cambridge.

Campbell, D. (1960) Blind Variation and Selective Retention in Creative Thought As in Other Knowledge Processes. Psychological Review, 67:380–400.

Campbell, D. (1974) Evolutionary Epistemology. In: Schilpp, P. (ed) The Philosophy of Karl Popper. Open Court: La Salle, Ill.

Dennett, D. (1995) Intuition Pumps. In Brockman, J. (ed) The Third Culture. Simon & Schuster: New York.

Derrida, J. (1976) On Grammatology. Johns Hopkins University Press: Baltimore.

DeVries, R. (1994) The Place of Hayek's Theory of Mind and Perception in the History of Philosophy and Psychology. In Birner, J. and Van Zijp, R. (eds) Hayek, Coordination, and Evolution: His Legacy in Philosophy, Politics, Economics, and the History of Ideas. Routledge: London.

Edelman, G. M. (1987) Neural Darwinism: The Theory of Neuronal Group Selection. Basic Books: New York.

Galeotti, A. (1987) Individualism, Social Rules, Tradition: The

Case of Friedrich Hayek. Political Theory. 15:163–181.
Garrison, R. (1987) The Kaleidic World of Ludwig Lachmann. Critical Theory, 1:77–90.
Gould, S. J. (1989) Wonderful Life. W. W. Norton: New York.
Gray, J. (1986) Hayek on Liberty, 2nd ed. Basil Blackwell: Oxford.
Hayek, F. (1952) The Sensory Order. The University of Chicago Press: Chicago.
Hayek, F. (1967a) Notes on the Evolution of Systems of Rules of Conduct. Studies in Philosophy, Politics, and Economics. The University of Chicago Press: Chicago.
Hayek, F. (1967b) Rules, Perception, and Intelligibility. Studies in Philosophy, Politics, and Economics. The University of Chicago Press: Chicago.
Hayek, F. (1967c) The Theory of Complex Phenomena. Studies in Philosophy, Politics, and Economics. The University of Chicago Press: Chicago.
Hayek, F. (1973) Law, Legislation, and Liberty: Rules and Order. The University of Chicago Press: Chicago.
Hayek, F. (1978a) Dr. Bernard Mandeville. New Studies in Philosophy, Politics, Economics and the History of Ideas. The University of Chicago Press: Chicago.
Hayek, F. (1978b) The Errors of Constructivism. New Studies in Philosophy, Politics, Economics and the History of Ideas. The University of Chicago Press, Chicago.
Hayek, F. (1978c) The Primacy of the Abstract. New Studies in Philosophy, Politics, Economics and the History of Ideas. The University of Chicago Press: Chicago.
Hayek, F. (1984a) Competition As a Discovery Procedure. In Nishiyama, C. and Leube, K. (eds) The Essence of Hayek. Stanford University Press: Stanford.
Hayek, F. (1984b) Philosophical Consequences. In Nishiyama, C. and Leube, K. (eds) The Essence of Hayek. Stanford University Press: Stanford.
Hayek, F. (1988) The Fatal Conceit, The University of Chicago Press: Chicago.
Hayek, F. (1989) Order: With or Without Design? The Center for Research Into Communist Economies: London.
Herrmann-Pillath, C. (1992) The Brain, Its Sensory Order, and the Evolutionary Concept of Mind: On Hayek's Contribution to Evolutionary Epistemology. Journal of Social and Evolutionary Systems, 15:145–186.
Herrmann-Pillath, C. (1994) Evolutionary Rationality, 'Homo Economicus', and the Foundations of Social Order. Journal of Social and Evolutionary Systems, 17:41–69.
Jonker, A. (1991) F. A. Hayek: The Sensory Order. Cognitive Science, 3(2):103–127.
Kohonen, T. (1995) Self-Organizing Maps, Vol. 30. Springer: New York.
Kukathas, C. (1989) Hayek and Modern Liberalism. Clarendon Press: Oxford.
Lorenz, K. (1977) Behind the Mirror. Methuen: London.
Lorenz, K. (1982) Kant's Doctrine of the A Priori in the Light of Contemporary Biology. In Plotkin, H. (ed) Learning, Development, and Culture. John Wiley & Sons: New York.
McColloch, W./Pitts, W. (1943) Logical Calculus of Ideas Immanent in Nervous Activity. Bull. Math. Biophysics, v. 5.
Minsky, M. (1995) Smart Machines. In Brockman, J. (ed) The Third Culture. Simon & Schuster: New York.
Nagel, T. (1974) What Is it Like to Be a Bat? Philosophical Review. 83:435–451.
Nicolis, G./Prigogine, I. (1989) Exploring Complexity. W. H. Freeman: New York.
Popper, K. (1963) Conjectures and Refutations: The Growth of Scientific Knowledge. Harper Tourchbooks: New York.
Popper, K. (1972) Objective Knowledge: An Evolutionary Approach. Clarendon Press: Oxford.
Popper, K. (1984) Evolutionary Epistemology. In Pollard, J. (ed) Evolutionary Theory: Paths into the Future. John Wiley & Sons: London.
Popper, K. (1987) Natural Selection and the Emergence of Mind. In Radnitzky G. and Bartley, W. (eds) Evolutionary Epistemology, Rationality, and the Sociology of Knowledge. Open Court; La Salle, Ill.
Rosenblatt, F. (1958) The Perceptron: A Probabalistic Model for Information Storage and Organization in the Brain. Psychological Review. 65:386–408.
Ryle, G. (1945–46) Knowing How and Knowing That. Proceedings of the Aristotelian Society. 46:1–16.
Smith, B. (1996) The Connectionist Mind: A Study of Hayekian Psychology. In Frowen, S. F. (ed) Hayek the Economist and Social Philosopher: A Critical Retrospect. Macmillan: London.
Sober, E. (1993) The Nature of Selection: Evolutionary Theory in a Philosophical Focus. The University of Chicago Press: Chicago.
Sober, E. (1994) Preface. In Sober, E. (ed) Conceptual Rules in Evolutionary Biology 2nd ed. MIT Press: Cambridge.
Streit, M. E. (1993) Cognition, Competition, and Catallaxy: In Memory of Friedrich August von Hayek. Constitutional Political Economy. 4:223–262.
Vanberg, V. (1994) Rules and Choices in Economics. Routledge: London.
Weimer W. (1982) Hayek's Approach to the Problems of Complex Phenomena: An Introduction to the Psychology of the Sensory Order. In Weimer W. and Palermo D. (eds) Cognition and the Symbolic Processes, Vol. 2. Lawrence Erlbaum: Hillsdale: N.J.

Wolfgang Wildgen

How to Naturalize Semantics (in the Spirit of Konrad Lorenz)?

Abstract

The naturalization of semantics is primarily a search for causes, forces and processes, which allow an understanding of what language is. It demands a transdisciplinary metalanguage, the acceptance of different levels of self-organization and insights into the biological and behavioural processes which control them. After a short characterization of three levels, the static and a priori concept of a 'universal grammar' is deconstructed and a naturalized alternative is deve-loped. It is first elaborated for the semantics of verbs of locomotion / control of objects / interaction and later expanded to the analysis of noun-phrases. The statistical dynamics of grammaticalization are modelled in the framework of linguistic synergetics. Finally, some hypotheses concerning the origins of language are proposed.

Key words

Naturalization, self-organization, semantics of verbs, language origin, valence

Semantics" asks the question how sign and meaning are correlated in humans (animals) and what the nature of meaning and its correlation with signs is. The verb "naturalize" refers to natural causes, i.e., causes originating in human nature or nature in general. The philosophical background of the question is given by the proposals for a naturalization of epistemology. How can we give a natural, explanatory account of true knowledge; the question of "real, true" meaning is a subpart of this epistemological question. The major candidates for a naturalization of epistemology have been psychology (which may again be naturalized by cognitive science, e.g., neurobiology) and evolutionary biology (based on DARWINIAN principles). These two programmes exhibit different time scales and therefore different types of causality:

- a statistical causality based on genetic mutation and the selection pressure exerted by an environment;
- a life-space causality, via learning, and contextual adaptation.

The critics of such programmes have pointed to the fact that evolutionary conditions only select a very broad frame of possibly true knowledge and that psychological conditions can only explain attitudes and opinions. Therefore, true knowledge would not be explainable by naturalization. In order to reach this goal a more general common-sense and language (culture) background should be considered (cf. KOPPELBERG, 1996).

The case of semantics or of semiotics in general is similar. Any programme of naturalization must go beyond the explanatory levels (e.g., biology and/or psychology) *and* integrate these levels into one *coherent* descriptive and explanatory whole. This can only be achieved if we apply a meta-language which is more general than the specific (historically grown) languages of the single disciplines. In the context of semiotics and semantics it is obvious that:

- physics, chemistry and biology are necessary in order to specify the denotation of perceptual and semiotic categories and processes,
- physiology (of the motor and perceptual mechanisms and the brain) is necessary in order to specify internal correlates of denoted entities and uttered signs,
- behavioural and social sciences are necessary in order to specify the conventional surface of semiotic systems and the culturally and historically specific shape of individual languages.

Consequently an explanatory model of meaning (= a naturalized semantics) must find a way to integrate all these disciplines (and relevant subdisciplines), their laws but also the specific conceptualizations into *one* framework. If the reader imagines that these sciences and their conceptualizations urge us to ask the same questions again, he could begin to feel dizzy and reproduce the panic reaction of many epistomologists and linguists who say: let us forget all the other disciplines and just do the job that we can do in our own discipline.

Since the question of a transdisciplinary solution of basic problems is not new, there also exist traditional strategies for such a solution. An optimistic one was developed by BERTALANFFY in his "General Systems Theory". More specific enterprises were:

- WIENER's "Cybernetics"
- SHANNON & WEAVER's "Information theory"
- *"Gestalt* theories"as those proposed by KÖHLER , KOFFKA , LEWIN etc.

In the last few decades a group of models have evolved which may be called morphogenetic, self-organizational, or synergetic. These form the background which my proposal will elaborate (cf. THOM, 1977; HAKEN, 1983; NICOLIS/PRIGOGINE, 1989; PENROSE, 1990).

The specific model I propose refers in several respects to the ideas of Konrad LORENZ. One of his basic intuitions was that the experience of a penetrating observation of animals as beings which are similar, and therefore understandable, to humans gives us far-reaching insights into our own thinking and even our own culture (and ethics). This presupposes a kind of continuity between animals and humans (animal societies and human societies). This basic move made by Konrad LORENZ links animal ethology – human psychology – linguistics (via animal communication) and sociology. I will accept this LORENZIAN hypothesis of an explanatory continuum, it is a basic presupposition of any programme of naturalization (cf. KOPPELBERG, 1996).

A more general continuum hypothesis could link life in general with basic chemical and physical processes. A hypothesis which is easier to subscribe to says that the environment to which an animal, a human (a mind) is adapted is specified in many respects by conditions at a more primitive ontological level, e.g., by laws of gravitation, fields of electromagnetic waves (light), fields of sound and odour, etc. A theory of the living organism must at least include some aspects of this context in order to understand the specificity of behaviour. In the following this weaker hypothesis of explanatory continuity will be used.

As opposed to a reductionist view of naturalization, I shall use a device already introduced by *Gestalt* theorists and subscribed to by Konrad LORENZ. This postulates a kind of independence for every stratum of organismic organization. This (relative) independence is covered by the term "self-organization," i.e., there are local (level-specific) processes which, by means of a self-referential effect, partially eliminate the determination from outside (from the more primitive layers of organization) and ensure that internal features gain dominance. This reduces the force of the global continuum of cause and effect and creates "islands of order."The intuition of autonomy which was first in opposition to any strategy of naturalization is now embedded as substructure into the general enterprise.

At the beginning of this section I indicated the necessity of a neutral, trans-disciplinary language which is a necessary ingredient of any programme of naturalization. I think that we have only one choice (with a rich field of sub-choices): The meta-language must be grounded in space and time. This suggests topology (geometry) and analysis (dynamical systems theory). In recent discussions LAKOFF (1987) has argued against "objectivistic" and formal semantics. I agree with his rejection of purely algebraical models (that lack a natural concept of space and time) but I do not agree with his rejection of all mathematical models, because only mathematical models allow for a neutral meta-language which has the power of abstraction necessary in the case of transdisciplinary naturalization.

The meta-language must be based on general notions of spatial and temporal organization for several reasons:

- our personal experience, and thus the phenomenological basis of knowledge, rests on conceptualization of space and time;
- the common background of all causally relevant processes and entities is space and time.

Moreover, space and time-domains are a decisive parameter in our theoretical accounts of phenomena. This was shown for optics by BERRY (1988). I shall, therefore, start my analysis of meaning phenomena with a spatio-temporal stratification.

Levels of self-organization in language

The basic phenomenon of self-organization is the emergence of new and stable patterns. Although we may find causes for the emergence in prior structure, there is a specific macro-dynamics which goes beyond the (multiple) causes, its "cause" (in a more general sense) lies in a complex interaction between local effects, the selection of very few "slaving factors", and a kind of self-stabilization by which the system creates a memory of the emergent properties and is able to reproduce the selected mode easily. In HAKEN and STADLER (1990) the laser paradigm and the fluid dynamics paradigm describe prototypes of such self-organization.

Evolutionary self-organization: the emergence of the language capacity

Language is a type of behaviour in social animals and Konrad LORENZ has shown convincingly that behavioural patterns in animals are more stable than the specific morphological patterns (body shape, colour, size, etc.). This means that animal behaviour is already a self-organised, highly stable pattern which emerges under the conditions of interaction and communication in animal societies. Now, some of these behaviours have straightforward functions for the immediate survival of the species, other behavioural patterns are rather free and functionally neutral. These patterns are the resource from which sign-behaviour can emerge. The type of behavioural pattern may be linked to specific sensory channels:

- kinetic patterns are linked to visual perception (and self-perception in the animal which moves);
- noises which are produced may be understood as signs and be used for recognition and as signals (e.g., of alarm). In the first case a specific tuning has been observed, e.g., the new born animal is tuned to the mother and this guarantees mutual identification (cf. Konrad LORENZ being identified as mother goose by the newly hatched gosling after an appropriate response);
- odours and touch sensations may be used for specific purposes of social co-ordination (cf. PRIGOGINE, 1976). The basic dynamic principle which underlies the emergence of communicative patterns can be called a principle of tuning: the behaviour of A is adapted to major rhythms and figures in the behaviour of B and vice versa. The reinterpreted behaviour must be functionally rather neutral or free. This freedom means that there is no functional determination, but it does not mean that no general laws are applied. On the contrary, if specific functions are lacking, more general laws (which apply to all behaviours of this type) become dominant.

These reflections show two things:

1. Communicative behaviour is not strictly determined by functions; therefore, it is not strictly controlled by DARWINIAN principles of survival (of the fittest). It is, therefore, astonishing that the patterns which are selected are even more simple and more systematic than morphological properties which may be the result of DARWINIAN selection.
2. There is an internal selection of the new patterns, which comes from the fact that the co-ordination of many subsystems (e.g., individual animals) only has very few stable solutions (resting points). The selection of stable solutions underlies again a DARWINIAN principle, since animal societies without stable patterns of co-ordination and communication would not survive (in conflicts with other animal societies which have developed stable solutions). In more technical terms, every social interaction between individuals, which have a large domain of freedom, tends to chaos (or noise) and domains of order are so rare that the few that do exist must be found and memorised. In an evolutionary scenario language patterns are (like behavioural patterns in general) chaos-controllers and therefore very forceful order parameters.

The specificity of human communication vs. primate communication may be due to a higher degree of freedom of humans. It originated first from brain growth which created large areas not primarily functional for motor and perceptual activities, and secondly from cultural (social) diversity due to the adaptation forced by new and rapidly changing environments (the ecological niche of humans was further transformed by the activities of human societies themselves). Language capacity is in this view not some God-given or chance generated blue-print as in CHOMSKY's "Universal Grammar", it is rather one type of solution which controls the chaotic results of increased freedom.

Historical self-organization and language acquisition

The evolutionary dimension of languages may be fixed on a scale between 100,000 and 1 million years. Historical developments cover time spans of thousands of years and may, under specific conditions, be rather rapid, as studies on change in progress by LABOV (1963) and studies on immigrant languages have shown (language loss can occur after three generations). The most selective zones of this process are critical transitions like: first language acquisition, and some minor transition points like the transition from adolescent to adult (say between the age of 15 and 17 years) where the plasticity of the brain changes dramatically, or socially defined transitions: to school – to work – to child-rearing – to retirement. The specific conditions imposed on these transitions (mainly in first language acquisition) and general (thermodynamic) processes of loss and fluctuation specify the historical development of languages. If cultural techniques like writing systems, public education, etc. are added, these basic regulatory mechanisms may be modified or higher complexities may be allowed. In a LORENZIAN perspective the process of language acquisition can be compared to the phenomenon he calls "Prägung", i.e., the quick accommodation of inborn schemata to specific patterns offered by the environment. In humans the phase of "Prägung"is much longer and the unfolding of a semantic space may be described as a channelling and bifurcation of basic inborn schemata during primary interactions with the care-giver (cf. MOTTRON, 1987).

The basic laws of language change were the major area of research in nineteenth century linguistics. Although we possess a huge mass of comparative and historical data, the laws of historical self-organization are still obscure. Rather intuitive accounts were made applying the "hidden hand"metaphor of Adam SMITH (cf. for a summary of the discussion KELLER, 1990). This metaphor is, however, only an eighteenth century forerunner of modern synergetic concepts. In order to find a model of the social dynamics of language one must consider the chance fluctuations in language use, the restrictive conditions of first language acquisition and the effects of language contact (cf. BECHERT and WILDGEN, 1991 for a summary of basic results).

Spontaneous language use and the situative emergence of meaning

In fact-to-face communication one specific language is used as an unproblematic background, therefore the dynamics described under 1.1 and 1.2 are (statistically) irrelevant. Nonetheless, new sign-forms appear and new meanings are shaped. The most obvious emergence of "new"complex sign-forms has to do with the ad hoc creation of complex structures such as phrases, sentences, texts, and dialogue sequences. This aspect is only partially covered by CHOMSKY's concept of "creativity"which he tried to model by a (deterministic) generative device, a generative grammar. What is lacking in this type of models is a proper treatment of the situational and spontaneous emergence of these structures, triggered by cues in the context, by ongoing discourse and information flows in and between participants (which may be much richer than the uttered signs). In a less dramatic way, words are subjected to ad hoc formation, as in nominal compounds and spontaneous derivations from one word-class to another in discourse (cf. WILDGEN, 1982b for the self-organization of ad hoc compounds in German).

VUKOWICH (1995) tells us that the emergence of meaning has been discussed since St. AUGUSTINE (400 A.D.) and J.J. ROUSSEAU (18th century). He points out that in everyday discourse very subtle plays with ad hoc meanings occur and that very often an interpretation given by the decontextualised sign-sequence is insufficient; it has to be constructed ad hoc and the ad hoc constructed meaning is much more natural and vivid than purely conventional sign-meaning mappings. As a consequence, the self-organization of language use and meaning becomes dependent on processes of social interaction and discourse, i.e., there is some top-down process of slaving which shapes situational utterance forms and meanings instead of them being simply the result of a generative mechanics which constructs wholes from parts (using some general rules). In the spirit of Konrad LORENZ we can say that human discourse has a behavioural macro-structure which slaves the use of linguistic patterns and their meaning. Without this behavioural Gestalt human language would be a noise without behavioural effect.

As an initial conclusion we can say that global behavioural patterns and functions are omnipresent in the self-organization of human language.

The myth of "universal grammar": a deconstruction and reformation

At every level of semiotic self-organization we observe the existence of highly stable patterns are selected and memorised (in a broad sense including biological/genetic memory). They are, however, not autonomous. They do not form an independent crystalline substructure, but are in a state of equilibrium "far from thermodynamic equilibrium,"in BERTALANFFY's terms they are in a "Fließgleichgewicht". Only if context contributes information and if energetic resources (of the mind, of social co-operativeness) are available does this equilibrium exist. It would immediately be lost, if the behavioural and biological context changed dramatically. "Universal grammar" becomes a myth, because we do not understand the complicated genesis of this stability and the forces which sustain it. Thus a successful theory of the genesis and self-organization of language will destroy this myth. I shall first propose a new perspective on basic lexical and syntactic patterns, the valence of the verb, and then turn to the basic patterns of noun-phrases. Together, these two types of structures constitute the kernel of what a "universal grammar"would have to cover. If we find a new analysis of these two domains in terms of an interdisciplinary account of self-organization, the myth of "universal grammar"can be dissipated.

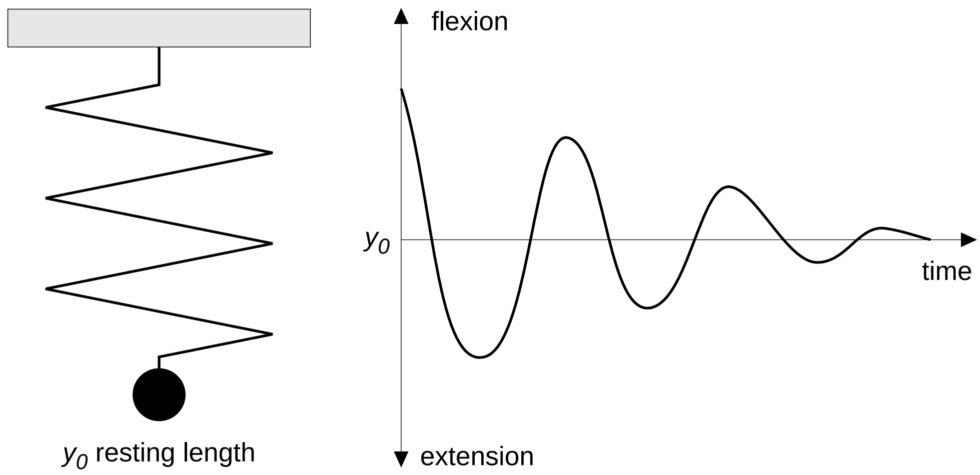

Figure 1: Ttion of an elastic pendulum and of a limb.

scale. Dynamical systems theory with its many specifications is the ideal candidate for such a neutral, flexible but nevertheless powerful meta-language. The basic ideas come from René THOM and models he proposed (between 1968 and 1977). These proposals were elaborated in WILDGEN (1979; published in WILDGEN, 1982a and 1985) and in PETITOT (1985; his dissertation of 1982). A parallel line of research is documented in BALLMER (1982) and BRENNENSTUHL (1982). I shall integrate the two lines with more recent proposals made by KUGLER, KELSO, and TURVEY (1980) who follow the tradition of biomechanics (BERNSTEIN) and ecological psychology (GIBSON; cf. WILDGEN, 1994, chapter 3).

Verbal valence and the fundamental dynamics of sentence production

In the theoretical debates of recent decades three image-schemata were used that hint at relevant underlying forces (causes, laws):

■ The chemical image-schema of "valence." An atom has specific possibilities of binding, of saturation, which predict the possible molecules (macrostructures in relation to the atomic level);

■ TESNIÈRE's notion of "actance": the sentence may be conceived as a mini-drama, in which one distinguishes central roles (protagonists–antagonists)—they are called "actants"by TESNIÈRE (1959)—and helpers, relevant background conditions ("circonstants"). This image-scheme was elaborated by the Paris school of GREIMAS.

■ CHOMSKY's notion of "government"(rectio). A traditional form of drama is that which involves kings, queens, a court, etc. If "rectio"refers to "rex" (king) government indicates rather a president and his council. The central role of the verb (predicate) corresponds to the president, the nominal roles constitute a council.

Now, if we reanalyse the three image-schemata, we see that they correspond to a scale from physical (chemical) patterns, basic behavioural patterns (the drama can have animals in specific roles as in the fable) and political patterns; as a whole it is an image of the continuum hypothesis which was discussed in the first section.

Instead of elaborating one of these image-schemata, I shall present a picture which integrates physical, behavioural, and socio-historical dynamics. Such a holistic enterprise presupposes, however, that we possess a meta-language which covers the whole

The lexicon of simple verbs (basic predicates) in natural languages shows a natural stratification along an axis called "evolutionary complexity" by BALLMER (1982, p73). He distinguishes four large areas: physical being/life/direct influence/indirect influence. I shall focus on just three prototypical subdomains:

a. locomotion (it presupposes the evolution of organisms and bodily organs),
b. direct influence/control on an object (it presupposes an asymmetry between the entity which controls and the entity which is controlled);
c. interaction (it presupposes a social structure which selects stable patterns of interaction and it constitutes the functional background of sign-usage).

For the analysis of the first subdomain I will take lexical material (German verbs) from BALLMER and BRENNENSTUHL (1986) and apply the dynamical models proposed by KUGLER et al. (1980); the other two subdomains are analysed in the same spirit but using the morphodynamic schematizations of René THOM (cf. WILDGEN, 1982a, 1985).

Bodily motion and locomotion

Movements of living bodies and body parts are subject to two types of control:

■ The non-linear control of movements, which is largely independent of specific contextual factors

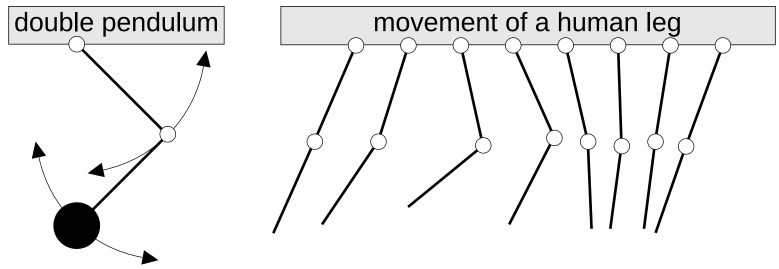

Figure 2: The motion of a double pendulum and of a human leg.

and defines the goal of a movement. Non-linear controls involve catastrophes, i.e., sudden changes in the evolution of a process.

■ The linear control adapts the movement in its metrical detail to specific contextual features, it "tunes" the qualitative motion-schema.

If we consider simple movements with one or two limbs and look for analogies in physical mechanics, we find the elastic pendulum and the double pendulum. Figure 1 shows the analogy between the motion of an elastic pendulum (left) and the motion of a limb (right) (compare Kugler et al, 1980). The right-hand side of Figure 1 shows that the peripheral mechanism of a muscular system controlling the movement of the limb is a damped oscillator of the kind given by the elastic pendulum. This means the higher (e.g., cerebral) controls only specify this peripheral quasi-autonomous system and do not govern it in detail.

Figure 2 shows the analogy between a double pendulum and the movement of a human leg. The right-hand side of Figure 2 shows phases in the movement of the human leg while the person is walking (experimental results of JOHANNSON, 1976, p386). The dynamical system of the human leg is comparable to a double pendulum (strongly damped and with restricted domains of freedom).

If a person performs a locomotion which is composed of a number of limb-motions two levels can be distinguished:

■ The rhythm of the composed movements, which is a code for the categorical perception of moving agents.

■ The overall "Gestalt" of the movement. In the case of a simple locomotion there is the initial phase which starts the locomotion. It destabilizes the system in its position of rest and creates a steady evolution until the system is at rest again. The coarse topology of locomotion has three phases:

1. Loss of position of rest, beginning of motion
2. Steady motion
3. Gain of a new position of rest, end of locomotion.

The steady motion in phase B is the basic schema, which underlies the semantics of simple verbs of locomotion like *go, run,* or *drive*. These have been traditionally characterized as *durative*. The process of locomotion of a body may involve an implicit or explicit boundary and an orientation of the process relative to this boundary. The introduction of an orientation defines a goal and introduces a kind of *intentionality*.

The cognitive schemata that have been classified here are not only relevant for the lexicon of the verb, they also form the cognitive basis for causative constructions (cf. WILDGEN, 1994, chapter 3) .

The influence/control by one agent on an object

In the prototypical situation there is one agent (a more highly organized body with its periphery) who acts on an entity which has a lesser degree of agency (matter, solid objects, living beings dominated by the agent) and which is not an inalienable part of it. We can distinguish three major aspects:

■ The configurational aspect. This aspect only concerns the spatio-temporal relationship, the topologico-dynamic "connectivity" in the scene.

■ The energetic aspect. Here the forces controlling the process, the irreversible path of an effect by an agent, are considered.

■ The intentional aspect. The direction of the energy of an agent towards some global goal is experienced by the (human) agent as an intention which either reaches or fails to reach its goal. In our framework, intentionality is a subjective interpretation of the energetic aspect.

BALLMER and BRENNENSTUHL (1986) distinguish two main groups of verbs at this level of control:

■ the creation, the destruction and the regeneration of entities (pieces of the environment),

■ the effect of an agent on the state of entities in his environment.

The first group clearly mirrors the fundamental schemata of emission and capture derived in catastrophe theoretic semantics (see WILDGEN, 1982a, p42–45 and 1985, p118–136); two sub-types of emission can be distinguished:

■ an agent creates something,

■ something appears (against a background)

In the first case, the verbal frames can take one or two nominal roles, as the following examples show:

Alan tells	a story	Charles eats	the soup
M1	M2	M1	M2
Doris sews	(a dress)	Fritz reads	(texts)
M1		M1	

Table 1: Verbal frames with one or two nominal roles

The incorporation into the verb of features pertaining to the created/observed object is a very general procedure. A converse strategy chooses a semantically poor verb such as "make" and combines it with a noun specifying the product. They have the effect that the configurational structure on the level of process schemata is not isomorphic with structures of the lexicon or the syntax. For the subtype called "regeneration" by BALLMER and BRENNENSTUHL (1986) I must refer to WILDGEN (1994, chapter 3).

The asymmetry between agent and non-agent becomes more pronounced in a group of verbs which describes the effect of an agent on the state of entities in its environment. Here the energetic/intentional aspect is dominant. Following the same strategy as above I would like to start with the description of a mechanical analogue. Two systems of pendulums are shown in Figure 3:

■ Two pendulums A and B. A gives its impulse to B (punctual transfer).

■ Two pendulums dynamically coupled. The coupling can be either rigid or elastic.

The mechanical analogue can be interpreted as a model of basic causality (by contact of rigid bodies), such that (a) stands for a chain of causes and (b) for a system of coupled causes.

The following sentences exemplify processes of type (a) and (b) in linguistic category formation. Many types of complex propulsion use a series of mechanical couplings whereby an initial force giver can cause the final locomotion. Thus, in the case of a bicycle the vertical motion of the legs is transformed into the circular motion of different wheels and finally into the horizontal locomotion of the person on the bicycle:

a. punctual transfer	b. rhythmic coupling
The player kicks the ball	The sexton tolls the bell
The man pushes the chair	The man pushes the rocking-chair
The girl throws the ball	

Table 2: Sentences expressing punctual transfervs. rhythmic coupling

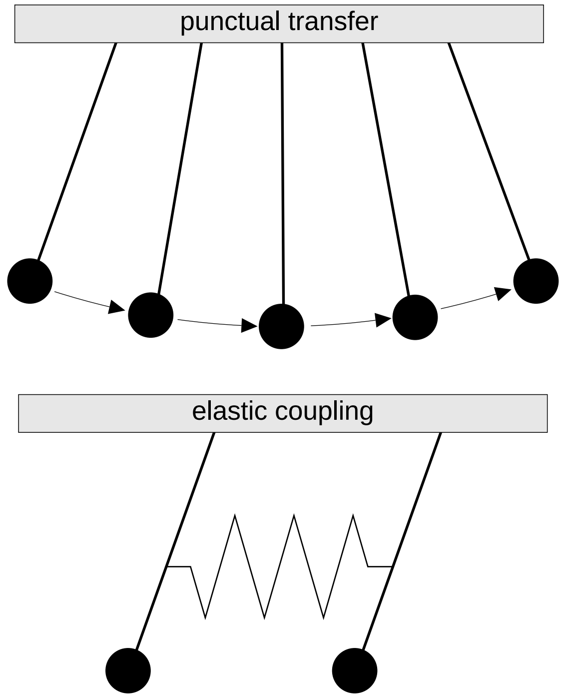

Figure 3: Two basic types of dynamic coupling.

The general configurational schema is a modified schema of emission. Two systems (the agent and the object) are co-present. The agent is in motion and the object begins to move under the control of the agent (emission of momentum). In the complementary case the momentum of an object is absorbed by the agent who stops the motion. In all cases the existence and stability of the agent and the object are not changed by the action. Due to the fact that the agent (M1) is energetically superior to the object, the energy of M1 is not absorbed totally by M2 (this would correspond to the physical example (a) in Figure 3). In Figure 4 every point on the lines of the diagram is an entity in movement.

The transfer of energy/intentional direction from M1 to M2 (via some mediator) can be either isolated (as in "kick"), repeated or continuous (as a sum of rhythmic actions).

In the case of movement, the object changes from rest to movement under the effect of the agent. In a

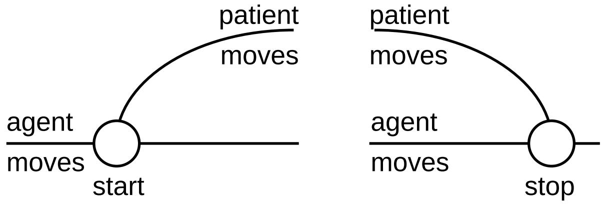

Figure 4: Emission and capture of motion/energy

similar way the object can change its shape and even its qualities. Thus if we introduce a quality space we obtain a very rich field of actions on objects which can be labelled and organized in the verbal lexicon.

The interaction of agents

The level of interaction between human agents cannot be strictly separated from the level of movement or manipulation of objects. Expanding these domains results in a higher level of organization with specific controls on the co-ordination of several (free) agents. An initial clue as to the basis for such patterns can be found in animal behaviour. FENTRESS (1982) and GOLANI (1982) show that very specific paths exist for the contact behaviour of mammals. The paths and their attractors can be lines of contact (between the tip of the mouth and the body of the partner) or lines followed in the bodily orientation of one animal (the direction of its head and its eyes). These lines follow stable paths and stabilize in very specific regions. Thus a very small sub-field of the body surface is selected for allowing contacts. Furthermore in the course of repeated contacts very specific symmetries and asymmetries in the relative behaviour appear in the interaction so that a highly ritualized pattern is created (cf. the analysis of the behaviour of wolves by FENTRESS, 1982). The punctual attraction in the relative movement of two agents plays a similar role to the body-joints in Figure 1. Thus even the patterns of interaction reproduce basic mechanical processes.

Different types of social contacts make use of different "joints":

■ the eyes of the mother are an attractor for the baby and are essential for the first contacts with the mother (humans attract humans specifically by the white parts of the eyeball and the movements of the eyes);
■ the bodily contact zone (at short distance) using the lips (compare the suckling activity of the baby); in the same way the breast of the mother is an attractor for the baby;
■ the contact at a certain distance using the hands (grasping, petting);
■ the contact of exchange (using the hands, controlled by the limbs and the eyes);
■ the communicative contact (using the mouth and the ears as instruments).

The co-ordination of the interactive processes exploits these kinematic and energetic sources and elaborates them. One specific process in this field will be more closely analyzed: the process of giving (receiving/exchanging).

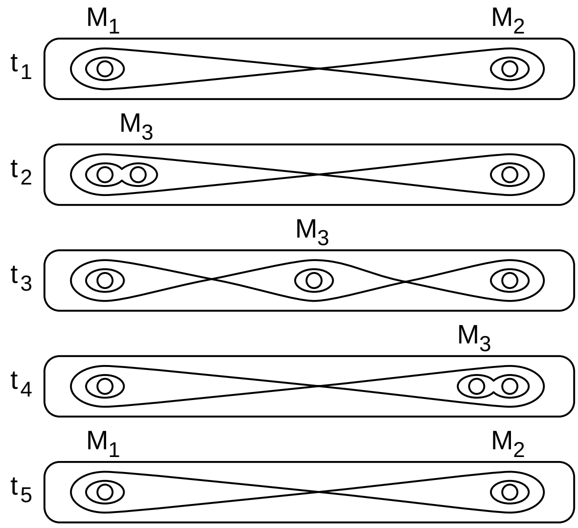

Figure 5: A topological model of the transfer schema

The basic schema or prototype of "giving" can be configurationally described by a sequence of snapshots: t_1 to t_5. Each snapshot represents an instantaneous three-dimensional configuration in which the specific positions of sender, receiver and object define a plane. The third dimension is a density (or relevance) function. This density is a correlate of the subjective focus in the perception or the motoric control of a specific region of the scene. At the beginning and at the end of the series density has two attractors (maxima of attention, relevance), in the middle of the series (t_2 to t_4) a third attractor appears, grows and finally disappears (the participants focus on the exchanged entity).[2]

The intermediate, symmetric scene at t_3 is the most unstable one. Both agents concentrate their control on one target, and their control must be coordinated in order to secure the smooth exchange. Thus, if A releases his control before B takes the object, or if A holds the object tight, although B seizes it, the character of the process is dramatically changed and degenerates to "A loses, drops the object" or "A and B compete for the object C". Thus the unstable state of exchange is the "junction" of the process, the point of maximum co-ordination of the controls. It can be a metastable state, if the object gains some autonomy, for example if it lies on a table between A and B such that it is within the reach of both but is not strictly controlled by either of them. This configuration corresponds to the topological schema of transfer (see WILDGEN, 1985,p185).

The two concepts of control and intentionality allow the construction of an evolutionary scale for behavioural and semantic scenarios:

a. a simple control is a function inside the agent and its immediate parts (limbs),
b. a second type of control is created if the agent takes into consideration an entity which is not part of himself and has its own dynamics (own forces),
c. the control necessitates a recognition of the "intention", the goal of another agent,
d. a complex scenario with objects and other agents is integrated into a higher "Gestalt". It creates a social schema which is the presupposition for the evolution of communicative routines and of language.

Complex nominals, noun-phrases and descriptions

If we extrapolate the contents of the verbs of basic events and actions to a global linguistic function, we arrive at narratives, reports of lived experience, etc. Thus the schemata described in the last paragraph constitute a basic dimension of text and communication: the recreation of past experience. Narratives serve emotional goals, thus hunters prepare or conclude hunting experiences by the exchange of narratives; they may also be directive, didactic, status relevant, etc. (cf. WILDGEN, 1994, part two).

The biological function of descriptions is basically an orientation. Thus in a narrative the speaker must first describe the scene in which events will happen. I suppose that orientation on a small scale (for large scale orientation temporal paths become prominent) is the major function of descriptions. Orientation is itself biologically rather complex. There is a basic body-centred orientation, i.e., our body is a schema for local orientation. It defines:

- a centre (the trunk, the skeleton) and a periphery (the limbs, the muscles);
- a topology of distances (in touch with the trunk, the limbs), at grasp-distance, at calling distance, etc.

This phenomenological space has global dimensions (above-below, in front-back, left-right) and it has Gestalt-features which respond to optimal, ideal types of ambient space.

Konrad LORENZ has pointed to the existence of basic internal spaces triggered by "inborn releases."A "releaser"is defined as "a differentiation, which is used for sending specific stimuli, to which a parallel differentiated receptive correlate of the member of the same species reacts selectively."(LORENZ, 1950/1965, p444; author's translation). "False"reactions of humans to animals are very revealing, e.g., their interpretation of the camel as arrogant, the eagle as proud, etc. Humans apply their "release" mechanism to phenomena outside the domain in which they are functionally relevant. A possible deep structure of this internal space is OSGOOD'S space of connotative meaning with the axis: Potency (strong-weak), Evaluation (good-bad), Activity (active-passive). Although OSGOOD proposed an interpretation in terms of evolutionary functions, a naturalization of the inner space is still lacking. His and LORENZ'examples show that the basic patterns are bimodal (as in: strong-weak) or apply an implicit scheme of negation as in active and passive (lacking activity).

The grammar of noun-phrases and complex nominals give as a good indication of the basic laws of self-organization. Thus many structures are endocentric, i.e.,

- there is a head (vs. modifier),
- there is self-similarity, i.e., the head-modifier dichotomy is iterated,
- there is a natural topology in the sequence of modifiers.

The question which one of two nominals is the head of a noun-phrase is governed by natural hierarchies, as Table 3 shows:

the tail of the dog	the dog of the tail
the lustre in the hall	the hall in the lustre
the fir-tree	the tree-fir

Table 3: Asymmetry in noun-phrases and complex nominals

In the first case the natural part-whole-hierarchy of the body predicts the head-modifier asymmetry; in the second case the relative size predicts which object may be contained in the other (the contained is the head); in the last case the biological hierarchy predicts that "fir"is more specific and should therefore take the head position. The general principle is already known from perceptual dynamics: foreground vs. background.

The specific human, and therefore new, phenomenon is iteration. However, this innovation is heavily restricted. Thus nominal compounds in German may have many subordinated constituents, as in "Donau/dampf/schiffahrts/gesellschafts/kapitäns/witwen/renten/kasse"(8 stems), but in oral speech, compounds with more than two constituents tend to be shortened or give rise to mistakes. The slower and more transparent syntactic constructions with embeddings also call for a natural topology which controls the complex array. Thus SEILER (1977) showed that in a noun-phrase like: *the beautiful red wooden ball* the sequence of modifiers: beautiful (evaluative adj.)—red (colour adj.)—wooden (material adj.) is triggered by the artefact "ball". The different qualities (material, colour, evaluation) are more or less typical ("prägnant") for the object labelled in the head position.

A more general problem of any nominal construction is given by the condition of maximal coherence of descriptive space. As every modifier adds specific dimensions or even image-schematic components, a semantic whole can only be constructed if a properly coherent spatial frame (including quality spaces) is found, into which the individual modifiers (and the head) fit.

More general principles of self-organization in grammar

Beyond lexemes, we can find processes of semantic "transport". Thus TALMY (1991) has shown for verbs that the specific manner of a process may be integrated into the verbal stem or may be coded by a "supporting satellite"(ibid, p6):

i. I *blew* out the candle (supporting verb).
ii. Apagué la vela *de un soplido* (supporting satellite).
(I extinguished the candle with a "blow")
Even within one language one may have alternatives as in German:
iii. Eine *Reise* machen/unternehmen (to go on a journey).
iv. *reisen* (to travel).

In (iii) the specific type of locomotion (travelling) is coded in the direct object of the sentence, whereas in (iv) it is contained in the verbal stem.

These examples show that the domain of verbs (predicates) underlies the same principles of complex information packaging as that shown for noun-phrases, i.e., the analysis of valence should abstract from these processes which are common to verbal and nominal structures.

If we try to naturalize semantics we must at least consider the following, very general, types of self-organization:

a. The world of designata (things, events) has been structured by evolution. The types of structures produced by evolution are mostly divergent fields (cf. for harmonically divergent fields and some analogies created by convergence, RIEDL, 1987, chapter 3). The categorical mapping of the world into linguistic signs tends to preserve this basic order; in synergetic terms we could say the major principles of evolution (relative to the world we live in and are able to react to) are slaving parameters of semantic categorization and the mental lexicon.
b. Since antiquity the human mind has been experienced as a kind of "common sense" (sensus communis) which integrates the categorical output of different senses. The semantic system which is created in the synergetic co-ordination of different sense modalities has to resolve conflicts and ambiguities and to create a multi-sensual space in which cross-sensual integration is possible.

As sophisticated models of perceptual dynamics, e.g., FREEMAN's model of smelling, show, there is a quick transition between a pre-categorical chaos attractor and a breakdown of chaos which creates categorical order. A similar transition presumably occurs in a semantic categorization and in the construction of semantic complexes (cf. WILDGEN, 1994, chapter 4 and 1997).

In the process of communicative interaction, which is partially ritualised and partially spontaneous, a profile (several concurring profiles) of relevance is created. Thus one observes asymmetries of topic-comment, head-modifier, etc. Discourse dynamics (cf. topic-comment patterns) and (more ritualised) grammatical patterns are the result of such a profiling. In discourse two major sub-functions seem to apply.

- The anchoring of the utterance in the situation by deictic and demonstrative gestures. The determiners in the noun-phrase (articles, quantifiers, etc.) are the "memory" in grammar of innumerous cases of spontaneous self-organization processes governed by this subfunction.
- The descriptive unfolding of an entity which is the topic of discourse. In section "Complex nominals, noun-phrases and descriptions" I show some aspects of this very rich phenomenon.

The historical self-organization of language

The analogy between the patterns of language change and evolutionary mechanisms was already obvious to the first generation of comparatists who reacted to DARWIN's theory (cf. SCHLEICHER, 1863). But the price of such an analogy was that languages were treated as natural organisms, such as plants and animals. This is fundamentally inadequate. A "natural" theory of language change must be more abstract. Languages must at least be seen as features of populations, which are created, lost or changed. I shall propose, therefore, a preliminary model taken from population dynamics.[3]

In order to avoid the organismic "trap,"I shall not consider the analogy between animal populations and language directly but via the interface of a formal model of population dynamics. The mathematical model is an abstract device which may be interpreted in a biological, in an economic or in a computional setting. The underlying organismic image becomes irrelevant (it is only a memory aid).

In a language we may distinguish two populations of elements which interact:

a. The *lexical entities* of a language. Their core is related to concrete and experiential concepts. We may say that this population is "fed by prelinguistic conceptualizations" in the same way that herbivores graze in the meadows.
b. The *grammatical entities* of a language. These are higher level entities, insofar as they evolve by the specialization of lexical entities. In the process of grammaticalization they loose concrete and experiential features and take over more abstract functions in the organization of linguistic utterances.

As an example of grammaticalization we can ask where prepositions with the content of: ON, UNDER, IN FRONT, BACK come from. In an analysis of 125 African languages HEINE et al. (1991) found the following dominant source domains for local prepositions:

Source domain	**No. of type**	**Example of source**
lexemes for body parts	321	head, eye
relational lexemes	101	top, front
topographical lexemes	86	sky

Table 4: Source domains of local prepositions in 125 African languages

Even in the synchronic analysis of one language we can find different stages of grammaticalization as Table 5 shows.

Table 5 shows that the content and even the signform decreases dramatically. Now, we can consider the population (a) as the set of prey and the population (b) as the set of predators. A stable equilibrium is reached if the rate of grammaticalization does not exceed the rate of reproduction of lexical items (by lexical innovation). In Figure 6 this situation is shown schematically; the unimodal systems are only realised as pathological cases.

on top of something	"top"still has its own lexical meaning
below something	only "low"has a free distribution
on something	"on"is a preposition with a spatial core-meaning
of something	"of" is fully grammaticalised and may replace a genitive
house plant	the grammatical relation is marked by juxtaposition (no morpheme as marker)

Table 5: Different stages of grammaticalization

A more refined model would have to consider grades of grammaticalization and frequency of use. Certain aphasias could be examples of the normally not observed transition to a unimodal field and I suppose that different centres of the brain or even hemispheric dominances interpret the two modes.

The LOTKA-VOLTERRA model establishes a relation between loss and growth of populations. Let us consider first a "population" of lexemes. By *loss* (coming out of use) and *growth* (lexical innovation) a net difference is produced (*growth* – *loss*). The population growth is proportional to the size of population n_1: $growth_1 = \alpha_1 \cdot n_1$.

The grammatical morphemes also have a growth/loss rate β but their growth depends additionally on the number of lexemes (proper for grammaticalization), i.e., $growth_2 = \beta_1 \cdot n_1 \cdot n_2$. The loss of the lexemes is itself proportional to the number of prey *and*

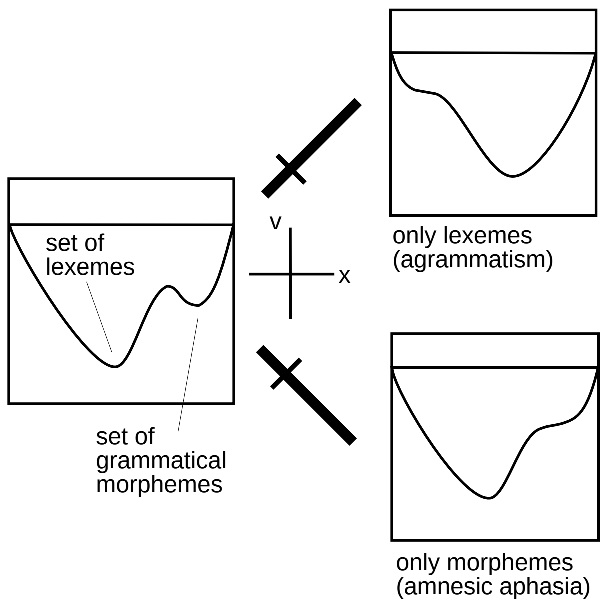

Figure 6: A bipolar field in which lexemes and grammatical morphemes coexist

predators: $loss_2 = \alpha_2 \cdot n1 \cdot n2$. The predators on their part have a loss ratio dependent only on the size of their own population ($loss_2 = \beta_2 \cdot n_2$). In general we get two equations which may be informally described as:

(i) dynamics of n_1 (lexemes) = $growth_1 - loss_1$
(ii) dynamics of n_2 (grammatical morphemes) = $growth_2 - loss_2$

In the terms $loss_2$ and $growth_2$ both population sizes are linked. LOTKA and VOLTERRA showed that this model fitted the population dynamics of sharks and shrimps in the Mediterranean (cf. STEWART, p264 f., and for a mathematical treatment HAKEN, 1983, chapter 10). A stable situation typically shows time-delayed but coupled oscillations of the two populations. In the case of grammaticalization the situation is more complicated:

- the transition between two types has intermediate stages (as Table 5 shows),
- language borrowing may bring in external grammatical morphemes,
- whole languages (or domains and functions of these languages) may fall prey to other languages and complicated patterns of language mixture at the level of lexical or grammatical entities may occur (compare the relexification discussion in the context of Creole studies, cf. WILDGEN, 1996).

These reflections on language change and grammaticalization have shown that models of self-organization exist which may explain the historical shaping of a vocabulary and a grammar.

An outlook: possible scenarios of language origin

In section "Verbal valence and the fundamental dynamics of sentence production" I argued in favour of a motor-schematic origin of basic sentential frames and in the following section the body-centred topology and an internal quality space prestructured by release-mechanisms was postulated as an area from which complex nominals, noun-phrases and finally the description of entities may have started. The last section has provided biological arguments for a functional split in grammar between lexemes (open class items) and grammatical morphemes (closed class items). What is still lacking is a coherent picture of language origin. I can only contribute some ideas to such a discussion:

a. Before linguistic capacities evolved, there were means of co-ordination at the motor-level and probably simple sounds for identification because we can observe these two phenomena in many animal societies.
b. The *Homo erectus* not only changed its habits of nutrition; his upright locomotion changed, in the long run, the geometry of the head and the face, concentrating semiotic activities in this small and specific area (mimic gestures and vocalization).
c. The social organization of prehumans became, at a certain moment, critically dependent on the improved communicative skills; a kind of forced DARWINIAN selection based on social/communicative skills was triggered. The sound communication won the competition against gestural communication and the human ear became more sophisticated (cf. categorical perception of sound patterns in the domain of vowel formants).
d. The periods of adaptation became longer for humans and neural areas without specific motor or perceptual functions appeared. Humans "managed"(in an evolutionary process) to functionalise these areas for cognitive growth and communicative skills.

There were probably one or more critical zones in this development which drew mankind towards extinction, but the biological basis (the accumulated potentials) was sufficient to select a surviving and highly competitive species. The future of this success story could be as insecure and dangerous as the evolution of man probably was, and no one knows if the biological potential is broad enough to guarantee the survival of the human species in further millennia.

If we look back to the topic of the naturalization of epistemology and other disciplines concerning the human mind, we can conclude that a purely intuitive or common sense analysis of language misses many insights which can be derived from evolutionary and behavioural biology, and from cognitive analysis in the domain of motor-programmes and perception. The fact that these insights are *a posteriori* and can be disproved in future does not diminish their value, because absolute knowledge is only an illusion of some philosophers. The more difficult challenge is the translation of disciplinary languages into some unified language in the science of man. The remarkable fail-

Author's address

Wolfgang Wildgen, University of Bremen, Fachbereich 10, Sprach- und Literaturwissenschaften, Postfach 33 04 40, D-28334 Bremen, Germany.
Email: wildgen@linguistik.uni-bremen.de

ure of the programmes of NEURATH, CARNAP, and logical structuralism (STEGMÜLLER) in the field of working scientists, and the growing frontier-lines between the natural sciences and the interpretative sciences show that this is the central challenge. Transdisciplinary enterprises like *Gestalt* theory, constructivism, morphodynamics, synergetics, autopoesis, etc., try to respond to this challenge. We need more such enterprises and a more systematic and more penetrating application of them, especially in the human and social sciences, in order to understand ourselves and the world we live in. The priority of such a complete understanding which can then direct specialised research beyond an unreflected continuation of working paradigms is also a heritage of Konrad LORENZ.

Notes

1 HERDER proposed an intuitively similar hypothesis in his treatise "Vom Ursprung der Sprache" (1772).

2 At a conference in Urbino (in 1992) Jean PETITOT showed that the technique of neural net dynamics can be used to simulate the cognitive process of finding such stable patterns in the perception of three-dimensional scenes.

3 SCHNELLE (1990) misses the point of the models I have proposed in WILDGEN and MOTTRON (1987) if he identifies them with SCHLEICHER's enterprise. The direction, the search for a naturalized theory of language, is similar, but we do not apply DARWINIAN biology to linguistics, we use very abstract mathematical devices which have interpretations in many sciences including a theory of evolution.

References

Ballmer, T. T. (1982) Biological foundation of linguistic communication. Benjamins, Amsterdam.

Ballmer T. T./Brennenstuhl, W. (1986). Deutsche Verben. Eine sprachanalytische Untersuchung des deutschen Wortschatzes. Narr, Tübingen.

Bertalanaffy, L. von (1982) General systems theory. Foundations, development and applications (revised edition). Braziller, London.

Berry, M. (1988) Breaking the paradigms of classical physics from within. In: Petitot, J. (ed) Logos et théorie des catastrophes. À partir de l'oeuvre de René Thom. Patino, Geneva.

Brennenstuhl, W. (1982) Control and ability. Towards a cybernetics of language. Benjamins, Amsterdam

Bechert, J./Wildgen, W. (1991) Einführung in die Sprachkontaktforschung. Wiss. Buchgesellschaft, Darmstadt.

Chomsky, N. (1986) Knowledge of language. Its nature, origin and use. Praeger, New York.

Greimas, A. J./Courtès, J. (1979/1986) Sémiotique. Dictionnaire raisonné de la théorie du langage, vol.1 (1979), vol. 2 (1986). Hachette, Paris.

Freeman, W. J. (1995) The creation of perceptual meanings in cortex through chaotic itinerancy and sequential state transitions induced by sensory stimuli. In: Stadler/Kruse, 1995:421–437.

Haken, H. (1983) Synergetics. An introduction (third revised and augmented edition). Springer, Berlin.

Haken, H./Stadler, M.(eds) (1990) Synergetics of cognition. Proceedings of the international symposium at Schloß Elmau. Springer, Berlin.

Heine, B./Claudi, U./Heinmeyer, F. (1991) Grammaticalization. A conceptual framework. The University of Chicago Press, Chicago.

Herder, J. G. (1772/1966) Abhandlung über den Ursprung der Sprache. Reclam, Stuttgart.

Johansson, G. (1976) Spatio-temporal differentiation and integration in visual motion perception. Psychological Research 38: 379–393.

Keller, R. (1990). Sprachwandel. Von der unsichtbaren Hand in der Sprache. Francke, Tübingen (UTB 1567)

Koppelberg, D. (1996) Erkenntnistheorie und Wissenschaft—Naturalistische Positionen (presented at the 5thBremer symposion of philosophy of science). Paper.

Kruse, P./Stadler, M. (eds) (1995) Ambiguity in mind and nature. Multistable cognitive phenomena.. Springer, Berlin.

Kugler, P. N./Kelso, I. A. S. /Turvey, M. T. (1980) On the concept of coordinative structures. In: Stelmach, G.E./ Requin. I. (eds.), Tutorials in Motor Behavior. North Holland, Amsterdam: 3–47.

Labov, W. (1963) The social motivation of a sound change. Word 19: 273–309.

Lakoff, G. (1987) Women, fire, and dangerous things. What categories reveal about the mind. Chicago U.P., Chicago.

Lorenz, K. (1965) Über tierisches und menschliches Verhalten. Aus dem Werdegang der Verhaltenslehre. Gesammelte Abhandlungen Bd. I und II. Piper. München.

Lorenz, K./Leyhausen, P. (1968) Antriebe tierischen und menschlichen Verhaltens. Gesammelte Abhandlungen. Piper, München.

Mottron, L. (1987) Morphogenese und Bedeutungsstruktur. Ein dynamisches Modell des Spracherwerbs beim normalen und beim autistischen Kind. In: Wildgen / Mottron, 1987: Part II.

Nicolis, G./Prigogine, I. (1989) Exploring complexity. An introduction. Freeman, New York.

Osgood, Ch. E. (1976) Focus on Meaning. Mouton, The Hague.

Penrose, R. (1990) The emperor's new mind. Concerning computers, minds, and the laws of physics. Oxford U.P. (Vintage), New York.

Petitot, J. (1985) Morphogenèse du sens. Pour un schématisme de la structure. Presses Universitaires de France, Paris.

Prigogine, I. (1976) L'ordre par fluctuation et le système social. In: Vorträge der Westfälischen Akademie der Wissenschaften (260). Westdeutscher Verlag, Opladen: 7–48.

Riedl, R. (1987) Begriff und Welt. Biologische Grundlagen des Erkennens und Begreifens. Parey, Berlin.

Seiler, H. (1977) Determination: A Functional Dimension for Inter-Language Comparison. Arbeiten des Kölner Universalien-Projekts (akup) 23.

Schleicher, A. (1863) Die Darwinsche Theorie und die Sprachwissenschaft. Böhlau, Weimar.

Schnelle, H. (1990) Die Natur der Sprache. Die Dynamik der Prozesse des Sprechens und Verstehens. De Gruyter, Berlin.

Stewart, I. (1990) Does God play dice? The mathematics of chaos. Penguin Books, London.

Talmy, L. (1991) Path to Realization: a Typology of Event Conflation. Buffalo Papers inLinguistics 1.

Tesnière, L. (1959) Éléments de syntaxe structurale. Klincksieck, Paris.

Thom, R. (1977) Stabilité structurelle et morphogénèse. Essai d'une théorie générale des modèles. Deuxième édition revue, corrigée et augmentée. Interédition, Paris.

Vukowich, A. (1995) The emergence of meaning by linguistic problem solving. In: Kruse and Stadler, 1995: 241–252.

Wildgen, W. (1982a) Catastrophe Theoretic Semantics. An Elaboration and Application of René Thom's Theory.Benjamins, Amsterdam.

Wildgen, W. (1982b) Zur Dynamik lokaler Kompositionsprozesse. Am Beispiel nominaler *ad hoc*-Komposita im Deutschen. Folia Linguistica 16 (1–4):297–344.

Wildgen, W. (1985) Archetypensemantik. Grundlagen einer dynamischen Semantik auf der Basis der Katastrophentheorie.Narr, Tübingen.

Wildgen, W. (1994) Process, image and meaning. A realistic model of the meaning of sentences and narrative texts. Benjamins, Amsterdam.

Wildgen, W. (1996) Die Dynamik des Haiti-Kreols: ein Fallbeispiel für die Interaktion von Sprachen und Kulturen. In: Ulfers, H. / M. Franzbach (eds) Togo, Kamerun und Angola im euro-afrikanischen Dialog von Geschichte, Kultur und Politik. Universitätsverlag, Bremen:112–124.

Wildgen, W. (1997) Chaos, fractals and dissipative structures in language or the end of linguistic structuralism. In: Koch, W. / Altmann, G. (ed) Systems: New paradigms for the human sciences (in print). de Gruyter, Berlin.

Wildgen, W./Mottron L. (1987) Dynamische Sprachtheorie. Sprachbeschreibung und Spracherklärung nach den Prinzipien der Selbstorganization. Brockmeyer, Bochum.

Jean Choi / Irwin Silverman

Sexual Dimorphism in Spatial Behaviors: Applications to Route Learning

Abstract

Recent studies with humans have revealed sex differences in preferred strategies for route learning and other spatial mapping tasks, with males disposed to Euclidean methods (the use of distance concepts and cardinal directions) and females to topographical techniques (the use of landmarks and relative directions). The present study assessed whether females' preference for topographical strategies stemmed from their lesser Euclidean abilities, or whether the difference represented a manifestation of an evolved dimorphism whereby females acquired superior capacities for object location recall based on their ancestral roles as caretakers of the habitat and foragers for food. The latter explanation was favored.

Key words

Spatial mapping, route learning, sex differences, estrogen, division of labor.

Introduction

Recent studies with humans have revealed sex differences in preferred strategies for route learning and other spatial mapping tasks, with males predisposed to Euclidean methods and females to topographical techniques. Specifically, males show a greater tendency to use distance concepts while females rely more on landmarks (Bever, 1992; Galea/Kimura, 1993; Holding/Holding, 1989; McGuiness/Sparks, 1983; Miller/Santoni, 1986). Males are also more inclined to use cardinal directions; that is, north, south, east and west, rather than relative directions such as right, left, front and back (Ward/Newcombe/Overton, 1986). Analogous sex differences have been found in rodents. In solving radial arm mazes, females use objects in the surrounding environment as landmarks, while males depend more on cues based on the geometric shape of the room (Williams/Barnett/Meck, 1990; Williams/Meck, 1991).

One view of the origins of these differences derives from well-established findings in humans that, in general, males outperform females in spatial abilities (see Halpern, 1992; Kimura/Hampson, 1993, for reviews). These findings have also been shown to apply to rodents (Binnie-Dawson/Cheung, 1982; Dawson, Cheung/Lau, 1975; Gaulin/FitzGerald, 1986; Gaulin/FitzGerald/Wartell, 1990). Thus, reliance on topographical, rather than Euclidean, methods in spatial mapping may represent a compensatory mechanism employed by females in species where they possess less of the spatial capacities required for effective Euclidean-based solutions. This notion is strongly implied in the conclusions of Galea and Kimura (1993) and Miller and Santoni (1986).

An alternate perspective in the human case, however, stems from the theory of Silverman and Eals (1992; Eals/Silverman, 1994) that spatial sex differences in general represent an evolved dimorphism based on division of labor. These authors maintained that the spatial tests on which males outperformed females tended to measure attributes associated with successful hunting. They further demonstrated that women surpassed men in original spatial tasks which mimicked the female's evolutionary roles as caretakers of the habitat and foragers for plant foods; namely, tasks involving the ability to recall objects and their locations in spatial arrays. Thus, the female's disposition to use landmarks rather than distance concepts in spatial mapping may be a function of her particular proficiency in recalling object locations. This analysis would also apply to the use of relative rather than cardinal directions, inasmuch as the former includes directions based on the location of objects in relation to one another as well as to the observer.

Though a direct test of evolutionary causation appears infeasible, the “compensatory” and the “evolved dimorphism” theories do lead to differential predictions regarding patterns of relationships among spatial mapping variables. These differential predictions were the focus of the present study. For one, the assumption of the compensatory model is that EUCLIDEAN methods are the most effective strategy for spatial mapping and will be used and preferred by the individual to the extent that he/she is capable. Thus, it would be expected that, for both sexes, use and preference for EUCLIDEAN methods will show a positive relationship with route learning success. If, on the other hand, strategy preferences between sexes emanated from an evolved dimorphism, members of each sex would be expected to be most successful using the strategy for which they are innately inclined. In this case, males' performances will relate to the use and preference for EUCLIDEAN methods; females' to the use and preference for topographical methods.

Differential predictions may also be inferred regarding the relationships between performance on general spatial tests, which are essentially based on EUCLIDEAN principles, and route learning variables. The compensatory model would predict that general spatial test scores will relate positively to use and preference for EUCLIDEAN methods, as well as to route learning performance for both sexes. The evolved dimorphism theory, however, would presume that these relationships should pertain only to males. The expected counterpart relationships for females should be between tests of object location recall and use and preference for topographical methods and route learning success.

A final set of hypotheses based on the evolved dimorphism theory is that, for females, use and preference for topographical strategies and route learning success will increase with elevated estrogen levels. These predictions are based on prior studies showing increased performance levels on various female-biased cognitive tasks occurring with transient increases in estrogen levels, which have been invoked as evidence for the genetic and evolutionary bases of the abilities measured by these tests (see HAMPSON/KIMURA, 1992, for a review).

Method

Test Materials

■ *Spatial mapping.* This was measured by a route learning task, based on a map (Figure 1) adapted from that used by WARD et al. (1986). Some landmarks in the original map with uniquely American derivations were replaced with Canadian counterparts, and several new points of origin and destination were added. The latter was necessary because WARD et al. tested route learning from a single point of origin to a single destination, whereas the present study required each subject to construct four separate routes from a single point of origin (Kenrick Street Apartment Complex) to four different destinations (Marina Restaurant, Memorial Cultural Centre, Mercy High School, and Airport). As shown in Figure 1, the map provided a scaled metric measure and a cardinal direction indicator.

Written instructions for the test contained a sample map. Subjects were informed they would be asked to study a larger map of the same type for six minutes; then, "from memory, write out directions to various destinations from a single point of origin". They were told that the point of origin would be highlighted and each destination marked with an *X*. They were asked to pretend that, "the person you are giving directions to is driving, has never seen the map before, and is unfamiliar with the area". Six minutes were allowed for the writing of the four sets of directions.

Individual scores coded from the test protocols were:

(1) frequency of distance estimations in terms of either mileage or blocks (*Frequency Distance)*; (2) frequency of landmark references, comprising of buildings, traffic lights, gas stations, lakes and railways (*Frequency Landmarks*); (3) preference for landmarks over distances, measured in terms of the percentage of total landmarks and distance references which were landmarks (*Preference Landmarks*); (4) frequency of cardinal directions, comprising of north, south, east, west, or any combination thereof (*Frequency Cardinal)*; (5) frequency of relative directions, such as left, right, front, back (*Frequency Relative*); (6) preference for relative over cardinal directions, measured in terms of the percentage of total relative and cardinal directions which were relative (*Preference Relative*); (7) spatial mapping performance, measured in terms of the number of routes accurately described (*Mapping Performance*).

■ *General spatial skills.* This was measured by an abbreviated, 12 item form of the Three-Dimensional Mental Rotations Test adapted by VANDENBERG and KUSE (1978) from SHEPARD and MELTZER (1972). This test was chosen due to its tendency to show the strongest and most reliable sex differences, heritability estimates, and hormonal effects (SILVERMAN/PHILLIPS, 1993). The test requires subjects to determine which two, in a series of four drawings of three-dimensional objects depict a target object in rotated positions.

Subjects were told they would be scored a point for each correct response and a point subtracted for each incorrect response, yielding a possible range of

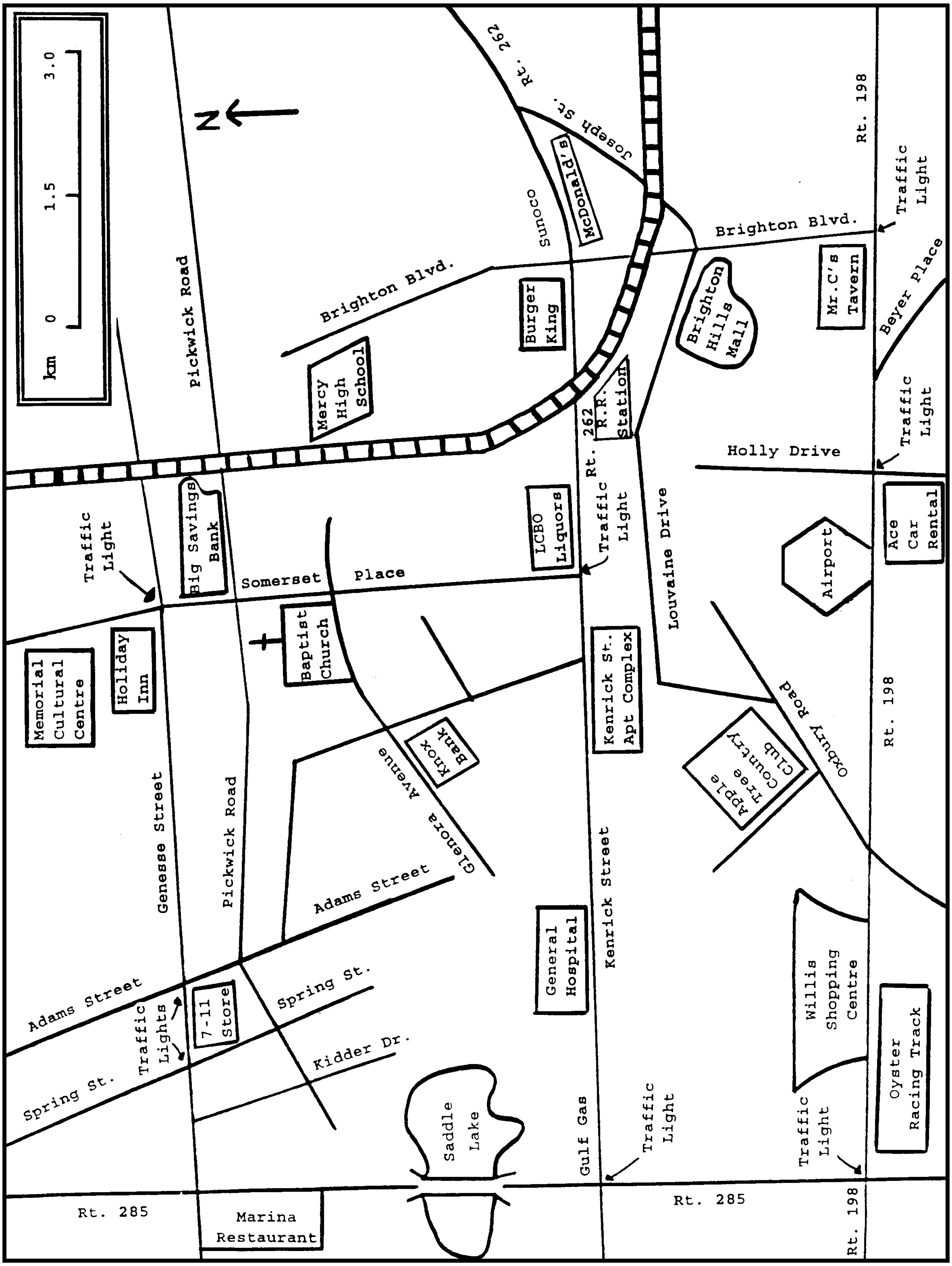

Figure 1: Map utilized in the Spatial Mapping Task.

scores from –24 to +24. Three and one half minutes were allowed.

■ *Object location recall.* This was evaluated by a simpler version of the Building Memory Test (EKSTROM/FRENCH/HARMAN/DERMEN, 1976), which is essentially a direct measure of landmark location recall. In the present version, subjects were asked to study a schematic of a city map containing drawings of twelve diverse structures in twelve different locations, for four minutes. They were then required to identify the location of each structure on a test map from memory. Subjects chose one of five alternative responses for each identification; therefore, one point was given for each correct response and one fifth of a point was subtracted for each error as a correction for guessing. Thus, the range of possible scores was –2.40 to +12.

Procedure

■ *Phase One:* Subjects in the first phase of the study were 36 male and 51 female paid volunteers from both day and evening undergraduate psychology courses at York University. The aforementioned tests were administered in the order in which they are listed, in three groups of approximately equivalent sizes. Mean ages of males and females were 25.5 (SD = 8.80) and 25.8 (SD = 5.7) respectively, a difference which did not approach significance.

■ *Phase Two*: Subjects in the second phase of the study were 20 female paid volunteers, undergraduate and graduate students from York University in Toronto and Canisius College in Buffalo, with a mean age of 24.05 (SD = 2.93). Each subject was administered separate spatial mapping tasks on two separate occasions; during menses, when estrogen levels are at their lowest in the menstrual cycle, and during the luteal phase (defined as 16 to 23 days following the onset of menses), when estrogen levels are at one of the two highest peaks of the cycle.

This design required construction of a second test map, identical to the first in all relevant aspects (number of available landmarks, directional references, routes to be learned, etc.) and administered in precisely the same manner. Performance means for the two test maps were virtually identical. Nevertheless, order of testing for menstrual cycle phase and test form were counterbalanced with respect to each other. Sessions for each subject were scheduled within the same menstrual cycle if the first testing session was during menses, or between two adjacent cycles if the first session was during the luteal phase.

Results and Conclusions

■ *Descriptive data*: Table 1 shows mean scores by sex for all of the measures of Phase One. For the spatial mapping variables, the pattern of results closely conforms to previously reported strategy differences between sexes. Males used distance concepts and females used landmarks significantly more frequently, and although both groups tended to use landmarks proportionally more than distances (which can be attributed in any given case to the specific nature of the map), females showed a significantly greater mean preference score in that direction. Similarly, sex differences for frequencies of both cardinal and relative directions showed the expected trends, approaching significance at p=.10 for the former. The preference measure was also in the expected direction, but remained statistically nonsignificant.

	Males (n = 36)	Females (n = 51)	F
Frequency Distances	1.75 (2.31)	0.29 (0.73)	17.82***
Frequency Landmarks	3.36 (3.15)	5.14 (4.23)	4.56**
Preference Landmarks	0.58 (0.41)	0.81 (0.36)	7.94**
Frequency Cardinal	5.81 (4.04)	4.31 (3.97)	2.94
Frequency Relative	4.36 (3.74)	5.55 (4.92)	1.49
Preference Relative	0.40 (0.33)	0.50 (0.40)	1.46
Mapping Performance	3.31 (1.06)	3.41 (1.00)	0.22
Mental Rotations	8.58 (5.31)	3.59 (4.09)	24.52***
Location Recall	8.17 (3.04)	7.90 (3.17)	0.15

*) $p < .05$ **) $p < .01$ ***) $p < .001$ (Two-tailed tests)

Table 1: Means and (standard deviations) of test scores by sex.

In general, sex differences in the strategy variables appeared to be at least as strong and consistent as prior comparable studies, such as MILLER and SANTONI (1986) and GALEA and KIMURA (1993). Both of these studies, however, showed superior male performance in route learning, whereas in the present data sexes were essentially equivalent in overall performance levels. Inasmuch as performance success in the present study was measured by the number of routes completed, a post-hoc analysis was undertaken to ascertain whether females may have taken less efficient routes. For all destinations, however, the proportions of males and females giving the most direct routes were equivalent.

Regarding spatial ability measures, the expected male bias in the three-dimensional mental rotations task was significant and robust, replicating copious prior findings. The measure of object location recall, however, failed to show a significant sex difference or trend of any degree. Thus, for whatever reason, this test did not serve as a discriminative measure of the attributes which underlie previously demonstrated superior object location recall for females, and these scores were not considered in any further analyses.

■ *Correlations with spatial mapping performance*: Table 2 shows correlations of spatial mapping performance with variables comprising of strategy use and preference, and mental rotations scores, separately by sex. Mental rotations, which constituted the measure of general spatial abilities, failed to achieve significant relationships to mapping performance in either sex. The correlation coefficient for females, however, approached statistical significance at $p < .10$, and the coefficient for males was similar in magnitude, implying that this confidence level may have also approached significance if not for the smaller N. Thus, the compensatory model may receive some support from these trends if they are borne out to a significant level in subsequent research.

	Males (n=36)	Females (n=51)
Frequency Distances	0.33*	0.10
Frequency Landmarks	0.27	0.47***
Preference Landmarks	0.04	0.53***
Frequency Cardinal	0.50**	0.32*
Frequency Relative	0.27	0.39**
Preference Relative	0.14	0.28*
Mental Rotations	0.22	0.23

*) $p < .05$ **) $p < .01$ ***) $p < .001$ (Two-tailed tests)

Table 2: Correlation Coefficients of Strategy Preferences and Mental Rotations Scores to Spatial Mapping Performance by Sex

Correlations involving mapping strategy variables, however, strongly supported the evolved dimorphism theory. Frequency of distance concepts was positively and significantly related to performance levels for males only, while both frequency and preference for landmarks were positively and significantly related to performance levels for females. Frequency of cardinal directions was positively related to performance in both sexes; however, both frequency and preference for relative directions were positively and significantly related to performance for females only. Thus, overall, males showed superior performance when using EUCLIDEAN methods; females performed better using topographical methods.

■ *Correlations of mental rotations to strategy preferences*: These are shown in Table 3. Patterns were definitive and largely similar between the sexes, but inconsistent with either the compensatory or evolved dimorphism model. Wholly unexpectedly, mental rotations scores correlated with the use of topographical rather than EUCLIDEAN strategies in both sexes.

	Males (n=36)	Females (n=51)
Frequency Distances	0.20	0.09
Frequency Landmarks	0.58**	0.36***
Preference Landmarks	0.43**	0.13
Frequency Cardinal	0.00	–0.13
Frequency Relative	0.38*	0.40**
Preference Relative	0.30*	0.30*
Frequency Distances	0.20	0.09

*) $p < .05$ **) $p < .01$ ***) $p < .001$ (Two-tailed tests)

Table 3: Correlation Coefficients of Mental Rotations Scores to Strategy Preferences by Sex.

■ *Hormonal effects:* No strategy variables showed significant effects of hormonal status, though differences in preference for landmarks approached significance. Means were .98 (SD = .05) and .88 (SD = .26) for menses and midluteal phases respectively ($p < .10$), indicating an increase in landmark preference with decreased estrogen level. This is in direct opposition to the direction anticipated by the evolved dimorphism model.

Discussion

Ambiguities in the data notwithstanding, the strongest patterns to emerge were consistent with the evolved dimorphism model. The data were decisive in their demonstration that, on an inter-individual basis, successful performance was positively related to the use of EUCLIDEAN methods for males and topographical methods for females. In light of these findings, it would appear that females do not necessarily use topographical methods in spatial mapping because they lack EUCLIDEAN capabilities, but because they possess a unique capacity for object location recall.

Authors' address

Jean Choi & Irwin Silverman, Department of Psychology, York University, 4700 Keele Street, North York, Ontario, Canada, M3J 1P3.
Fax: 416-736-5814, Email: isilv@yorku.ca

The hormonal data were inconclusive. The relationship of menstrual cycle phases with landmark preference did approach significance, but the preference was greater with decreased estrogen levels, which was in the reverse direction than predicted from the evolved dimorphism model. This finding is consistent, however, to recent data on menstrual cycle effects for Silverman and Eals' object location recall test by GAULIN, SILVERMAN, PHILLIPS and REIBER, which showed the same inverse relationship with estrogen at about the same confidence level. It may be that estrogen, at some levels, operates as a generalized inhibitor of spatial capacities, whether the specific capacity favours males or females.

In one sense, the results and conclusions of the present study are contradictory to both GALEA and KIMURA (1993) and MILLER and SANTONI (1986), who found expected strategy differences between sexes coupled with superior male performance in route learning. The expected differences in EUCLIDEAN vs. topographical strategies between sexes did occur in our study, but route learning performance means were virtually identical.

The dependent measures in the prior investigations, however, involved process variables, such as time to solution and deviations from a presumed correct route, rather than number of routes completed. Based on the present notion of an evolved dimorphism involving diverse spatial strategies, it would appear that measures of route-learning based on process, rather than outcome, may be readily biased toward either sex, depending on the placement of landmarks and vectors in the mapping test. This is consistent with BEVER'S (1992) observation in an actual direction-finding task using a labyrinth of hallways. Females navigated more rapidly than males when travelling in one direction, whereby landmarks operate as unequivocal cues. Males, on the other hand, were quicker when travelling to a target point and returning, whereby the directional cue functions of landmarks become equivocal in the return phase.

We were frankly puzzled by the findings that mental rotations scores, which are regarded as a foremost measure of general spatial ability, predicted the use of topographical, and not EUCLIDEAN strategies, for both sexes. We would be tempted to regard these findings as a quirk; however, James DABBS and his colleagues at Georgia State University have observed a similar pattern of relationships in their ongoing research (personal communication), suggesting that there are facets of mental rotations performance yet to be explored.

Overall, the present findings provide additional support, albeit with qualifications, to burgeoning data demonstrating sexually dimorphic spatial perceptual processing mechanisms originating with division of labour and roles in human evolution. More precise studies emanating from the questions raised by the present data, particularly in regard to the standardization of distance and landmark cues, may ultimately resolve some of these qualifications.

References

Bever, T. (1992) The logical and extrinsic sources of modularity. In: M. Gunnar and M. Maratsos (eds), Modularityand Constraints in Language and Cognition (Vol.25 of Minnesota Symposia on Child Psychology). Lawrence Earlbaum and Associates, pp.179–212.

Binnie-Dawson, J. L. M./Cheung, Y. M. (1982) The effects of different types of neonatal feminization and environmental stimulation on changes in sex associated activity/spatial learning skills. Biological Psychology, 15, 109–140.

Dawson, J. K. M./Cheung, Y. M./Lau, R. T. S. (1975) Developmental effects of neonatal sex hormones on spatial and activity skills in the white rat. Biological Psychology, 3, 213–219.

Eals, M./Silverman, I. (1994) The hunter-gatherer theory of spatial sex differences: Proximate factors mediating the female advantage in recall of object arrays. Ethology and Sociobiology, 15, 95–105.

Ekstrom, R. B./French, J.W./Harman, H. H./Dermen, D. (1976) Kit of Factor-Referenced Cognitive Tests. Princeton: Educational Testing Service.

Galea, L./Kimura, D. (1993) Sex differences in route learning. Personality and Individual Differences, 14, 53–65.

Gaulin, S. J. C./FitzGerald, R.W. (1986) Sex differences in spatial ability: An evolutionary hypothesis and test. TheAmerican Naturalist, 127, 74–88.

Gaulin, S. J. C./FitzGerald, R. W./Wartell, M. S. (1990) Sex differences in spatial ability and activity in two vole species. Journal of Comparative Psychology, 104, 88–93.

Gaulin, S. J .C./Silverman, I./Phillips/Reiber *Unpublished manuscript available from the present authors*.

Halpern, D. F. (1992) Biological hypothesis part I: Genes and hormones, In: Sex Differences in Cognitive Abilities, 2nd Edition. Hillsdale: Lawrence Earlbaum Associates, Inc., pp.99–135.

Hampson, E./Kimura, D. (1992) Sex differences and hormonal influences on cognitive function in humans. In: J.B. Becker, S.M. Breedlove/D. Crews (eds), Behavioral Endocrinology. Cambridge: MIT Press, pp.357–398.

Holding, C. S./Holding, D. H. (1989) Acquisition of route network knowledge by males and females. The Journal ofGeneral Psychology, 116, 29–41.

Kimura, D./Hampson, E. (1993) Neural and hormonal mechanisms mediating sex differences in cognition. In: P. Vernon (ed), Biological Approaches to the Study of Intelligence. New Jersey: Ablex Publishing, pp.375–397.

McGuiness, D./Sparks, J. (1983) Cognitive style and cognitive maps: Sex differences in representations of a familiar terrain. Journal of Mental Imagery, 7, 91–100.

Miller, L. K./Santoni, V. (1986) Sex differences in spatial abilities: Strategic and experiential correlates. ActaPsychologica, 62, 225–235.

Shepard, R. N./Meltzer, J. (1971) Mental rotation of three dimensional objects. Science, 171:701–3.

Silverman, I./Eals, M. (1992) Sex differences in spatial abilities: Evolutionary theory and data. In: J.H. Barkow, L. Cosmides/J. Tooby (eds), The Adapted Mind: Evolutionary Psychology and the Generation of Culture. New York: Oxford Press, pp.531–549.

Silverman, I./Phillips, K. (1993) Effects of estrogen changes during the menstrual cycle on spatial performance. Ethologyand Sociobiology, 14, 257–270.

Vandenberg, S. G./Kuse, A. R. (1978) Mental rotations: A group test of three-dimensional spatial visualization. Perceptualand Motor Skills, 47, 599–604.

Ward, S. L./Newcombe, N./Overton, W. F. (1986) Turn left at the church, or three miles north: A study of direction giving and sex differences. Environment and Behaviour, 18, 192–213.

Williams, C. L./Barnett, A. M./Meck, W. H. (1990) Organizational effects of early gonadal secretions on sexual differentiation in spatial memory. Behavioral Neuroscience, 104, 84–97.

Williams, C. L./Meck, W. H. (1991) The organizational effects of gonadal steroids on sexually dimorphic spatial ability. Psychoneuroendocrinology, 16, 155–176.

Manuela Delpos

Review of the Journal Cybernetics and Human Knowing

A Journal of Second Order Cybernetics & Cyber-Semiotics

This multi- and interdisciplinary journal is edited and published quarterly (now Vol.4) by Soeren BRIER in Denmark and has an international and interdisciplinary editorial board consisting of Peter Bogh ANDERSEN, Denmark; Evelyne ANDREWSKY, France; M.C. BATESON, USA; Stafford BEER, Canada; Heinz VON FOERSTER, USA; Ernst VON GLASERSFELD, USA; Ranulph GLANVILLE, UK; Louis KAUFFMANN, USA; Klaus KRIPPENDORFF, USA; George LASKER, Canada; Ervin LASZLO, Italy; Niklas LUHMANN, Germany; John MINGERS, UK; Humberto MATURANA, Chile; Edgar MORIN, France; Roland POSNER, Germany; Lars QVORTRUP, Denmark; Thomas SEBEOK, USA; Kjell SAMUELSON, Sweden; Robert VALLÈE, France and Francisco VARELA, France.

These scientists, the consulting editors and the authors reflect the wide range of how second order cybernetics, which is the cybernetics of cybernetics, is understood. They come from various disciplines, namely architecture, biology, chemistry, computer science, economics, education, librarian studies, mathematics, media studies, medicine, philosophy, physics, psychology, semiotics, sociology, and technics. This interdisciplinary approach has been influental on issues such as Systems Theory, Radical Constructivism, Theory of the Artificial to the Concept of Information, Semiotics and Consciousness.

The journal thereby focuses recent developments in understanding of self-organizing processes of information in human knowing, that have arisen through the cybernetics of cybernetics, or second order cybernetics. This development within cybernetics is a non-disciplinary approach. Through the concept of self-reference it tries to explore: the meaning of cognition and communication; our understanding of organization and information in human, artificial and natural systems; and our understanding of understanding within the natural and social sciences, humanities, information and library science, and in social practices as design, education, organization, teaching, therapy, art, management and politics.

Due to the interdisciplinary character of the journal the articles are highly original and are written such, that readers from other fields can understand them. This is enhanced by a narrative writing style and the occasional additional glossary. The composition of the themes in each issue includes a special section for articles from practitioners of cybernetic theory as well as from theoreticists. As a result there are different, sometimes controversial approaches to a specific theme which broadens the spectrum of understanding and invites discussion. In this respect, all relevant symposiums which are strongly related to the topic respectively second-order cybernetics and semiotics are announced at the end of the journal. In order to emphasize the very new scientific approach towards diverse topics, the articles are accentuated by related art drawings and poems. This altogether gives an interesting example of how a new scientific concept can be realized in a journal.

Subscriptions & further informations

Soeren Brier, M.Sc.,Ph.D., Assoc. Prof. in Information Studies in Science and Arts, Royal School of Librarianship, Langagervej 4, 9220 Aalborg Oest, Denmark. "Cybernetics & Human Knowing" homepage: http://www.db.dk/dbaa/sbr/cyber.htm, email:sbr@db.dk