

**METAZOAN COMPLEXITY AND EVOLUTION:
IS THERE A TREND?**

Daniel W. McShea

January 1996

Mailing Adress:
Santa Fe Institute
1399 Hyde Park Road
Santa Fe, NM 87501

Phone: 505-984-8800
E-mail: dwm@santafe.edu

January 1996

LRH: DANIEL W. McSHEA

RRH: METAZOAN COMPLEXITY

METAZOAN COMPLEXITY AND EVOLUTION: IS THERE A TREND?

DANIEL W. McSHEA

Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501

ABSTRACT

The notion that complexity increases in evolution is widely accepted, but the best-known evidence is highly impressionistic. In this paper, I propose a scheme for understanding complexity which provides a conceptual basis for objective measurement. The scheme also shows complexity to be a composite term covering four independent types. For each type, I describe some of the measures that have been devised and review the evidence for trends in the maximum and mean. In metazoans as a whole, there is good evidence only for an early-Phanerozoic trend, and only in one type of complexity. For each of the other types, some trends have been documented but only in a small number of metazoan subgroups.

Key words.--complexity, trends, hierarchy, Metazoa, macroevolution.

The centerpiece of the case for a pervasive evolutionary trend in complexity has always been a story. Figure 1 tells one version: the first organisms were simple and single-celled. From these arose more complex multicelled invertebrates, which in turn were followed by primitive vertebrates, then mammals, and finally the most complex species of all, human beings.

Many find the story compelling, and to some it might make the existence of a trend seem too obvious to question. But there have always been reasons for doubt. First, what is complexity? Is it number of parts? Number of interactions among parts? Degree of functionality of those interactions? Or a combination of all three? The story does not say, and no general consensus exists. Even if a consensus could be reached -- say, on complexity as a joint measure of number of parts and functionality -- how would we demonstrate a trend? To compare parts-and-functionality in a human and a trilobite, for example, how would we proceed? These questions have no satisfactory answer at present, and therefore such comparisons (along with the trend inferred from them) are purely impressionistic.

Second, even if the sequence in Figure 1 *is* an increase in complexity in some sense, it documents just one case, not a pervasive trend. Many other (putative) cases of increase have been cited in the evolutionary literature (Cope 1871; Spencer 1890; Rensch 1960b), such as the increase in segment

differentiation in arthropods, in folding of the vertebrate brain, and so on. But marshalling cases does not document a pervasive trend either. The many increases could well be offset by an equal (or greater) number of decreases (McCoy 1977), such as the loss of parts in the evolution of many parasites, the reduction in number of skull bones in vertebrates, and so on. Without an unbiased sample, we cannot say which, if either, predominates. Only very recently have attempts been made to sample in an unbiased way (see below).

Historically, a great many rationales have been offered for why complexity ought to increase in evolution. For example, Rensch (1960a, b; Bonner 1988) argued that complexity should be favored by natural selection, because complex organisms are mechanically more efficient, having more parts and presumably greater division of labor among parts. Waddington (1969; Arthur 1994) suggested that as diversity increases, niches become more complex, and more complex niches are then filled by more complex organisms. Saunders and Ho (1976; Katz 1987) contend that component additions are more likely than deletions, because additions are less likely to disrupt normal function. (Others are reviewed in McShea 1991.) The combined effect of all of these speculations has undoubtedly been to reinforce the impression of a persistent trend. However, none has any solid empirical support.

For most of the history of evolutionary thought, there has been a near consensus on complexity (McShea 1991). Evolutionists who shared the impression of a general trend include Lamarck (1809), Darwin (1887), Cope (1871), Spencer (1893), Huxley (1953), Rensch (1960a, b), Stebbins (1969), Saunders and Ho (1976, 1981), Wake et al. (1986), Bonner (1988), Ayala (1988), Arthur (1984, 1988), Lewin (1992), Valentine et al. (1993), and many others. But in recent decades, some have expressed skepticism, including Williams (1966), Lewontin (1968), Levins and Lewontin (1985), Gould (1985), and Slobodkin (1992), perhaps a sign that the consensus is coming apart.

Rising doubts about trends offers an opportunity now to reformulate concepts and to evaluate the evidence neutrally. The mood of this paper is skeptical, but the point is not to make a case that complexity has not increased. Possibly it has, in some sense. Rather, the point is to rescue the study of biological complexity from a swamp of impressionistic evaluations, biased samples, and theoretical speculations, and to try to place it on more solid empirical ground. The immediate goal is to determine what we can say now and with confidence about trends, and what evidence is still needed.

Trends in Means and Maxima

A preliminary issue concerns the distinctions among trends. A trend might refer to directional change in a single lineage, such as the increase in brain size from *Homo*

habilis to *Homo sapiens*, or in the mean in a diversifying group of lineages, such as the increase in mean brain size in primates. This paper is concerned mainly with trends in the mean, in particular, mean complexity for multicellular animals (metazoans) over the Phanerozoic Era, essentially their entire history. Some closely related topics are not covered, such as trends in plants and trends at the scale of ecosystems.

Also of interest will be a trend in the maximum, in other words, in the complexity of the most complex metazoan. Maxima are of special interest, because a leveling off of the maximum in a diversifying group suggests the presence of a boundary, or an upper limit to complexity. In principle, means and maxima are independent, so here they are treated separately.

Definitions, Evidence, Causes, and Limits

This paper has four parts: 1) First, I propose a narrow definition and a general scheme for understanding complexity. The scheme reveals complexity to be a compound concept encompassing four independent aspects or types. Thus, the question of a trend will ultimately have four answers, not one. 2) Then, I describe some of the measures that have been devised for the four types and review the evidence for trends in each. Some evidence exists for an early-Phanerozoic trend in the metazoan maximum and mean, but only for one type of complexity and only at one scale.

For other types and scales, trends have been documented but only in groups within the Metazoa.

3) Another issue concerns the causes of trends, if trends in fact occurred. The standard explanation has been general tendencies driven by the supposed selective advantages of complex structure and development (Gould 1994). But trends can also occur "passively," even if complexity is not generally advantageous. In this section, I explain these two categories of causes further and consider the evidence for each. 4) Finally, if trends occurred, they may have been short-lived; in other words, limits may have been reached, perhaps long ago. Here I review the available evidence.

DEFINING COMPLEXITY

The word complexity has been applied to various objects and processes, or more generally, systems. The weather, a watch, and the world economy are said to be complex systems. In biology, DNA, the human brain, and rain forests are usually considered complex. In search of a common theme, some broad definitions of complexity have been proposed. For example, the complexity of a system has been defined as the length of the shortest complete description of it (Löfgren 1977; Papentin 1980, 1982) and the length of the shortest algorithm that will generate it (Kolmogorov 1965; Chaitin 1975). (For others, see Bennett 1988, 1990.) Each may be useful in some context. But no broad definition has been

offered that is both operational, in the sense that it indicates unambiguously how to measure complexity in real systems, and universal, in the sense that it can be applied to all systems. What is the length of the shortest algorithm that will generate an actual rain forest?

A Narrower View

Thus, in order to study complexity empirically -- to measure it, test for trends, and investigate its properties -- some evolutionists have adopted a narrower view: the more differentiated a system is, the more complex it is. More precisely, the complexity of a system is some increasing function of the number of different types of parts or interactions it has. The opposite of complexity is simplicity; systems with few types of parts or interactions are simple. Thus, an organism and an automobile are both complex, both having many different parts, while a raindrop and a rubber ball are simple. This narrow view has been advocated by some theoreticians in biology (Hinegardner and Engelberg 1983; Katz 1986, 1988; Kampis and Csányi 1987; Wicken 1987; Slobodkin 1992) and is implicit in recent empirical studies (Cisne 1974; Boyajian and Lutz 1992; McShea 1993; Valentine et al. 1993).

The narrow view is purely "structural," in that complexity depends only on number of different parts and interactions and not on their functionality. Thus, a working automobile and a demolished one, a live organism and a dead

one, and a natural community and a compost heap, are all complex. To many this view will seem wrong-headed, to miss the essence of complexity by failing to distinguish between a functional differentiated structure like an automobile and the (ordinarily) functionless heap of parts in a demolished one. The complaint is serious and will be addressed later. I will argue that *not* making this distinction is *essential* for empirical research.

Order and organization.-- These words have been used casually in biology, and interchangeably with complexity, creating much confusion. Order is especially troublesome, because simple systems like crystals and complex ones like organisms are both said to be highly ordered. Some have suggested that organization should denote degree of functionality (Atlan 1974; Wicken 1979; McShea 1993), independent of complexity, but this usage is not yet widely accepted. For clarity, neither word will be used in this discussion.

Four Types of Complexity

Even narrowly defined, complexity is still a compound term, encompassing four distinct types, based on two dichotomies: object versus process, and hierarchical versus non-hierarchical structure (McShea in press). The four possible combinations of these terms generate the four types: 1) Non-hierarchical object complexity; 2) non-hierarchical process complexity; 3) hierarchical object

complexity; and 4) hierarchical process complexity.

Objects and processes.--Object complexity refers to the number of different physical parts in a system, and process complexity to the number of different interactions among them. For processes, a collision between two billiard balls is simple, while an avalanche is complex. Parts do the interacting, but the interactions can be considered on their own, independent of the parts. Indeed, there is no necessary correlation; one part may participate in essentially one (major) interaction, as does a heart, or many, as does a liver. In Figure 2, A has greater object complexity than B, because it has more different parts. C and D have the same object complexity, but D has greater process complexity.

In biology, the object might be the genome, and the parts genes or nucleotides, or it might be the entire organism (its entire morphology), and the parts cells or organs. Processes might be developmental or physiological, and the component interactions might be morphogenetic events and metabolic steps, respectively. *Here, only morphology will be considered under the heading of object complexity, and only development under process complexity.*

Hierarchical and non-hierarchical structure.-- Hierarchical object complexity is the number of levels of nestedness of parts within wholes. A possible object hierarchy in biology is the series: ... organelle, cell, organ, organism ..., a sequence of what have classically

been called "levels of organization." Upper-level entities or individuals physically contain the lower and constrain their behavior somewhat (Eldredge and Salthe 1984; Salthe 1985, 1993; O'Neil et al. 1986). In Figure 2, E and F appear to have the same number of levels, but in E the upper level (the large circle) is more completely "individuated" (see below) and thus has (fractionally) more levels.

Hierarchical process complexity is the number of levels in a causal specification hierarchy (Salthe 1993). An army chain of command is such a hierarchy, with the highest ranking officers issuing the most general orders, causing the lower ranks to give more specific orders. Likewise, development is (partly) a causal hierarchy (Arthur 1988; Gould 1993; Salthe 1993). (See McShea [in press] for further discussion.) The arrangement of interactions is hierarchical in Figure 2G but non-hierarchical in Figure 2H.

Non-hierarchical complexity is the number of parts or interactions at a given spatial or temporal scale. It is thus a scale-relative property. No scalar level is privileged *a priori* (Salthe 1985), and thus no system has a single true or essential non-hierarchical complexity. In particular, the molecular or genetic level in organisms is no more privileged than any other.

Configurational complexity.--A third dichotomy could be recognized also, differentiation versus configuration. The four types of complexity above are differentiatinal.

Configurational complexity is irregularity of arrangement of parts and interactions, independent of their differentiation (Katz 1986). For example, a parade is typically highly differentiated, consisting of many differently dressed individuals, but configurationally simple, in that individuals march in regular rows. (In Figure 2, A is more differentiated than B, but less complex configurationally.) This dichotomy introduces four more types of complexity, a configurational version of each of the four above. However, the configurational types have received little attention in biology (but see Yagil 1985; McShea 1992), and will not be considered here.

Overall complexity.--Is a human more complex than a trilobite *overall*? The question seems unanswerable in principle because the types of complexity are conceptually independent. The aspects of other measures, such as size, have this same independence: a balloon can be larger than a cannonball in volume but smaller in mass. Likewise, a trilobite might have fewer parts but more interactions among parts. Thus, it is hard to imagine how a useful notion of overall complexity could be devised. This is not to deny that the types might be related empirically. Just as volume is often correlated with mass, so morphological and developmental complexity might turn out to be correlated.

Objections to the Narrow View

Complexity and randomness.--Three senses of randomness

are relevant here. 1) Colloquially, in calling a compost heap random, we might mean that its composition has no functional significance. Accordingly, it might seem appropriate to define complexity so as to exclude systems, or those portions of them, that have no function. However, our ignorance of function is profound in biology, and the attempt to separate the functional from the "random" would be frequently frustrated. We do not know, for example, whether the placement of the human heart on the left or the number of fingers in the hand is functionally significant. One virtue of the narrow view is that complexity can be measured while judgments about function are deferred.

2) The narrow view might seem to equate complexity with entropy (Thomas in press), another sort of randomness. A compost heap is entropic in that a large number of different microstates (possible combinations and configurations of parts) correspond to the same macrostate (the same compost heap). However, calling a compost heap complex for this reason would be a category mistake. Entropy is a relationship between microstates and macrostate, while complexity in the narrow view is a property of a single microstate, of one specific composition and configuration, and involves no such relationship (Wicken 1987). In the narrow view, complexity is not entropy.

3) Some have argued that systems contain both a "regular" and a "random" component (Crutchfield and Young

1989; Crutchfield 1991) and that only the differentiation of the regular portion ought to contribute to complexity. Here, random refers to the unique features of systems, such as the precise number of hairs on the arm of a particular human individual, while regular refers to shared features, such as a five-fingered hand. The intent is to restrict complexity to features that are "rule-based," in other words, to features produced either by natural law acting in the present or by irrevocable, contingent events in the past (frozen accidents) (Gell-Mann 1994). This approach is actually consistent with the narrow view. For example, to identify types of parts is to discern first-order regularities. More concretely, a decision that two cells are the same type can be construed as a decision that their similarities are rule-based and their differences not.

Dependence on descriptive frame.--The number of parts in a system depends on scale, as discussed, but also on descriptive frame, in other words, on how we define a part at a given scale (Kampis and Csányi 1987). The choice may seem arbitrary, which will make complexity measures seem arbitrary as well. One solution might be to define parts functionally, based on the plausible notion that distinctiveness of parts in organisms is the evolutionary result of selection for functionality (Wagner 1995). The problem is that identifying function is difficult, especially for parts that are no longer functional.

Alternatively, parts could be defined as regularities, in the sense discussed. Elegant algorithms have been devised for discovering regularities in bit strings (Crutchfield 1991), and in principle these could be applied to organisms. However, the algorithms make assumptions, such as statistical stationarity, which have not yet been justified for organisms. At present, our best strategy is probably to proceed intuitively, that is, to identify parts and discover regularities using our precognitive perceptual skills. We are obviously sensitive to many cues, such as boundaries and morphological commonalities among parts. We can articulate only some of them, but this does not make the partitioning arbitrary. The complexity studies reviewed below apply this method opportunistically. Scalar levels and organisms (or their substructures) are selected in such a way that the partitioning is intuitively unambiguous.

Differing research agendas.--A view that classifies a demolished car as complex will seem somewhat perverse to those with certain research agendas. In particular, a goal for many students of complexity has been to find the critical structural and dynamical commonalities among systems that are known to be highly functional, in the sense that they are able to self-organize, compute, evolve, and so on. For example, in Boolean networks (Kauffman 1993) and cellular automata (Packard 1988; Langton 1990), high levels of functionality seem to occur when some frozen structure is

present but change occurs as well, in the middle range between monotonous regularity and chaotic irregularity (cf., Mitchell et al. 1993). And accordingly, an appropriate complexity scale would seem to be one on which middle-range systems score highly. In effect, the point of this research agenda is to discover the essence of complexity by investigating the common features of highly functional systems. This is very much science in the exploratory mode.

The approach outlined in this paper supposes a very different agenda. The criteria for complexity are fixed in advance, and the point is to measure the complexity of systems (and thus to *discover* which are complex) and then to test empirically for trends and for relationships with other variables (e.g., stability). In principle, this tactic leaves open the possibility that humans, for example, will prove not to be especially complex relative to other species. However unlikely such a finding may seem, the existence of the in-principle possibility is essential for getting non-trivial answers to questions about trends. In the exploratory agenda, however, the criteria for complexity are sought with the capabilities of highly functional systems such as humans in mind, and thus a trend with humans at or near the zenith is virtually inevitable.

Notice that the narrow, *a priori* definition of complexity advocated in this agenda (or some close analogue of it) is essential for investigating the relation between

complexity and other variables. Is complexity correlated with intelligence? Are complex organisms more evolvable? To answer such questions, we must be able to measure the two variables involved -- complexity and intelligence, or complexity and evolvability -- *independently*. Only then can we plot one against the other and find out if the suspected relationship really exists. This agenda is very much science in the hypothesis-testing mode. Both agendas are worthwhile and they need not conflict.

Complexity and generating processes.--One suggestion has been that the complexity of a system ought to be a function of the process that generated it (e.g., Lloyd and Pagels 1988). In biology, a common notion is that the complexity of an organism is the information content of its DNA, which is assumed to be a kind of generating mechanism. One problem is that much of the information in development is cytoplasmic and not present in DNA. Another is that simple systems can have complex generating mechanisms, and vice versa. For example, mayonnaise is a simple, homogeneous substance (at ordinary scales of observation), but it has a very complex recipe (Rombauer and Becker, 1974). Finally, we would like to learn how complexity of systems and complexity of their generating processes are related empirically, and defining one in terms of the other permanently obscures that relation.

In early studies of DNA complexity, the more "advanced"

organisms (such as humans) were assumed to be more complex, in some unspecified sense, and the expectation was that generating their greater complexity would require more genes and larger genomes (Britten and Davidson 1969; Sparrow et al. 1972). For the most part, expectations have been frustrated. Based on limited data (Cavalier-Smith 1985; Szathmáry and Maynard Smith 1995), the correlation between "advancement" and genome size (also called C value) is poor. The correlation with gene number is better, but current estimates place humans in the same range as lungfish (Szathmáry and Maynard Smith 1995).

The unruly behavior of these variables has been called the "C-value paradox," but for complexity no paradox exists. Indeed, the absence of a correlation between structure and generating mechanism is unsurprising, especially across the wide scalar gap that separates DNA molecules and whole-organism morphology. (Conversely, the possibility should not be brushed aside that lungfish and other high-C-value species are complex in some respect still undiagnosed. Impressionistic assessments of "advancement" could be wildly off.)

EVIDENCE FOR TRENDS

A Research Strategy

For investigating a trend in the mean, ideally a metric applicable to all metazoans would be developed for each type of complexity. However, a (nearly) universal metric has been

devised only for non-hierarchical morphological complexity (see below). The alternative is to develop a variety of metrics, each targeting a specific group *within* the Metazoa. If a trend occurred in the Metazoa, it should emerge as a statistical regularity or bias, that is, significantly more groups should show increases than decreases. (A limitation of this approach is discussed below, under "causes.") The recent studies reviewed below fit neatly into this strategy.

Non-hierarchical Morphological Complexity

Measures.--Various approaches to counting parts are possible. For example, Thomas and Reif (1991, 1993) devised an ingenious classification of design elements. Using their "skeleton space," the complexity of an organism might be just the number of different elements it contains. For greater resolution, parts can be weighted by frequency of occurrence using formulae from information theory (e.g., Gatlin 1972; Cisne 1974). Where variation is continuous and types intergrade, complexity is degree of differentiation among parts, which is measurable using the range of variation, the variance (Bookstein et al. 1985), or a variance analogue (McShea 1992). As a complexity measure, a variance could be construed as a way of counting parts that weights each according to its distinctiveness from a typical part, the mean.

Evidence.--1. Valentine et al. (1993) used cells as parts and measured complexity as number of cell types