The Nature of Prejudice

A Philosophical, Sociological, and Biological Evolutionary Inquiry into the Relationship of External Reality and the Human Brain

Barry R. Zeeberg

Nuns with the Little Sisters of the Poor give a thumbs up at a National Day of Prayer event with President Donald Trump and other religious leaders in the Rose Garden at the White House, in Washington, on Thursday. Trump signed an executive order aimed at easing restrictions on political activity by tax-exempt churches and charities.

https://www.thestar.com/opinion/commentary/2017/05/08/trump-finds-religion-for-political-gain-of-course-coren.html

The Nature of Prejudice

A Philosophical, Sociological, and Biological Evolutionary Inquiry into the Relationship of External Reality and the Human Brain

Barry R. Zeeberg

Revised March 2003

Privately Published 1997

ISBN-13: 978-0-9657219-1-2

The Nature of Prejudice

Barry Zeeberg

The point of departure for this manuscript is the classical book of the same title by Allport (1954). As excited as I was when I first read this book a number of years ago, I now feel that the book was too limited in scope. It is not clear to me whether the present manuscript could have been written at a substantially earlier time, since perhaps the thinking that lead to it is based upon the technology in which our culture is immersed. On the other hand, as an exercise in pure logic it seems that it could have been written at any time, but there might not have been the motivation or inspiration to do so. I can't put my finger directly on it, but there is something here akin to what we would now perceive as Thomas Jefferson's ownership of and patronizing attitude towards slaves.

In any event, it is painfully clear to me that Allport, immersed in the culture in which he was immersed, was taking for granted a particular assumption. This assumption forms the basis of this manuscript. The assumption of which I speak is that the nature of prejudice solely concerns the prejudice of humans towards other humans. In a sense, it is dangerous to write about prejudice, because in so doing, you are very likely to betray your own prejudices in spite of every attempt to be as objective as possible. Unfortunately, Allport's assumption that the nature of prejudice solely concerns the prejudice of humans towards other humans betrays a prejudice on his part, as I will show. I point this out not so much to impugn the merits of Allport's work, but to start off with a disclaimer to show that at least I am aware that what I am about to write, in spite of my best intentions, is probably more a reflection of my own prejudices than it is a reflection of my insight. Furthermore, this will probably be pointed out at some time in the future, unless (which is more likely) this manuscript is ignored altogether.

I will now summarize what I consider to be Allport's main points by means of a fairly extensive section of selected quotations from his book. Although his definitions sometimes refer to prejudice towards objects in addition to people, he appears to deal exclusively with prejudices towards people in the rest of his text.

One of Allport's fundamental concepts is that of category, defined and described (p.171) as

> . . . *an accessible cluster of associated ideas which as a whole has the property of guiding daily adjustments*. Categories, of course, overlap. . . There are combinatorial categories, overlapping, superordinate, and qualified. . . A category, in short, is whatever organizational unit underlies cognitive operations.

Allport (pp. 175,176) summarizes that

> Impressions that are similar, or that occur together, or that are spoken of together, especially if a label is attached . . . tend to cohere into categories (generalizations, concepts).
>
> . . . Categories help us to identify a new object or person, and to expect from it (him) a certain kind of behavior to accord with our preconceptions.

> . . . A rational category is built around the essential or defining attributes of the object. But nonessential and "noisy" attributes often enter into the category, lessening its correspondence to outer reality.

Further properties of categories are (pp. 20-22):

> The human mind must think with the aid of categories (the term is equivalent here to *generalizations*). Once formed, categories are the basis for normal prejudgement. We cannot possibly avoid this process. Orderly living depends upon it. We may say that the process of categorization has five important characteristics.
>
> (1) It forms large classes and clusters for guiding our daily adjustments.
>
> (2) Categorization assimilates as much as it can to the cluster.
>
> (3) The category enables us quickly to identify a related object.
>
> (4) The category saturates all that it contains with the same ideational and emotional flavor.
>
> (5) Categories may be more or less rational.

Allport (p. 27) summarizes the relationship of categorization to prejudice:

> . . . man has a propensity to prejudice. This propensity lies in his normal and natural tendency to form generalizations, concepts, categories whose content represents an oversimplification of his world of experience. His

> rational categories keep close to first-hand experience, but he is able to form irrational categories just as readily. . . .

One definition of prejudice offered by Allport (p. 6) is

> *A feeling, favorable or unfavorable, toward a person or thing, prior to, or not based upon, actual experience*.

Finally (pp. 8,9),

> *Overcategorization* is perhaps the commonest trick of the human mind. Given a thimbleful of facts we rush to make generalizations as large as a tub. . .
>
> There is a natural basis for this tendency. Life is so short, and the demands upon us for practical adjustments so great, that we cannot let our ignorance detain us in our daily transactions. We have to decide whether objects are good or bad by classes. . .

Allport (apparently assuming that the reader of this passage is not a Chinese person, since *Chinese* is unlikely to stand out in the mind of a Chinese person as the symbol of primary potency) emphasizes (p. 179) that

> Most people are unaware of this basic law of language - that every label applied to a given person refers properly only to one aspect of his nature. You may correctly say that a certain man is *human, a philanthropist, a Chinese, a physician, an athlete.* A given person may

> believe all of these; but the chances are that *Chinese* stands out in your mind as the symbol of primary potency. Yet neither this nor any other classificatory label can refer to the whole of a man's nature.
>
> . . . Thus each label we use, especially those of primary potency, distracts our attention from concrete reality. The living, breathing, complex individual - the ultimate unit of human nature - is lost to sight. . . the label magnifies one attribute out of all proportion to its true significance, and masks other important attributes of the individual.

The concept of categorization leads to that of group. Allport does not seem to directly define group, but rather defines in-group (p. 31):

> It is difficult to define an in-group precisely. Perhaps the best that can be done is to say that members of an in-group all use the term *we* with the same essential significance.

Allport (pp. 95 - 104) describes the types and degrees of difference of groups in detail:

> . . . This scheme has the merit of containing within four divisions all the types of group differences that have been established. It likewise enables us to grasp the fundamental logic of group differences. The scheme holds that every known difference between human groups fits into one of the four following types:
>
> 1. A *J*-curve of conformity behavior

2. A rare-zero differential

3. Overlapping "normal" curves of distribution

4. A categorical differential

. . . 1. *J*-curve of conformity behavior. Many groups are marked by the prescription that every member (because he is a member) engage in some particular form of conduct. . . A frequency curve drawn to the histograms looks approximately like the letter "J." . . . The characteristic thing about the *J*-curve is that only the members of a given group can be fitted to it. It is simply not applicable to nonmembers.

. . . The logic of the *J*-curve, then, may be stated as follows: Whenever there is a strongly prescribed action for members of an in-group they will, by virtue of their membership, tend to conform.

. . . The outstanding and obvious differences that mark off one group from another are of this order. . . The law is: *the essential attributes of a group-those characteristics that define the group-tend to follow a J-curve type of distribution.*

. . . 2. Rare-zero differentials. Some traits that are ascribed to a group are actually rare within the group, but they never exist within other groups.

. . . It is obvious that the danger of speaking of rare-zero differential traits lies in the assumption that what is in fact a rare characteristic is actually universal among members of the out-group under consideration.

. . . 3. Overlapping normal curves. Some differences between groups can best be represented in terms of two overlapping curves of the familiar type of "bell-shaped distribution." These are the cases where we know the incidence of a given trait throughout two populations.

. . . we note an almost universal principle in respect to overlapping group differences: *the differences within the same group are greater (i.e., the range is wider) than the differences between the averages of the two groups.*

. . . [4.] **Categorical differentials.** There is one remaining type of quantitative difference. It exists when some single attribute is found with differential frequency in various groups. . . Like the rare-zero differential, the attribute in question is uncommon in either group; but unlike the rare-zero differential, it is actually present in both groups to some extent.

. . . How great does a group difference have to be to be a *true* difference? In most of the sample results presented we notice on the whole rather small differentials. *Probably in no case can it ever be said that a group difference marks off every single member of a group from every single nonmember*. . . There is probably not a single instance where every member of a group has all the characteristics ascribed to his group, nor is there a single characteristic that is typical of every single member of one group and of no other group.

In the case of *J*-curve differences, to be sure, we are dealing with highly *probable* characteristics. In the case of overlapping normal curves the differences are less striking, as a rule. Rare-zero differentials and categorical differentials yield appreciable differences, but their order of magnitude is seldom very great. Strictly speaking, therefore, every statement concerning a "group difference" (unless suitably qualified) is an exaggeration.

> Probably the chief source of error in everyday discussions of the matter is the tendency for people to imply that all group differences follow a *J*-curve tendency. . . Some of these alleged attributes may be entirely fanciful . . . some may be rare-zero, or categorical, differentials. But the implication is that they fall high on a *J*-curve. These traits are thought to be . . . distinctive of the group as a whole. Any stereotype concerning any people is thought to mark the entire group, somewhat in the manner of the *J*-curve, but the ascription is an exaggeration, and may be wholly false.

Visibility and strangeness are important elements of categorization (pp. 129-136).

> . . . Unless there is some visible and conspicuous feature present in a group we have difficulty in forming categories concerning it, also in calling upon the category when we encounter a new member of this group. Visibility and identifiability aid categorization.
>
> . . . perceptible differences are of basic importance in distinguishing between out-group and in-group members. A category needs a visible sign. So urgent is this requirement that visibility is sometimes imagined to exist where it is actually absent.
>
> . . . Where visibility does exist, it is almost always thought to be linked with deeper lying traits than is in fact the case.
>
> . . . This tendency to amalgamate the symbol and the thing it stands for may be called *condensation*. It takes various forms and has many consequences.

> . . . skin color to a white person is a salient feature, as visible as a shooting star, and symbolically important.
>
> . . . Why then should people favored by nature with dark skin be regarded with repugnance rather than with admiration? It is not because of their color but because of their lower status. Their skin implies more than pigmentation, it implies social inferiority. . . on both sides of the fence visibility operates as a vastly important symbol, activating categories that have little to do with the visibility itself.

I suspect that there are forms of prejudice that are more overt, and other forms of prejudice that are extremely subtle. For example, a black person being lynched without a fair trial (or being lynched with a fair trial), or a woman in the early twentieth century being denied access to higher education simply because she was a woman rather than a man are overt forms of prejudice. Another overt form of prejudice is to think or believe or say that a black person is less intelligent than a white person, merely on the basis of skin color and without valid statistical data supporting this assertion. A subtle form of prejudice would be to conduct a technically well-designed study to produce statistical data whose intent was to disprove this assertion. It may seem surprising at first that this apparently well-intentioned effort would be an example of prejudice, yet the grouping of the data sets into "white" and "black" calls into question why the people being tested would be categorized by their skin color in the first place. Why not have the two groups be "has large liver" and "has small liver" instead? This makes about as much sense.

Regardless of my preference and strong desire for this not to be true, it may be the case that all of existence is simply an abstraction of my own mind, a movie, as it were, being projected by my imagination. Of course, if this is true, then there is little point in writing this, except, perhaps, as a "simulation" of what I imagine the real world to be. On the other hand, if this is not true, then there really are other people that exist quite independently of my own imagination. Consequently, there may then be some point in writing this. Furthermore, if there really are other people, then each one would presumably be faced with the same dilemma that I have just posed. Each would be uncertain whether there was a reality beyond his or her own imagination, and would find little to justify assuming that this reality exits. In any event, it is important to realize that each one of us (if there is an "us") is trapped in the dilemma. At least I have a starting point, namely the existence of my own consciousness and awareness, which, in some way, is "more" than an absence of consciousness and awareness. This is reminiscent of the existence of a vacuum in four-dimensional space-time. Even though there is, by definition, nothing within this vacuum, the fact that the vacuum is embedded within four-dimensional space-time is not the same as the vacuum being embedded in nothing. That is, there is a physical infrastructure, even though it is void, and this is different from a void existing (if it could) in the absence of this infrastructure or in a different infrastructure (Atkins, 1992; Barrow and Tipler, 1994).

Thus, I have as a starting point my own consciousness and a decision to make as to whether there is anything else. I do not say this to make myself special; anyone else with a consciousness could equally well say the same thing. From this point on, I will assume that what I observe and experience of outside reality in some way reflects or represents outside reality, and that it is not a projection of my imagination. This assumption is not based on any ascertainable evidence, but it seems that no harm can come of it, at least not a practical sense. For if the assumption is wrong, that is, if there really is nothing else besides my own imagination, then at least there is no one else to be injured by any mistakes I make. On the other hand, if what is sought is absolute truth as a matter of principle, whatever the cost, then this assumption would be completely unwarranted. Also, it seems that there would then be little more to say.

From time to time during much of my life, I have realized the role that a prejudice has played in my life. This is not to say (also it is not to deny) that prejudice has been an important factor in my life or even composes the entire substance of my life, but rather I mean to say something less profound. That is, that occasionally, I have realized that I have had a prejudice that I had previously been unaware of. More precisely, I had realized that I had had it, but I hadn't known that it was a prejudice. Also, I do not intend this passage to be a sort of confessional, but rather it is a point of departure. The particular prejudice I have in mind is the absolute correctness of science over religion, of the rational over the not rational. I feel and have felt that the belief in a religion is an act of faith. For it is difficult to accept that someone is of the opinion that he or she has actually had direct experience of God, and if he or she is of that opinion, then it may be that that person or another may question the sanity of that person.

On the other hand, the belief in science, I now realize, is also an act of faith.

At this point, I will interject a comment concerning what I occasionally perceive to be a rather curious admixture of religion and science. In particular, there have been published attempts by religious adherents to disprove, using (sometimes possibly invalid) scientific methodology, what they would consider to be an incorrect or invalid scientific conclusion (an example of this would be the age of the Earth). Although the motivation on the part of the religious adherent may simply be to correct what they view as a scientific error, it is more likely that the motivation is to form conclusions that have a scientific basis and that support their own religious principles. There are two problems with this approach:

First, the use of scientific methods of research implies the acceptance of science as a valid approach for learning about reality. This would have two consequences. The corpus of scientific knowledge that had not been disproven by the religious adherents would need to be accepted by them as true. Also, if one of their own experiments or interpretations were shown to be incorrect, they would have to accept the alternative, which might very well go against their own religious principles.

Second, it would probably be a better "strategy" for a religious adherent to simply ignore any conclusions of scientists that conflict with their own religious principles; there is no need to confront them. After all, the point for the religious adherent is not to win a debate, but to adhere to the religion, regardless of the (perceived) errors of other people.

But it is perhaps important at this point also to say a few words about science. For each individual, "science" can be partitioned into two sub-parts. This partitioning may be different for different individuals. The first sub-part would contain things like friction and gravity. For example, if we did not have a belief in the existence of friction, we would not only refuse to ride in a car, but we would refuse to cross the street on foot. For without this belief, we would not believe that the brakes of a car would actually stop the car. Further, we believe that, unless the brakes are defective, they will stop the car each time they are used, not just on some statistical basis (say 87% of the time). If we did not believe in the existence of and reproducibility of gravity, then we would not be willing to walk. With each step that we take, we are at risk of falling off the face of the planet and floating off into space. We would have to sink hooks and ropes into the ground at each step we took.

In fact, we do not consciously think about these matters. At least if we do, it is when the mechanic sends us a large bill for repairing our brakes. The second sub-part of science, for most people, would consist of things like positrons, neutrinos, the phase diagrams of igneous rocks, or the types and subtypes of neuroreceptors in the corpus striatum of laboratory rats. Unless we happen to work with any of these in our profession, it is not likely that we have direct experience of them in our everyday life. However, we are presumably willing to believe that they exist and are of some importance, provided that someone mentions them to us who seems reliable to us and does not have a demeanor that would lead us to suspect that he or she is trying to trick us in some way.

What is the essential difference between the two sub-parts of science? The first sub-part consists of phenomena that we do not need to go out of our way to experience, so we are more likely to be comfortable in accepting that they exist (at least if someone called this question to our attention). On the other hand, the second sub-part consists of phenomena that we are not aware of experiencing, and someone would have to go out of the way to experience them. In fact, it is a great understatement to say that someone would have to go out of the way to experience them, since it may be required to undergo a substantial number of years of specialized study to experience some of them. For others, it may require the expenditure of millions or even billions of dollars to be able to provide experimental evidence for their existence or for their elucidation.

To say that the belief in science is an act of faith is to say, at a minimum, that an individual has made a choice, either consciously or otherwise, that the phenomena in the first sub-part are real and are to be relied upon. That is, that we can ride in a car or walk on the sidewalk. It is probably safe to say that once this article of faith is accepted, the phenomena in the second sub-part should also be considered as real and to be relied upon. After all, they are really no different from the phenomena in the first sub-part except that they are beyond our everyday experience, and it is only the fact that the human mind, generally speaking, has a bias (and I don't mean this in any sort of a judgmental way) towards accepting what is more directly experienced rather than what is more abstract. Certainly an abstraction can be turned into a reality when someone contracts a disease and is examined using positron tomography, or, even more dramatically, is treated psychopharmacologically.

To be more precise, faith in science depends upon two factors. The first factor is extrinsic, in that we would believe certain assumptions about reality to be facts. In particular, we expect a certain continuity of external reality, that the existence of a gravitational force between times t_0 and t_1 ensures the existence of a gravitational force at time $t_1 + \varepsilon$; that the transformation of kinetic energy into thermal energy by the brakes of our car the last several thousand times that we applied our brakes implies that kinetic energy will again be transformed into thermal energy by the brakes of our car the next time we apply them. These assumptions are easily accepted as facts pertaining to external reality, and it is only our propensity for the concrete over the abstract (a property of the human mind, since probably only the human mind *can* conceive of the abstract) that prevents this from being applied to an acceptance of the reality of neutrinos, etc.

The second factor is intrinsic, in that we must first assume that there is an external reality that, as discussed above, is not simply a projection of a single mind. But more than this, we must believe that our brain and sensory organs have certain properties *vis a vis* external reality. That is, we must assume that our sensory organs (or actually our sensory organs plus any observational or measuring instrumentation that our technology allows us to develop) provide us with a complete and an accurate picture of reality. Finally, even granting completeness and accuracy of observation, we must assume that our brains process, analyze, interpret, understand, remember, etc. the perceived information accurately.

In contrast to the relatively esoteric train of thought leading to the realization that a belief in science requires a certain degree of faith, it appears that it is more obvious that a belief in religion requires an act of faith. It seems that there is no fundamental difference in the acts of faith required for science and religion. These two areas also have in common another feature in relationship to the human mind. A particular human brain may confuse a strong preference or need for belief in either science or religion as being equivalent to the object of that belief (be it science or religion) as being an external reality. That is, it may seem compelling to my brain (not really mine, just a figure of speech) for its perception of science (or, for that matter, for its perception of religion) to be an external reality. By this, I mean that I would believe that such and such a system of science or such and such a religious doctrine actually exists in the outside world, independently of the existence of my brain. That this system of science or religious doctrine would exist even if the human brain did not exist; that it is not the product of an idea emanating from a human brain. Of course, it could be the case that an integral part of some particular system of science or of religion would be the existence of the human being, and therefore of the human brain. If this is the case, then it would be difficult or impossible for a practitioner of this particular system of science or of religion to imagine the world without human brains and their thoughts being present. The human brain doesn't actually need to be nonexistent, it's just a question of whether such and such exists in external reality or whether it is the product of the human brain or imagination.

To differentiate the concept of a strong belief from the concept of external reality, consider that there are currently a number of different religions in the world, but that there is only one external reality. It is entirely irrelevant that the properties of this external reality may be unknown or only incompletely or inaccurately known to us. Now of the set of different religions, there is at least one nonempty subset of religions that are "mutually exclusive." By "mutually exclusive" I mean that each religion in the subset contains beliefs that contradict the other religions in the subset.

Suppose that the subset of religions is composed of $\{r_0, r_1, \ldots, r_n\}$, and that religion r_i is any religion in this subset. Let the totality of beliefs of religion r_i be represented by the set $\{b_{i,j}\}$. Then, by the hypothesis of mutual exclusivity, there is a member of the set $\{b_{i,j}\}$, represented as $b_{i,c_i(k)}$, such that for any k different from i, $b_{i,c_i(k)}$ contradicts some belief of religion r_k. That is, $c_i(k)$ picks out that value of j for the set $\{b_{i,j}\}$ such that belief $b_{i,j}$ contradicts some belief of religion r_k.

Now assume that there is at least one religion r_{true} in the subset (in fact, there will be either zero or one; we don't know whether there is zero or whether there is one; if there is one, we don't know which one it is; and if there are zero, then the conclusion below is that much stronger) whose beliefs accurately depict external reality. Consider any other religion r_k in the subset. As just shown above, there then exists a belief of religion r_{true}, represented as $b_{true,c_{true}(k)}$, that contradicts some belief of religion r_k. Thus, all the religions in the subset except for r_{true} do not accurately represent external reality. But all the religions in the subset are religions that are strongly believed by some people. Thus, there are some religions that are strongly believed but that do not accurately represent external reality. This means that a strong belief in something should not be confused with that "something" as accurately representing external reality.

Is there in fact at least one subset of actual religions that is in accord with my hypothesis? Well, one such subset would be {Christianity, Judaism, Hinduism}. Since Christianity and Judaism are both based, at least in part, upon the Old Testament, and the Old Testament establishes (almost immediately) a hierarchical arrangement in which humans (not counting God) are at the top of the hierarchy, Christianity and Judaism thus both contradict and are contradicted by the non hierarchical beliefs of Hinduism: According to Zehavi (1973), " . . . all rivers and mountains are sacred, notably the river Ganges. Cattle, monkeys, snakes, and to a lesser extent, certain other animals are sacred, and to an orthodox Hindu their lives are inviolate. . . . in contrast with Western religions, Hinduism holds that all animals have souls, fundamentally similar to those of human beings." Since Christianity is based, at least in part, upon the New Testament, Christianity clearly has beliefs that contradict and are contradicted by Judaism. Now we can hypothesize that the beliefs of one of the religions Christianity, Judaism, and Hinduism accurately represent external reality. Then it follows that the beliefs of the other two religions do not accurately represent external reality. What this amounts to is that hundreds of millions, if not several billions, of people have strong beliefs that are incorrect. I want to emphasize that the purpose of this argument is not to try to convince anyone that what they believe is incorrect. Rather, the purpose is to try to convince people to not assume that their own strong beliefs are an accurate representation of external reality.

Anecdotally, we can consider three people in a room: someone who believes in Christianity (person "C"); someone who believes in Judaism (person "J"); someone who believes in Hinduism (person "H"). Because of the particular beliefs of each religion, each of these three must necessarily consider that the other two have an inaccurate or incorrect perception of external reality. That is, C considers that J and H have an inaccurate or incorrect perception of external reality; J considers that C and H have an inaccurate or incorrect perception of external reality; and H considers that J and C have an inaccurate or incorrect perception of external reality. Furthermore, C knows that J considers that C and H have an inaccurate or incorrect perception of external reality, and C also knows that H considers that J and C have an inaccurate or incorrect perception of external reality. A similar statement is also true for J and C.

This problem can be solved if each of the three were to realize that their own and the other two persons' beliefs, however strong, are an internal abstract reality that seems true for them personally but is not necessarily an accurate representation of external reality. That is, each should realize that the human brain is capable of having strong beliefs about external reality that are, in fact, not accurate representations of external reality. After all, these beliefs did not exist before the existence of the human brain, and we cannot assume that the advent of the human brain and beliefs that it created caused underlying external reality to change.

To clarify this concept, consider three classes of items. The first class consists of those items that certainly exist in external reality, independently of the existence of the human brain. Of course this "certainly" is subject to the assumptions that there is an external reality and that the human brain is capable of some degree of accuracy in perception and interpretation. The second class consists of those items that exist only as abstractions of the human brain, and would "certainly" not exist if there were no human brains. The third class consists of those items where there is a reasonable question whether they would exist if there were no human brains, but for which this question cannot be decided either at the present time or perhaps ever.

A very safe example of the first class would be any item that existed before the human brain had evolved. Contemporary examples may be things like the Earth's moon. A tree would be a good example, except that a human may have grown the tree or may have performed some action that indirectly lead to the existence of the tree, such as cutting down other trees, etc. Perhaps the existence of the core of the Earth would be a good example, since I doubt if humans have had an influence on its existence in any manner whatsoever.

An example of the second class would be this manuscript. I am confident that in the absence of a human brain this manuscript would not have been written. Another example would be the consciously motivated acquisition of material objects beyond that which serves to provide for fundamental survival. There may be animals other than humans that engage in acquisition, but in those cases the activity is genetically controlled behavior, not a consciously motivated effort such as with some humans, where presumably there is the option of a choice or an act of will involved.

The third class would consist of items like the possibility that four-dimensional space-time happened by chance (that is, that we might have been embedded in two spatial dimensions instead of three), the set of real numbers, or the existence of God.

The possibility that four-dimensional space-time happened by chance and what physics and existence would have been like in other numbers of dimensions have been considered in great detail in a surprisingly large number of papers (Atkins, 1992; Barrow and Tipler, 1994; and references therein). This really raises two questions. The first is whether our own three-dimensional spatial dimensionality is actually a subspace or a projection of a larger-dimensional spatial dimensionality, and our senses are such as to only be able to perceive the three-dimensional spatial dimensionality. The second is whether anything material could exist within a two-dimensional spatial dimensionality. I suppose that a corollary to this question would be whether a sentient being living within a four-dimensional spatial dimensionality would have the same questions about whether anything material could exist within a three-dimensional spatial dimensionality.

To reiterate the first question, this is whether our own three-dimensional spatial dimensionality is actually a subspace or a projection of a larger-dimensional spatial dimensionality, and our senses are such as to only be able to perceive the three-dimensional spatial dimensionality. This hypothetical limitation of our senses could have arisen in various ways, depending upon an interpretation based upon faith in science or faith in God. An explanation based upon faith in God would be that, for some reason, God did not want us to have access to whatever it is that is in the fourth dimension.

Two scientific explanations suggest themselves. First, the materials for perceiving the additional dimensions might just not have been available (say, in the Earth's crust or dissolved in sea water) during biological evolution. Second, they might not have conferred a selective advantage for survival at the time when they would have been incorporated, and after that time they might have been incompatible with the other parts of the organisms. I will explain some aspects of evolution and the origins of life below, and that section will clarify what is meant.

To reiterate the second question, this is whether anything material could exist within a two-dimensional spatial dimensionality. Of course, mathematicians and physicists can "simulate" a two-dimensional reality, but I tend to doubt whether this could exist materially. The basis for my doubt is that atoms and subatomic particles (at least in our three-dimensional world) have mass, and even energy has a mass-equivalent (for example, the wave-particle duality of electromagnetic radiation). Ordinarily we think of density as the ratio of an object's mass divided by the volume it occupies. Of course, the density need not be isotropic, and in three-dimensional reality it is easy to envision an object that is very dense vertically and of low density horizontally. However, there are limits to this. In two-dimensional reality a material object would need to have infinite density vertically (that is, in a direction perpendicular to the surface of the two-dimensional space), since this material object is not permitted to have any height whatsoever. So these objects could not be made up of atoms as we know them. If two objects in two-dimensional space were moving towards each other and collided, since they have no height they would simply pass "through" each other and never know that the other one was there (for that matter, each one might not even know that itself was there).

Mathematically, we can define abstract objects that are two-dimensional (such as is done in Euclidean geometry), but other than an abstract triangle, we would need to have a triangle at the minimum printed in ink that was one atom thick, and this would no longer be a two-dimensional triangle, but it would be a three-dimensional triangle. Thus, although an empty two-dimensional space might exist, my question is whether the existence of a material object in two-dimensional space is possible in external reality or whether it is a concept which is the abstract product of the human brain. Phenomena could be defined in two-dimensional space as a mathematical "projection" of phenomena in three-dimensional space, but this procedure would clearly be an activity of the human imagination.

As a simple specific example of the generalization that I will make thereafter, imagine two objects in three-dimensional space that are free to move about independently in any direction. Now imagine a fixed sheet of white paper, which can be either flat (if there is such a thing as "flat" given that space is curved) or curved (that is, it could form the surface of a sphere, or it could have hills and valleys, etc.). Finally, consider a fixed point source of light that is at a distance from the surface of the paper such that the two objects could fit between the paper and the light source.

There are four possible cases to consider. At a given point in time, the positional relationships of the surface of the paper, the objects, and the light source may be such that neither object casts a shadow upon the surface of the paper, only one object casts a shadow upon the surface of the paper, only the other object casts a shadow upon the surface of the paper, or both objects simultaneously cast a shadow upon the surface of the paper. The edges of the shadows may not be sharp, but at least if the darkness of the shadow falls off monotonically as a function of radial distance from the center of the shadow, then we may define the edge of the shadow as being where the darkness is equal to some predefined percentage of the maximum darkness in the center of the shadow. Then we could say that the two objects have "collided" in the two-dimensional space of the surface of the paper if their shadows have collided in the two-dimensional space of the surface of the paper.

A computer program that "knew" the equations of motion of the two objects in three-dimensional space could then determine when the two objects have "collided" in the two-dimensional space of the surface of the paper. Some of these two-dimensional collisions would correspond, in fact, to an actual collision in three-dimensions, whereas others would just correspond to the three-dimensional objects passing by each other between the light source and the surface. In any event, the computer would then have to compute the result of the two-dimensional collision, and change the trajectories of the three-dimensional objects so that the motions of their shadows would be consistent with this result of the two-dimensional collision. It is not clear to me whether there would be a unique pair of such trajectories or whether there would be a set of pairs of trajectories such that the motion of the shadows of the three-dimensional objects would satisfy the "physics" of the two-dimensional collision.

In any event, this is a very unusual situation, since we started with the three-dimensional objects being the "independent" variables which determined phenomena on the two-dimensional surface, but we end with the phenomena on the two-dimensional surface being the "independent" variables that determine the subsequent motion of the real objects in three-dimensional space.

The generalization of this procedure is so obvious as to hardly need mention, except to say that its conceptualization may be easier than its implementation. The generalization would consist of two functions, one for the first object and the second for the second object, whose inputs are the three-dimensional coordinates of the points of the corresponding objects (this may be either an equation of the surface of the object or a rather tedious listing of these coordinates). The output of each function would be a set of numbers (perhaps even a set of ordered pairs or triplets) which would then serve as the input to a third function whose output would indicate whether the two two-dimensional projections had collided or interacted in some other way, and what the two-dimensional outcome of this collision or interaction is. Finally, another function would determine (either by analytical inversion or by iterative simulations) what the two three-dimensional objects need to do in order for the two-dimensional outcome to happen.

It is clear that we lose track of which is the "real" object and "real" space, and which is the "projection" of the real object and the projection of the "real" space. Furthermore, the functions mentioned above must be subjected to certain constraints. For example, when the three-dimensional objects collide in three-dimensional space, we must have the two-dimensional objects collide in two-dimensional space. Otherwise, we would have two material object in three-dimensional space passing through each other. Of course, in a computer simulation there may be no objection to this, and if the three-dimensional space is embedded within a four-dimensional space then perhaps the three-dimensional objects "really" could pass through each other.

At the risk of running off (momentarily) on a slight tangent, the interplay between the two dimensional world and its set of rules and the three dimensional world and its set of rules may have an analogy in the interplay between the genetic encoding within the sequence of nucleotides within the DNA of living cells and the implementation in the real (external) world of the products of the decoded DNA. Hundreds of millions of years of biological evolution have lead to a human brain that "functions" in the real (external) world of the products of the decoded DNA. Now that the technology is becoming available to decode the genetic encoding within the sequence of nucleotides within the DNA of living cells, one might ask the question: why bother to play it out; why not just simulate it? Presumably the answer is that the human brain, being the end product of several billion years of biological evolution, survived as a real object in the real world because of certain properties is has. One of those properties, perhaps one of overriding importance to its survival and existence, is its need to play it out rather than to simulate it. That is, a brain that would be satisfied with simulation rather than with implementation might not survive or exist for very long. In addition to the question "why bother to play it out; why not just simulate it?" one could ask "why even bother to simulate it?"

Returning now to the issue of the set of real numbers, would the set of real numbers exist if the human brain did not exist? Of course, the abstract concept of real number would not exist in the absence of the human brain, since all abstractions are the product of the human brain. However, in a sense real numbers would exist, since a nonhuman animal moving through space would traverse distances which numerically equal real numbers. Also, the passage of an interval of time could presumably be chosen so as to be equal to any given real number. The question that I have difficulty with is whether something like the *set* of real numbers exists or whether this is entirely an abstract product of the human imagination. I would say, at least tentatively, that it is entirely an abstract product of the human imagination, but I could conceive of being convinced otherwise by someone who has a clearer perspective on this.

The issue of the existence of God is somewhat different. While the question of the independent (of the human brain) existence of the set of real numbers may be something that we can only approach philosophically or through logic, God actually either does or does not exist. If God does exist, then those humans who, by possessing religious faith, believe in the existence of God, will turn out to correctly believe in the existence of God. Furthermore, those humans who, by not possessing religious faith, do not believe in the existence of God, will turn out to incorrectly believe in the nonexistence of God. On the other hand, if God does not exist, then those humans who, by possessing religious faith, believe in the existence of God, will turn out to incorrectly believe in the existence of God; while those humans who, by not possessing religious faith, do not believe in the existence of God, will turn out to correctly believe in the nonexistence of God. Of course, if there happened to be no human brains, then this dilemma would not occur, since, whether God exists or not, the question of God's existence would not have arisen.

Although it is not known whether God exists or not, there is a persistence throughout the ages and throughout different cultures of the phenomenon of a belief in the existence of God, particularly the belief of God as an external reality rather than as a product of the human brain. This consistent insistence by the human brain of a belief as being a reality provides an important insight into a facet of the human brain. Consider those religions in which the member of the religion is in a special relationship to God, such as in the Judaeo-Christian family of religions. My hypothesis is that, in this case, the human brain's belief in God is essentially self-serving. The human brain is desperately seeking to make itself feel special or important, and it does this by supposing the existence of a God which is all-important, and by supposing that it has a special relationship with that God. By special, I mean that God presumably considers humans as the real reason for the existence of everything else, so that everything else becomes subordinated to the human brain.

Now consider those religions in which the member of the religion is not explicitly in a special relationship to God. These probably include mainly religions that are not traditional Western religions. For example, perhaps a religion of this type will suppose that all living beings are equivalent in the sense that humans are not superior to other living beings. This type of religion may still be self-serving for the human brain, since the human brain is now creating a class of objects, namely living beings, among which it places itself, and which has a special position in relationship to God. Even a religion of this type cannot possibly be consistent with its principles, for bacteria are living beings and the immune systems of the humans who are members of this religion are killing bacteria. Any pharmacological intervention aimed at "correcting" this inconsistency by knocking out the immune system would lead to the death of all the members of this religion and therefore this religion would cease to exist.

Thus we see that the human brain has a compelling need to think of itself as of particular importance or of occupying a special place. This observation brings us to a critical juncture, and some sort of a decision must be made regarding faith in science and faith in God. The following possibilities seem to be available to us: (1) to assume that faith in science and faith in God are mutually exclusive, or (2) to assume that faith in science and faith in God are potentially compatible with each other.

The possibility that faith in science and faith in God are mutually exclusive would primarily be based upon the ideas that faith in science involves a belief in the accuracy of our perception and interpretation of external reality, and a certain degree of reliability (at least in a statistical sense) upon things like gravity, the transformation of kinetic energy into thermal energy, etc. On the other hand, God is presumably not perceptible to us and could at any moment "change the rules" of what we call "science." A particularly contentious example is provided by human attempts to determine the age of the Earth, or the geological times at which various biological evolutionary or extinction events occurred. If science uses radioactivity measurements in rock and fossil samples to determine their ages, then we implicitly have faith that the (statistical) rate of radioactive decay is invariant over several billion years and that no one "stacked the deck" at the beginning or along the way. God would be just the sort of agent to render these assumptions false. In fact, one could argue that God set things up to look a certain way, perhaps in order to fool scientists or in order to test one's faith in God *versus* faith in one's own perceptions.

Of course, the reason that the age of the Earth is such a contentious example is that scientists have determined an age for the Earth that is much older than some religious fundamentalists would agree to. Every time I see a geology book or a book about evolution that contains a chart of the ages of various geological or evolutionary events, and the chart is invariably presented as fact, I can just imagine the frustration of the religious fundamentalist whose faith in religion exceeds his or her faith in science. Perhaps I am sensitive to their perspective (or at least my perception of their perspective) since I had to deal for most of my life with being a nonsmoker whose perspective was ignored by smokers. I am sensitive to their perspective in spite of the fact that my own belief is that these charts are probably correct.

Another example that comes to mind based upon personal experience occurred several years ago when I participated in a march in the Washington, D.C. area to show my support for gay rights. There was a group of black people who were spectators at the event, and who were protesting the concept of gay rights. I found this to be particularly poignant, since the American heritage of the black people was likely to be one of experiencing discrimination based upon their biological makeup. Yet this very group was unable to relate their own experience of being discriminated against based upon their biological makeup with the experience of the marchers' experience of being discriminated against based upon their biological makeup. In this case, the outward differences between the relevant biological characteristics of the two groups (that is, being black and straight as opposed to being predominantly white and gay), in the minds of the black protesters masked the unity of the experiences of both groups (that is, being discriminated against for whatever reason of human prejudice). One would imagine (obviously incorrectly) that all people for whom a main theme of their lives was being discriminated against for whatever reason would empathize with another group of people undergoing the same type of experience.

It is amusing (to me, at least) to recall that one young man in the march was carrying a sign that read "I love butt fucking" (I don't remember if "butt fucking" was hyphenated, and my word processor's spelling checker doesn't have an entry for this). It seemed to me that if the purpose of the march was to enlist the sympathy and support of the mainstream culture, then this sign could have been edited, since it did not seem well chosen to help overcome mainstream prejudice, although a certain ambiance of directness, simplicity, and honesty seem to have been its virtues. There were also other signs that contained clever puns, such as "get your [I can't remember] out of my Bush." I though of a good sign based upon the name of a nearby Metro station: "Get your Federal out of my Triangle." While trying to expand upon this theme my other efforts were weaker ("Get your Foggy out of my Bottom," "Get your Shady out of my Grove," "Get your Metro out of my Center," etc.).

While on the subject of gay rights, I was disappointed and disturbed when the Commander-in-Chief of the Armed Forces (President Clinton) failed to ensure equal rights for gays in the military. It seemed more appropriate to me that people in the military with a bias against gay people should be given the options of leaving the military or having the opportunity to attend sensitivity and awareness training. In any event, it somehow seems paradoxical to me that gays would want to be in the military. Given the culture we are embedded in, it would appear to be an ennobling act to not hide (either from oneself or from everybody else) the fact of being gay. On the other hand, some of the functions of the military are to use force to ensure the implementation of (often self-serving and certainly always selective) foreign policy, to serve our own commercial interests, to serve special interests, or to serve political interests. It is difficult for me to reconcile the desire of someone who has been able to step outside of the system (often at great personal inconvenience) to be a member of an organization that is so much entrenched in the system.

The possibility that faith in science and faith in God are not mutually exclusive would presumably be based upon the idea that God actually exists in external reality, but follows certain rules and restrictions. A set of rules of this type might be that God gets to create everything and set things in an initial state with certain laws of physics and statistics applying. It may be that the laws of statistics actually follow from the laws of physics, so the last sentence might have been more accurately stated as " . . . God gets to create everything and set things in an initial state with certain laws of physics applying." There would be four possibilities here:

(1) The initial conditions were randomly chosen, without any regard for the eventual outcome, and there is no further intervention or interference by God in what happens.

(2) The initial conditions were chosen so that subsequent events would happen in a certain way, and there is no further intervention or interference by God in what happens, since there would be no need for such.

(3) The initial conditions were either randomly or non randomly chosen, and there is a substantial intervention by God along the way to make things come out a certain way. However, this intervention is carefully chosen so as to be not detectable as a departure from the ordinary laws of science.

(4) Similar to the preceding, but God gets to play with statistics in a certain way. For instance, there are certain events that are extremely improbable but that have a non zero probability. God could make one of these happen or prevent one from happening, which could have a significant outcome. For example, according to studies relating to biological evolution, human chromosome 2 came about as the result of a very unlikely and remarkable fusion of chromosomes 12 and 13 of the chimpanzee, with a consequent reduction in the haploid number of chromosomes from 24 in the chimpanzee to 23 in the human (Marks, 1992).

In all four possibilities save the first, the question then arises (to the human brain) as to whether the eventual existence of humans was God's purpose. Although this is a legitimate question to ask, there are three problems with it which could unduly influence an attempt to seek an answer:

(1) It is potentially a highly self-serving question, and we must be careful that we don't let the temptation of the satisfaction of a positive answer interfere with our attempts to answer this question.

(2) It would be equally valid to ask whether the eventual existence of aardvarks or cabbage or biotite mica was God's purpose. I am not aware of these questions being asked or asserted in the same manner as to whether the eventual existence of humans was God's purpose.

(3) It is a human brain, rather than a rat or a chimpanzee or a cow brain, that is asking the question. For example, every time a rat or a chimpanzee is used in a human biomedical experiment, the rat or chimpanzee brain would probably not be inspired to think of the creation of humans as God's supreme act. Likewise, every time a cow is slaughtered (either in a kosher, glatt kosher, or non kosher manner), she would probably not be inspired to think of the creation of humans as God's supreme act.

Perhaps the most sophisticated attempt at reconciling faith in science and faith in religion is found within a book by John Eccles, 1963 Nobel Laureate in Medicine (Eccles, 1989). Eccles is apparently in a dilemma, in that he is aware of the scientific evidence supporting the biological evolution of the human brain, and yet he has made up his mind, apparently long ago, that God exists in external reality, that God contrived to create humans, and that humans have a special relationship with God. Eccles' bias in this belief has flawed his book in two ways. First, after presenting all the evidence about biological evolution of the human brain, the statements he makes regarding the involvement of God constitute a substantial *non sequitur*. Secondly, it is my feeling that he has abused his privileged position both as author of a published book entitled *Evolution of the Human Brain* and as a Nobel Laureate to put forth his personal bias in an inappropriate forum.

He abused his position, since most people have beliefs about God (or about atheism, etc.) but because they are not sufficiently articulate to publish a book, or because they do not have the recognition that a Nobel Laureate would have, they cannot equally communicate their beliefs publicly. The forum was inappropriate, since the title of the book implies that this was to be a scientific work, not a work containing personal belief presented as fact. It would have been much more appropriate for him to publish these beliefs in a separate book whose title would reflect the non-scientific nature of the contents. It would even be more admirable if Eccles had published this separate book either anonymously or under a pseudonym, so as to not unduly influence readers by his prestige as a scientist and a Nobel Laureate. As an afterthought, I feel that I am not guilty here of these same transgressions, since the title of this does not imply that it is a work of science, and my name (Barry R. Zeeberg) does not enjoy the prestige that would unduly influence readers in the same manner that Eccles' name would.

Let us then leave behind the question concerning God, and proceed now under the assumptions that there is an external reality other than the projections of my own brain, that the human mind is capable of some degree of accurate observation and interpretation of external reality, and that God has not unduly interfered with our scientific researches. These are assumptions that are made in every scientific paper that has ever been published, but it has probably not been explicitly stated in any of them.

The following rather lengthy diversion was inspired by a film that I recently watched in which a person from the future (say at time t_2) was able to be transported back to a past time (say to time t_0). The rules (presumably some fundamental laws of physics) in other fictional works permitted the person from the future to alter past history, starting at time t_0 onwards. However, in this particular film, the rules did not permit the person to alter past history. In addition to these rules, a peculiar coincidence occurred: the person who was transported back to a past time was eventually shot and killed in the past at a time t_1 between t_0 and t_2. Furthermore, since he was transported back in time only by about 25 or 30 years, he was alive as a young child at t_0 and at t_1. At t_1, remarkably, this child was present and witnessed the shooting of his adult self that had been transported back from the future.

Quite apart from the story itself and the unusual coincidence just mentioned, the two facts - namely that a person can be transported back to a past time and that the person is not permitted to alter past history - have some remarkable consequences. Although these consequences are the inevitable result of these two facts, it is easiest to use the story itself to illustrate these abstract consequences in a rather concrete manner.

Because of the fact that the person who had been transported back in time is not permitted to alter past history, the little boy will grow up to become an adult man who will go back in time. This is easily shown using a proof by contradiction. Suppose that the little boy will *not* grow up to become an adult man who will go back in time. Then past history has been altered, which violates one of the rules.

Besides the man who was sent back in time, there were other people alive at time t_2. The fact that the man was sent back in time did not stop the lives of these other people from continuing forward from time t_2. On the other hand, the life of the little boy at time t_1 continued onwards too. In fact, his life continued onwards until he became the man who was sent back in time. There is thus an infinite cycle of events, given in what we would think of as chronological order: (1) A man appears at time t_0. (2) A little boy sees that man killed at time t_1. (3) The little boy grows up to become a man. (4) At time t_2, this man is sent back to the past. (5) The man appears at time t_0. (6) A little boy sees that man killed at time t_1. (7) The little boy grows up to become a man. (8) At time t_2, this man is sent back to the past. (9) The man appears at time t_0. (10) A little boy sees that man killed at time t_1. (11) The little boy grows up to become a man. (12) At time t_2, this man is sent back to the past. It is obvious that events (1) through (4) occur repeatedly, without end (this occurs a countably infinite number of times).

Now, there must be at least two separate universes that co-exist simultaneously. It is obvious that the little boy is growing older in one universe, starting at time t_0. At the same time, the people who were alive at time t_2 are growing older. They must be doing so in another universe than that of the little boy, since in this universe the little boy is now a grow man. But it then follows that there must be a countably infinite number of universes, since each time the cycle of events (1) through (4) occurs, an additional universe is required in order for time t_0 to have a place to occur.

Thus, each time one person is sent back in time, we require an additional countably infinite set of universes to co-exist, unless the time difference for this second person is in a kind of synchrony with that for a previous time traveler (then their cycles could share a common countably infinite set of universes). Now, we can ask whether the co-existence of the countably infinite set of universes is the *result* of the time traveler, or whether this set would have existed anyway. This is like saying that I took a trip to another country, but the country didn't exist until I arrived there. That is, the sole purpose of that country's existence was to give me a place to arrive at; that it wouldn't have been there if I hadn't gone to it. Clearly (or rather, *almost* clearly) this is not how countries work.

But if we take the more reasonable approach that the countably infinite set of universes exists independently of the activities of the time traveler, then there is no reason for nature to have favored the differential between the particular times t_0 and t_2 as being the basis for a countably infinite set of universes. Then, there will be a countably infinite set of universes for each pair of times. Since there are an uncountably infinite number of times (not to mention *pairs* of times), we then have an uncountably infinite number of sets of countably infinite sets of universes.

The most parsimonious treatment of this situation is to think of time as we think of space. That is, all the times that ever did, do, or will exist actually do exist simultaneously. We have no problem thinking of two different places existing at the same time, and we must now think of two different times existing at the same time. Actually, it may be a little more complicated than that, since it is really different universes co-existing, each of which is aging at the same rate, and each of which is offset from some other one by any specified time interval that one cares to choose. Each universe is an exact replica of the others, and events happen absolutely identically in each one. That is, each one represents an "instant playback" or "instant playforward" of the others.

If all of this is actually true, then it is interesting to note that Darwinian evolution resulted in our having the ability to perceive two different spatial locations simultaneously (not counting the effects of special relativity) and to move from one location to another. In stark contrast, Darwinian evolution did *not* result in our having the ability to perceive two different temporal universes simultaneously nor to move between them. Apparently this did not represent a selective advantage as far as survival or existence is concerned.

But then, we might have to redefine what is meant by "survival or existence." Since all of time always exists simultaneously, anything that ever was for one fleeting moment always was and always will be (in one of the universes or another), that conclusion trivializes the entire notion of survival.

Now assume that a person can travel back in time and *is* permitted to alter past history. But this would have to be rejected because of paradoxes that would immediately result. For example, someone going back in time before their own birth and then killing one of their own parents.

So, we seem to be left with two possibilities: either travel back in time is not possible, or else there is an uncountably infinite set of universes somewhere.

First, this hypothesis should be compared with the hypothesis that God exists. Both hypotheses are similar in that either God does exist or does not exist. Likewise, either there is an uncountably infinite set of universes, or there is not an uncountably infinite set of universes. Both hypotheses are again similar in that, at the present time, there seems to be no way of rationally ascertaining the truth or falsehood of either hypothesis. Both hypotheses are again similar in that, regardless of the truth or falsehood of either hypothesis (in external reality), both *concepts* are the abstract product of the human brain. Finally, both hypotheses are again similar in that it seems unlikely that the truth or falsehood of either hypothesis will ever be ascertainable. Thus, it is no less reasonable to believe in the truth of one rather than of the other.

For the sake of nomenclature, let us refer to the universe that we are embedded in as U_0. Let us further refer to the current moment in our own universe as $U_{0,0}$. Then the general reference to an arbitrary universe will be $U_{u,l}$, where u is the index of the universe and l is the local time within universe U_u. Note that both u and l can be any real number ranging from $-\infty$ to $+\infty$. Note also, that $U_{u,l}$ could alternatively be represented as the ordered pair (u, l), a notation which, by analogy to (x, y, z) as a point in a three dimensional spatial space, makes it clear that we now have a point in a two dimensional temporal space. The symbol u is interpreted as being the offset of U_u relative to U_0, and we have the fundamental relationship

$$U_{u',l'} = U_{u,l} \text{ if and only if } u' + l' = u + l$$

Now, if it is true that there is an uncountably infinite set of universes, then there would be two remarkable consequences. First, time is two dimensional rather than one dimensional, and we would be embedded in five dimensional space-time rather than four dimensional space-time. Second, there would seem to be no such thing as free will: For someone embedded within an arbitrary universe at an arbitrary local time, $U_{u,l}$, anything that will happen in the future (say at time l_f) is either just now happening or has already happened in an uncountably infinite number of universes, any one of which is given by $U_{u',l'}$, where $u' + l' \geq u + l_f$. It seems that a minimal definition of free will is that (a) someone embedded within a given universe at an arbitrary local time, $U_{u,l}$, performs action A rather than action B; (b) that it wasn't known in the past (that is, at U_{u,l_p}, where $l_p < l$) whether action A or action B would be performed; and (c) something would be different at a future time U_{u,l_f} (or, more generally, at all future times l'' for $l'' \geq l_f$) if action A rather than action B had happened at $U_{u,l}$.

But it seems to be impossible even for the first two, let alone all three, conditions to hold: The time in the past represented by U_{u,l_p} would correspond to an uncountably infinite number of $U_{u''',l'''}$, where $u''' + l''' = u + l_p$. In other words, $u''' = u + l_p - l'''$. Although u and l_p are "fixed," l''' can be any real number ranging from $-\infty$ to $+\infty$. Thus, u''' can be any real number ranging from $-\infty$ to $+\infty$. So l''' can be chosen such that $u''' = u + l - l_p$. In other words, $u''' + l_p = u + l$. That is, at local time l in $U_{u,l}$, U_{u''',l_p} is experiencing local time l_p. This would contradict (b), since at time l_p it would then be known whether action A or action B is performed at time l. Intuitively, this is as if a woman from the future traveled back to the past and told someone in the past what she knew was about to happen, since in the future what was about to happen was already part of known recorded history. Thus, we have a proof by contradiction that there can be no such thing as free will. Furthermore, all events are predetermined.

It is obvious that regardless of whether this hypothesis is true or false, if the human brain were to *believe* that it is true, then the human brain would likely cease to exist. If I were to believe that all events are predetermined, then I would conclude that whatever I do makes no difference, and any "choice" I make is really predetermined; I only have the illusion of making a choice. Thus, I am absolved of any responsibility for what I do, and I can act irresponsibly. I would only do what gave me pleasure, and if everyone did this we would cease to exist, since most (but not all) of the goods and services that are of use or are essential to our existence are performed by people, not for the pleasure that the task gives them, but rather because of economic repression.

To complete the thoughts concerning two dimensional temporal space, assume for the moment that we are embedded in our own universe U_0, and we are aging along with our universe from the present $U_{0,0}$ to $U_{0,l}$, where $l > 0$. That is, we are moving through time parallel to the x axis. Alternatively, we might be moving vertically through the y axis. That is, we are moving from universe to universe along a vertical line that goes through $U_{0,0}$. In this case, we appear to age by going to $U_{l,0}$, which would represent the same increment in time as would $U_{0,l}$, given $U_{0,0}$ as the starting point in both cases. Horizontal movement along the x axis may possibly not be distinguishable from vertical movement along the y axis either experimentally or in principle. Moving back in time, into the past, would be represented by a diagonal line that goes through $U_{0,0}$. The reflection of this diagonal would be another diagonal line that goes through $U_{0,0}$, representing travel into the future, but at a rate that is faster than the normal rate at which we go into the future (that is, faster than the normal rate at which our universe ages).

The final comment concerning this time issue would involve the fact that when astronomers now (that is, at $U_{0,0}$) observe a spatially distant galaxy (say at a distance D) the light that is seen actually originated a long time ago, at $U_{0,(-D/c)}$, where c is the velocity of light. According to the fundamental relationship given above, this would be equivalent to viewing $U_{(-D/c),0}$; that is, this would be equivalent to a real-time view, at our own local time, of a universe other than our own.

The major stimulus for what follows is a passage in a book by Dawkins (1987; pp. 160-163) in which he explains why the human brain is "bad" at conceptualizing the vast expanse of time available for biological evolution to have occurred. Dawkins looks at the human brain in a particularly useful manner, namely as something that has survived the process of biological evolution. I feel that the correctness of this point, the general lack of awareness or denial of it, and the consequences of it constitute the single most important point in all that I have to say.

To establish the basis for appreciation of this requires a diversion into the standardly accepted scientific (unmixed with religious) view of biological evolution. Normally, biological evolution is treated as a separate question from the origin of life. However, I will treat both topics as a single unified whole. The part dealing with the origin of life is more speculative and represents my own thoughts with no empirical basis whatsoever. I was greatly influenced by two books: *Biogeochemistry an Analysis of Global Change* by William H. Schlesinger (1991) and *Genetic Takeover and the Mineral Origins of Life* by A. G. Cairns-Smith (1987).

We will be concerned with *resources*. Resources are those material things that can either be transformed into something different, that can sustain or participate in the growth of something different, or can participate in the reproduction (that is, the creation of new copies) of something. Resources can be plentiful or scarce, depending in part on what is available through processes not involving the "something different" and in part on the rate at which the "something different" uses the resource. In fact, the *survival* and *existence* of the resource and the "something different" are mutually exclusive, in that the resource is consumed by the creation, growth, or reproduction of the "something different." On the other hand, without the availability of the resource, the "something different" could not be created, grow, or reproduce. All of this is generally true for both the origins of life and for biological evolution. All of this is also generally true for both living and non living resources and (whatever the plural is of) "something different."

We could start with a lion ripping out and consuming the stomach, liver, and heart of a giraffe (with a great deal of attendant warm blood spurting out, which in no way disgusts the lion), but why don't we start elsewhere and leave this as something to look forward to? Instead, we will start with a standing pool of water and assume that the origin of life has not yet happened. The pool of water contains some suspended particulate materials (perhaps clay or sand particles) and some dissolved substances (perhaps leached from the rocks over which the pool stands, and perhaps containing minerals of which clay particles are composed, or containing simple organic molecules). The forces at work here are:

(1) The regular periodic *cycles of sunlight* resulting from cycles of day and night.

(2) The *gradient* of dissolved substances in the pool of water, which resulted from the rate of dissolution from or precipitation onto the surrounding rocks, the diffusion of the dissolved substances in random directions starting at the point of dissolution, and the consumption/release of the dissolved substances by chemical compounds or non covalent complexes that are composed, at least in part, of the dissolved substances or components of the dissolved substances and that are located throughout the pool either in random or non random locations. For example, assume that we have some limestone (solid $CaCO_3$) in the rock in contact with the pool of water at location X. The monolayer of water immediately adjacent to the limestone at X may very well contain dissolved $CaCO_3$ at such a high concentration that it is essentially saturated with $CaCO_3$. This monolayer of water is undergoing dynamic processes of redepositing $CaCO_3$ onto the rock, dissolving $CaCO_3$ from the rock, and serving as a reservoir and receptacle of $CaCO_3$ for the next monolayer of water.

(3) The *dynamic equilibrium* between the dissolved substances and chemical compounds or non covalent complexes that are composed, at least in part, of the dissolved substances or components of the dissolved substances and that are located throughout the pool either in random or non random locations.

(4) The *shielding* from the sunlight of substances that are at various depths beneath the surface of the water and that are beneath other substances that absorb portions of the sunlight's spectrum.

(5) The sinking and rising of the substances according to their *buoyant density*.

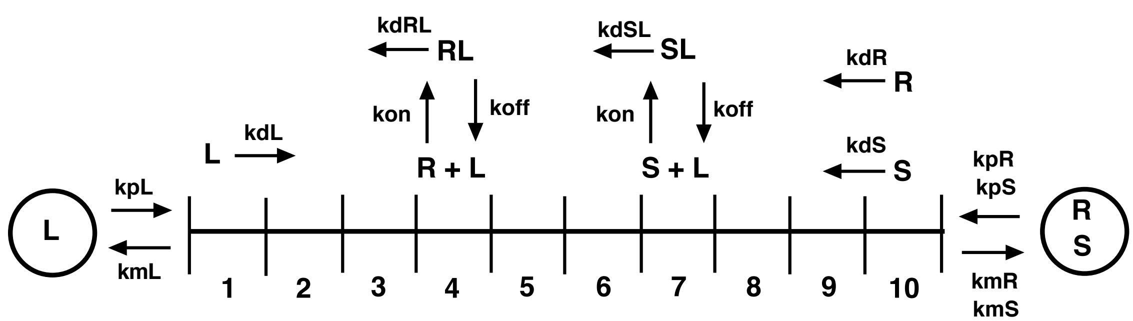
The net result of all of this could be to set up a ritualistic rearrangement of substances within the pool with a periodicity resonating with that of the prevailing periodicity of the cycles of sunlight and darkness. One interesting possibility is that the formation of a certain compound takes place near the surface of the pool of water because the substances that react to make this compound are plentiful near the surface (their buoyant density keeps them afloat, they are not decomposed by ultraviolet light, etc.), and ultraviolet light is required for the reaction to take place. But this compound, once formed, sinks to the middle of the pool of water because of its buoyant density.

This is fortunate, since it turns out that this compound would be decomposed by ultraviolet light and, thus, could not exist near the surface of the pool of water. I use the word "fortunate" ironically, since there is nothing fortunate or unfortunate about the existence or the nonexistence of this compound, as opposed to the existence or nonexistence (as independent entities) of the materials of which this compound is composed. It is only the human brain that would conceive of such an abstraction. That is, if the human brain did not exist, the ideas of "fortunate" or "unfortunate" would not exist; these are merely abstractions of the human brain. For that one compound that survived, there may be a million (or a billion) other compounds that didn't survive because they didn't have the buoyant density that would make them sink to the requisite depth, and they were immediately destroyed by the ultraviolet light from the sun. A human brain watching a time-lapse movie of this process would possibly find it satisfying when one compound "finally" "succeeded" in existing. Equally possible, a human brain watching a time-lapse movie of this process would possibly find it satisfying when the components of a potential compound "succeeded" in existing by "escaping" from being "trapped" by forming the compound. In the absence of a human brain, whichever happened would just simply happen, without the superimposition of an abstraction such as "success" or "failure."

The essence of evolution, be it prebiotic, biological, or molecular, is *competition*. "Competition" is probably a bad word to use, since it seems to connote a willful attempt to succeed or do better than something else. But, as it used here, it is not willful, it is not an attempt, there is no such thing as succeeding, and there is no such thing as doing better than something else. These are all abstract conceptions of the human brain which would not exist if the human brain did not exist. However, I will use the word "competition," but stripped of these connotations. For our specific example, I will use a variation on the theme discussed above. This would actually provide a very interesting (to the human brain) opportunity for performing either computer simulations or, if possible, for formulating an analytical solution.

Consider a pond of water with its now-familiar gradient of dissolved materials. For simplicity, let us consider a pond of water that is one-dimensional. We are only concerned with its length; that is, its depth and width are irrelevant. Consider a rock with a deposit of a material designated as "L," which forms one of the boundaries of the pond of water at its "left hand" end (Figure 1). Also consider a second rock with deposits of two materials, one designated as "R" and the other as "S," which forms the other boundary of the pond of water at its "right hand" end.

Figure 1



The length of pond is conceptually divided into ten equal lengths, starting at segment 1 adjacent to the rock containing L and ending with segment 10 adjacent to the rock containing R and S. L can be leached from its rock at a rate characterized by a rate constant kpL (the "p" stands for "plus," since this process adds L to the pond). Conversely, L can be deposited onto its rock at a rate characterized by a rate constant kmL (the "m" stands for "minus," since this process removes L from the pond). The corresponding rate constants are defined at the other rock for R and S. These rate constants are not just symbols, but have numerical values. Once these materials are dissolved in the pond, they can do several things. For one, they can diffuse randomly to the right or left. The rate constant for the diffusion is designated as "kdL," for example, for L, and similarly for the other materials. For another, they can react chemically to form new materials (kon), and these new materials can diffuse (kdRL) or they can revert back to the materials from which they were made (koff).

Technically, a set of 50 equations (known as differential equations) are formulated from this schematic. These 50 equations were solved numerically using standard methods (Press *et al.*, 1992) with various values assigned to the 13 different rate constants, as listed in Table 1.

These rate constants are given in arbitrary units, as are the distances between the two rocks in Figure 1 and the time points at which the evaluations were made. The time was treated as a continuous variable, whereas the distances along the gradient were treated as discrete variables. The procedure involves assuming that there is a "starting time" at which there is no L, R, or S dissolved in the pond, and then determining the amount of L, R, S, RL, and SL at each of the ten positions in the pond at a sequence of times. The strategy was to use the same kon for both R and S and the same koff for both RL and SL. The reason was that I wanted to show that, although R and S have identical chemical properties with regard to reaction with L, slightly different values for the diffusion rate constants for R and S can lead to the predominance of RL over SL (Figure 2).

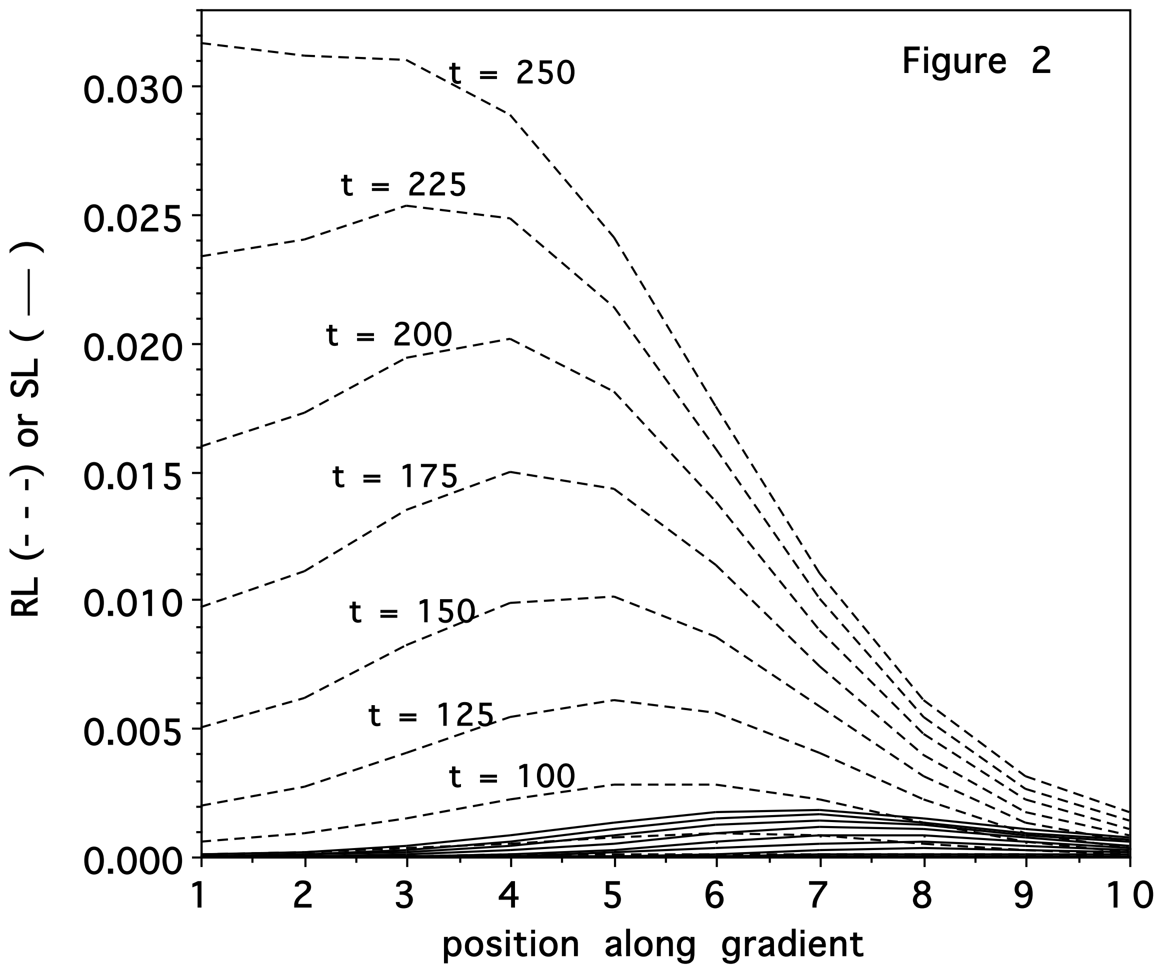
Table 1

Rate Constants used in the Simulations in Figures 2-4

rate constant	numerical value (arbitrary units)
kpL	0.001
kmL	0.001
kpR	0.100
kmR	0.100
kpS	0.100
kmS	0.100
kon	0.300
koff	0.001
kdL	0.100
kdR	0.050
kdS	0.010
kdRL	0.010
kdSL	0.010

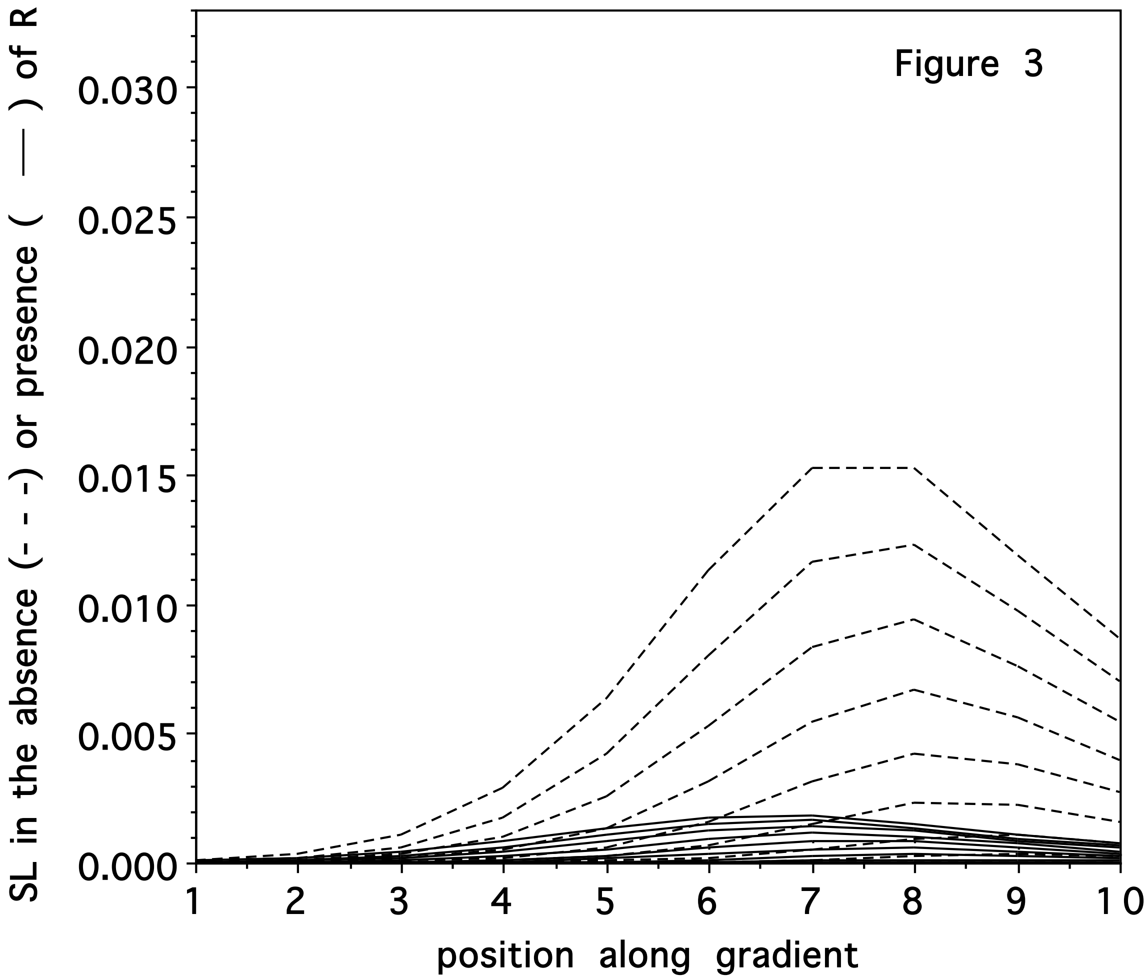
The time points at which the evaluations were made are (in arbitrary units): 25, 50, 75, 100, 125, 150, 175, 200, 225, and 250.

Seven of the ten times at which evaluations were made are indicated in Figure 2 for RL. The curves for the other three times for RL and all ten times for SL are shown in Figure 2, but the times are not indicated because of a lack of space at the bottom of the figure. However, the order of the times continues in the same sequence (that is, longer times at the top) for all the curves. This is also true in the

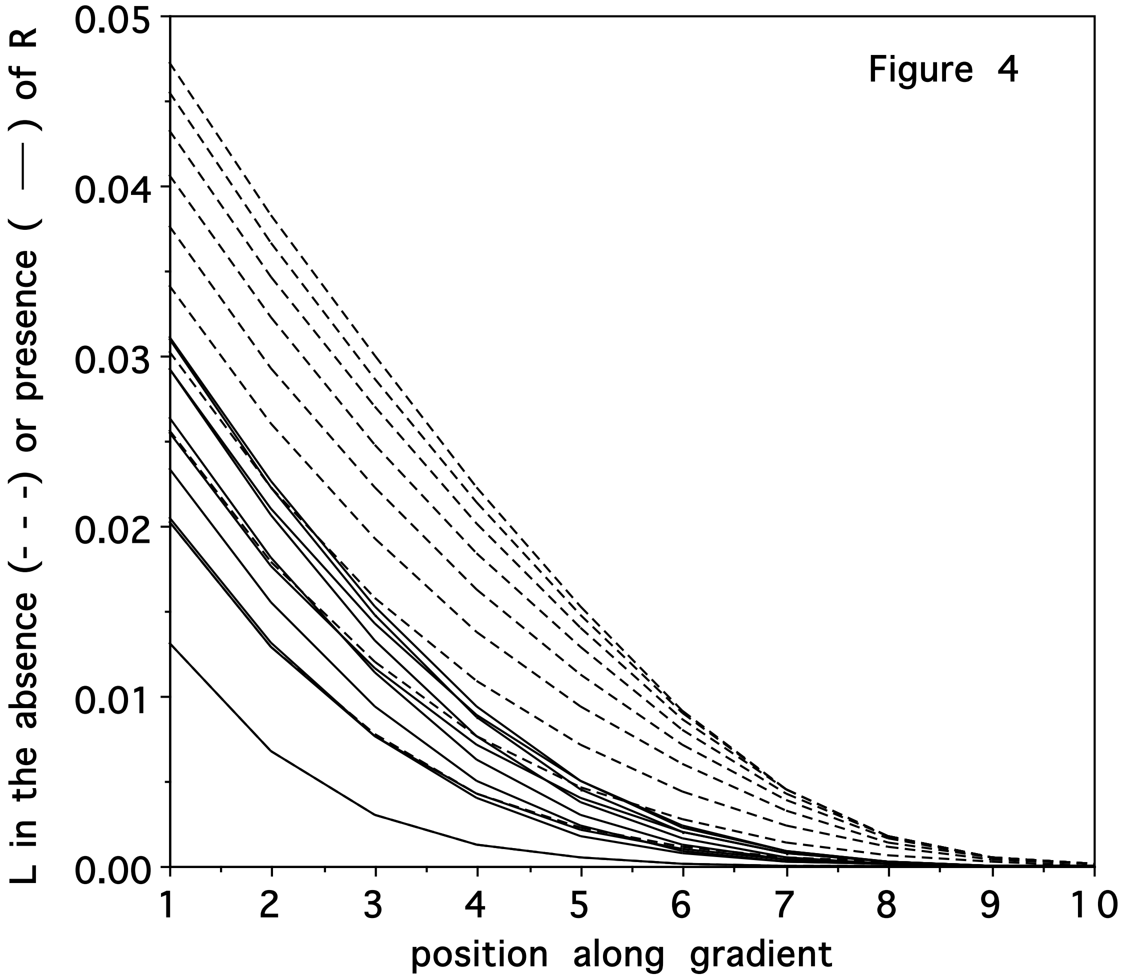


numerical values of the 13 different rate constants leads to the result in Figure 2. This result requires a rather special balance between these 13 rate constants (Table 1). On the other hand, it only took about an hour of trial and error to determine these rate constants, and there are likely to be many sets of other numerical values that would also yield interesting results.

The remarkable feature of Figure 2 is that at all times studied, RL greatly predominates over SL. In fact SL is hard to see (SL is indicated by solid lines that are scrunched together in the bottom right hand half of Figure 2). It is the somewhat greater diffusion rate constant for R as compared with S that causes RL to predominate. If R is absent, then there is a respectable amount of SL, as shown in Figure 3 which was drawn to exactly the same scale as Figure 2 in order to facilitate comparison. Figure 3 again contains the lines for SL in the presence of R (the same as in Figure 2) to emphasize the increase in SL when R is not present.



From the point of view of RL and SL, L is a *limiting resource*. In a way, R and S are competing for L. This is seen readily in Figure 4, where, in the absence of R, there is abundant L. This abundance of L results in the survival of substantial amounts of SL, as we had seen in Figure 3. In contrast, in the presence of R, L is *consumed* in the formation of RL, resulting in a diminshed amount of L. Again, this is seen readily in Figure 4, where, in the presence of R, there is a decrease in L. This decrease in L results in the insubstantial amounts of SL, as we had seen in Figures 2 and 3.



From the point of view of RL and SL, L is a limiting resource. In a way, R and S are competing for L, although we must be careful with our semantics here: since R is *more successful* at competing for L, R is converted to RL at the expense of R as well as at the expense of L. Thus, R is *less successful* at its own survival than is S. This is also a place where a *bias* may creep into our thinking: since we have focused on the competition between RL and SL, we may have lost sight of the fact that L is being lost whenever RL or SL is formed. Thus, the greater the success of RL and SL, the less the success of L at survival.

There are a number of considerations and ramifications relating to this example. First of all, this example is clearly what would be termed "prebiotic," since there is no living organism involved here. Yet this example illustrates many of the characteristics of biological evolution. There is a competition for a limiting resource, and the outcome of the competition determines an eventual survivor. If S had been around for a short while before R came into the scene, then the effect of the new arrival of R would be the *extinction* of SL. An *intrinsic property* of R, namely the faster diffusion rate constant for R than for S, confers a *selective advantage* for the survival of RL rather than SL.

The fact that the faster diffusion rate constant for R than for S confers a selective advantage is entirely dependent upon the *environment* in which R, S, RL, and SL are *embedded*. As a rather trivial example, if there were no L, then there would be no RL or SL no matter what the diffusion rate constant for R and for S happen to be. If the pool of water evaporated, diffusion would not be possible. If the temperature of the pool of water changed, then the diffusion rate constants might be subject to differential changes; this is also true of all the other rate constants, including kon and koff, which might then differ so that we now had to define the additional rate constants konR, konS, koffRL, and KoffSL. Another material T might come along and have the same effect upon R that the incursion of R would have had upon S.

Thus, in a way, the diffusion rate constants for R and for S are analogous to different *coding sequences* within DNA, when one considers that the *information content* of the coding sequences is entirely dependent upon the environment in which the DNA is embedded. This statement actually holds for two different levels of embedding. First, there must be the *internal cellular machinery* to decode the information content of the coding sequences and to implement the synthesis of the corresponding RNA or the corresponding mRNA and protein. This is analogous to there being water in the pool through which R and S can diffuse. Secondly, there must be an *external environment* in which the result of the synthesis of the corresponding RNA or the corresponding mRNA and protein can function: a digestive enzyme produced by a given organism, however much more efficient it is than the corresponding digestive enzyme produced by a different organism, is of no consequence if there is no food present to digest. This is analogous to there being no L present in the pool of water for R to react with in spite of R's prodigious diffusion abilities. Returning to the digestive enzyme example, if food is plentiful and is therefore not a limiting resource, then an organism with a run of the mill digestive enzyme may not fare any worse than the organism with the especially efficient digestive enzyme. This is analogous to there being an excess of L present in the pool of water for R and S to react with. In this case, S may form SL nearly as well as R forms RL, again in spite of R's prodigious diffusion abilities.

Returning now to the original situation depicted in Figure 2, let us assume that, of the possible *external environmental changes* that might occur, another material U comes along. Now U has the inherent characteristics of potentially reacting with RL to form RLU or of potentially reacting with SL to form SLU. Since there is virtually no SL available (Figure 2), U in fact reacts with RL to form RLU. Now assume that RLU has a great deal of stability; that is, there is little or no propensity to reconvert to RL + U. Let us assume that, unlike the diffusion rate constants, as described above, this stability of RLU is unchanged at all temperatures that are normally encountered. Let us further assume that U is present in limiting amounts, so that it is entirely consumed in the formation of RLU.

Now let us assume that the temperature eventually changes to a new temperature at which the diffusion rate constants for R, S, RL, and SL are not significantly different from each other. Recall that at this point, RLU is present, but SLU is not present, because when U appeared, RL was present but SL was not (Figure 2). However, if the temperature change had occurred before U appeared, then it would be a different story. As a result of the temperature change, as described above, both RL and SL would have been present, and both RLU and SLU would have been formed.

This example illustrates an important feature of the scientific view of biological evolution as compared with the religious view. In the scientific view of biological evolution, survival and extinction are based upon the *relationship* between the intrinsic properties, which can also be referred to as the "information content" (in the case of R and S, the difference in the diffusion rate constants; in the case of an organism, the difference in the coding sequences of DNA), the *intrinsic decoding mechanism* (in the case of R and S, the presence of water of a certain temperature in the pool; in the case of an organism, the biological machinery to implement the synthesis of the corresponding RNA or the corresponding mRNA and protein), and the *extrinsic environmental factors* (in the case of R and S, the presence of L in the pool; in the case of an organism, the presence of food, etc.). What I want to emphasize is the importance of *time* in this relationship: survival and extinction are based upon the relationship between the intrinsic properties, the decoding mechanism, and the extrinsic environmental factors at a given point of time.

RL is in a dynamic relationship with R and L, governed by the rate constants kon and koff (Figure 1). Thus, hypothetically, if one particular molecule of R is observed for a period of time, it would shuttle back and forth between being R and being the "R" component of various molecules of RL. Conversely, a particular RL may only exist momentarily, reverting back to the R and the L from which it was formed. This consideration raises the important question of what we mean when we talk about the survival or existence of an entity, since it is not really one particular RL that survives (for example, in the peak of the RL curve of Figure 2), but it is really a collection of RL's that survive, not necessarily the same ones from moment to moment. The same is true for a living object, since the components of that object are constantly "turning over" in processes of loss and renewal. The same is true for a nonliving object, such as rocks and components of the atmosphere (Schlesinger, 1991). Thus, an object that survives or exists is actually in a state of flux or change, although to us it may appear to be unchanging. We tend to ordinarily use either too coarse or too fine a scale (either temporal or spatial) for viewing objects that exist to be aware of the flux, and so we have the illusion of constancy.

When we speak of an object as surviving or existing, we really mean that a sort of dynamic infrastructure in a state of flux survives or exists.

This last point was rather technical (or perhaps philosophical), but the most important point of the RL and SL example is that, with the exception of the diffusion rate constant, everything about R and S is the same (Table 1). That is, the numerical values for kon and koff are the same, the numerical values for kpR and kpS are the same, the numerical values for kmR and kmS are the same, and even the numerical values for the diffusion rate constants for their products RL and RS are the same. Whereas the ratio koff/kon can be thought of in thermodynamic terms as representing what is called "free energy," which is a high-quality form of energy and which could potentially be directed towards accomplishing something useful, the diffusion rate constants represent the *random thermal motion* of matter, which is the lowest quality and most useless energy in physical terms. The ratio koff/kon is like a gallon of premium gasoline; the diffusion rate constants are like the friction that is dissipated when you have applied the brakes on your car. The remarkable thing is that the difference in the diffusion rate constants for R and S result in the survival or existence of RL and the extinction of SL.

I regard the "low quality" random thermal motion (or, more, precisely, the differential rate constants thereof) to have been the form of information that originated the prebiotic processes that eventually evolved into life. As noted above, the relationship between the intrinsic properties or information content (the difference in the diffusion rate constants), the intrinsic decoding mechanism (the presence of water of a certain temperature in the pool), and the extrinsic environmental factors (the presence of L in the pool) compose a *triad* that is associated with survival or extinction. Survival and extinction depend upon the ability to form structure out of chaos. Random thermal motion represents the ultimate in chaos. Yet there is a structure in the natural environment that is used to derive a structure out of the chaos of the random thermal motion. That structure consists of the free energy of binding of two substances and the arrangement in space of concentration gradients.

In contrast to the low quality of random thermal motion is gravitational energy, which is the form of energy that is considered to be the highest quality (Atkins, 1992). A particularly striking example of the above-mentioned triad is associated with the formation of the Earth's core, which is composed largely of iron (Hartmann and Miller, 1991). The primitive Earth was more or less homogeneous; the molten iron throughout the Earth percolating through to the center of the Earth. In this case, the triad consists of the intrinsic property of iron (that is, the high buoyant density in its molten form), the intrinsic decoding mechanism (Earth's gravitational field acting at a point located at its center of gravity), and the extrinsic environmental factors (the existence of molten substances with a lower buoyant density than iron's). In the cases of both the chemical gradient and the molten iron, a system that was initially chaotic was transformed into a more ordered system.

At this point, it may be appropriate to point out that my use of the words "chaotic" and "ordered" represent a kind of prejudice of the human brain. In the case of the chemical gradients, the human mind perceives the peak as a phenomenon representing an orderliness. The same is true of molten iron concentrated in one spot at the Earth's core as opposed to its being rather homogeneously distributed throughout the entire Earth. There are two levels of prejudice at work here. First is the human brain's recognition of something that it is "prewired" to detect as comprehensible. This form of "orderliness" may very well be the result of the intrinsic properties of the human brain (which is the product of its own biological evolution) rather than the "objective" properties of the external reality.

An interesting anecdote related to this "prewiring" can be put into the form of a story about four people playing bridge, although a version about an engineer, a mathematician, and a computer scientist, etc. type of story is probably possible. The dealer had all thirteen hearts, the left hand opponent had all thirteen clubs, the partner had all thirteen diamonds, and the right hand opponent had all thirteen spades. The left hand opponent, bidding first, simply passed, reasoning that when one hand is distributed so unevenly, the other hands are likely to be distributed unevenly also. Since clubs is the lowest ranking suit, what was the sense of bidding since someone else was likely to have thirteen cards of a higher ranking suit? The partner, reasoning somewhat differently and having a higher ranking suit than clubs, smoothly bid one diamond. The right hand opponent stood up, and with a flourish exposed all thirteen spades, expressing an awareness of the impropriety of this behavior, but unable to contain the excitement of so improbable a hand. The dealer, rather bored, just barely audibly murmured that the hand was no more improbable than any other hand whatsoever.

Since the exposure of the right hand opponent's cards resulted in a misdeal being declared, a redeal was called for. This time, the cards were distributed among the four players such that both the suit and high card distributions were about as even as possible. After three passes, the dealer stood up, and with a flourish exposed all thirteen cards, expressing an awareness of the impropriety of this behavior, but unable to contain the excitement of so improbable a hand. The dealer, rather excitedly, explained that the probability of his particular hand was approximately 1 in 2.5×10^{22}!

The second level of prejudice is that, even if the first prejudice was not a prejudice and there was in fact an objective meaning to the "chaotic" and "ordered" situations, the human brain, viewing the "chaotic" situation, would not see the potential "orderliness," that will occur after some period of time, as being part of a system that is destined to become orderly. The human brain will isolate the current view of the situation; it will freeze the current view in the current time. Whereas the human brain has a visual field of view that can encompass and integrate more than a single point of space, it does not have a temporal field of view that can encompass and integrate the "chaotic" and the "orderly" situations that will occur in a sequence of time. Some human brains might be able to do this by a process akin to computer simulation, but this is not one of our "senses."

The link between the prebiotic and the biotic is the existence of *polymers*. The transition from the prebiotic to the biotic (that is, the origin of life) was likely to involve particular polymers, although not necessarily (as a matter of fact, probably not) the same ones that are important (RNA, DNA, proteins, lipids, carbohydrates, cytoskeletal components) in living things today. In its simplest form, a polymer is a linear arrangement of identical *subunits* that are oriented identically and chemically bonded (either covalently or noncovalently) together within the polymer. That is, a subunit has a "left" side that can join with the "right" side of another identical subunit. This polymer, in its simplest form, has two ends, each of which is available for joining with another one of its subunits.

For this simplest kind of polymer, there are exactly nine variables, each having its own numerical value, that fully define the triad. These variables are the initial length of the polymer, the time interval (t), the diffusion rate constant for the subunits (kd), the association and dissociation rate constants for the subunits at one end (kp1 and km1), the association and dissociation rate constants for the subunits at the other end (kp2 and km2), the concentration of the subunits in the solution, and the length that one subunit adds to the polymer. In this example, from the perspective of the survival of the polymer, we again have the triad: the intrinsic properties or information content (the four rate constants kp1, km1, kp2, km2), the intrinsic decoding mechanism (the presence of water of a certain temperature in the pool), and the extrinsic environmental factors (the concentration of the subunits and their diffusion rate constants kd).

Before presenting the results of several computer simulation studies relating to the survival or existence of the polymer, I will provide some perspective on what these results will show. In effect, I have only performed several preliminary simulation studies to convince myself that the program works properly, and that there is a reasonable chance that various selections of numerical values for the set of nine parameters will provide a spectrum of behaviors. Since the initial length will always be fixed at one unit in these studies, there will actually only be eight variable parameters, allowing me to refer to a particular set of parameters as an "octet." A subset of the octet is the "quartet" consisting of the four rate constants kp1, km1, kp2, km2. Basically, what the simulation does is this: Given the octet, it gives the final position of the two polymer ends and the concentration of the subunits in each position in the solution surrounding the polymer.

What I would like to do is to show that for a particular octet, changes in the quartet can lead to either survival or extinction of the polymer. Furthermore, if one end of the polymer is very good at adding subunits (for example, if kp1 is sufficiently greater than km1), and the value of kd is not terribly high, then the portion of the solution that is surrounding that end will become depleted of subunits. There are then two possible scenarios. The first is that this polymer end is starved for subunits and is "stalled" at this position. The second is that this polymer end grows past the portion of the solution that is depleted of subunits, and will be located in a new portion of the solution that is not yet depleted of subunits; this would represent a vigorous growth of this end of the polymer as it is continuously depleting the local portion of the solution and then moving on to more "fertile" grounds and converting them to more "infertile" grounds. Of course, it is possible that the other end is losing subunits so quickly that it may overtake the first end and the polymer will become extinct. This second end can be thought of as "fertilizing" the solution.

Although the problem of a polymer in a solution is inherently three dimensional, it can be reduced to a two dimensional problem by taking advantage of the radial symmetry about an axis running through the center of the polymer along its length. Then a single half-plane can be studied. The half-plane extends from this axis of symmetry to "infinity" in the y direction ("radially"), and from minus "infinity" to plus "infinity" in the x direction ("axially"). "Infinity" can be defined in practice as a distance that is sufficiently great that there is no perturbation of the subunit concentration along the corresponding three edges of the half plane, and that the polymer ends did not approach the two edges in the x direction too closely. The entire half-plane is divided into 41 (axial direction) x 20 (radial direction) squares, as follows: The initial polymer is assumed to be of length one, and extends in the axial direction.

An example will make these rather abstract ideas more concrete. Let us start with a polymer both of whose ends have identical kinetic properties (that is, kp1 = kp2 and km1 = km2), an arbitrary initial subunit concentration equal to 100, and a numerical value for kd that was determined by trial and error to produce an attractive subunit gradient for demonstration purposes. The detailed numerical values for these and for the following related examples are tabulated in Table 2.

Figure 5 demonstrates a prototypical example of the representation of the subunit gradient. In order to avoid an almost inevitable confusion about what this representation represents, I will belabor the point. Although this appears to be a three dimensional representation, it is, in fact a two dimensional representation. The x

Table 2

Rate Constants used in the Simulations in Figures 5-20

figure	description	time	subunit diffusion rate constant	end 1 rate constants assoc	end 1 rate constants dissoc	end 2 rate constants assoc	end 2 rate constants dissoc	subunit concentration /polymer length	final subunit ends
5	prototype	1	10	20	1000	20	1000	1000	±1.533868
6	prototype	3	10	20	1000	20	1000	1000	±2.453449
7	prototype	10	10	20	1000	20	1000	1000	±5.490171
8	faster diffusion	1	20	20	1000	20	1000	1000	±1.662907
9	faster diffusion	3	20	20	1000	20	1000	1000	±2.877762
10	faster diffusion	10	20	20	1000	20	1000	1000	±6.753915
11	very fast diffusion	3	100	20	1000	20	1000	1000	±3.536163
12	very fast diffusion	10	100	20	1000	20	1000	1000	±8.374853
13	zero diffusion	1,3,10	0	20	1000	20	1000	1000	±1.050000
14	zero diffusion/ lower sub/poly length	3	0	20	1000	20	1000	49	±16.74773
15	treadmilling	1	10	20	1000	5	1000	1000	-1.685877 +0.526171
16	treadmilling	3	10	20	1000	5	1000	1000	-2.954939 -0.402457
17	treadmilling	10	10	20	1000	5	1000	1000	-7.081885 -3.567796
18	unidirectional	0 - 7	10	20	0	0	1000	1000	see fig 19
19	unidirectional	0 - 7	10	20	0	0	1000	1000	see fig 19
20	excretion	0 - 7	10	20	0	0	1000	1000	see fig 20

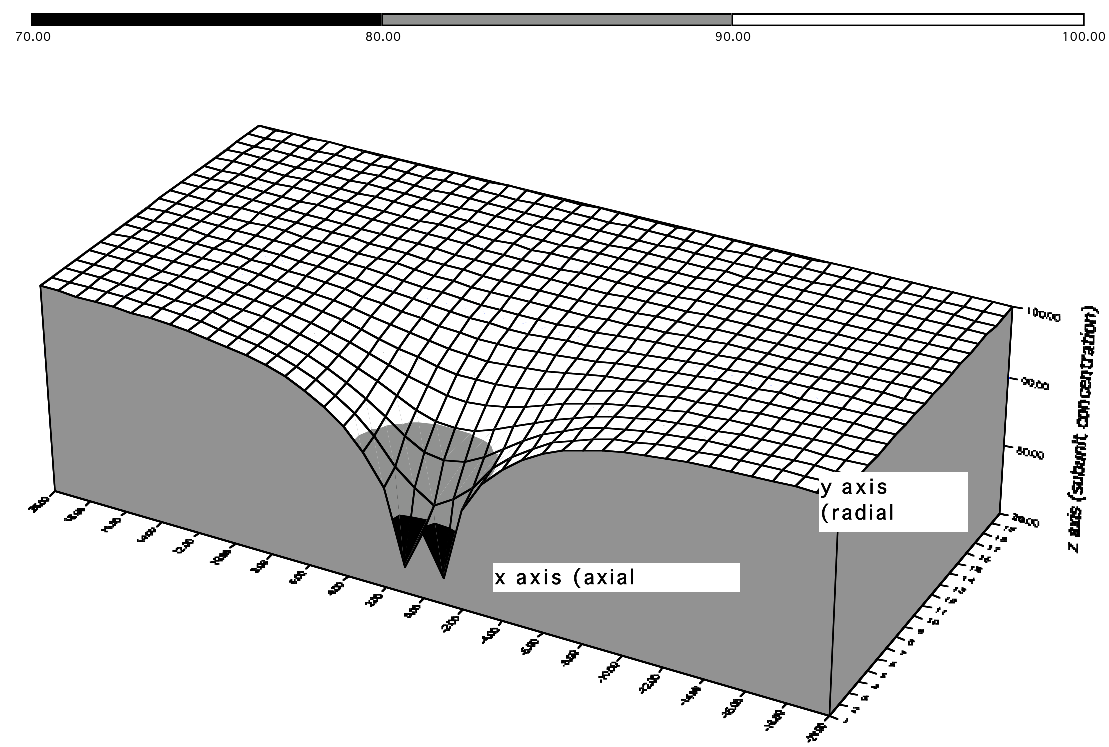


Figure 5

(axial) and y (radial) directions represent one radial plane taken from the "cylinder" of space surrounding the polymer. The polymer lies along the center of the cylinder; the polymer forms the central axis along the cylinder. Distance along the polymer and along the cylinder (passing through the ends of the cylinder) is the x (axial) direction. Distance away from the polymer (passing through the walls of the cylinder) is the y (radial) direction. Unlike the x and y directions, the z direction does not represent a spatial direction. It represents the concentration of subunits at the particular x and y position below it. For example, in Figure 5, at $x = -20$ and $y = 1$, the subunit concentration = 100.

The interesting portion of the figure is the subunit gradient that occurs from around $x = \pm 4$ and $y \geq 4$. Here we see a depletion of the subunit concentration below the initial value of 100. This depletion resulted from the growth of the polymer from its initial position of x = ± 1.0000000 to x = ± 1.533868, and is greatest in the immediate proximity of the growing polymer ends. Does this amount of depletion make sense? The polymer grew by a total length of 2 x 0.533868 = 1.067736. The subunit concentration/polymer length = 1000. Therefore, the subunit concentration theoretically must have decreased by a total of 1000 x 1.067736 = 1,067.736. In fact, after summing the subunit concentrations in the 41 x 20 = 820 squares within the x-y plane, I found a total decrease of 1,067.435 subunits, in excellent agreement with the theoretical decrease. The bulk of the decrease occurred in the region mentioned above from around x = ± 4 and y ≥ 4.

At longer times of t = 3 (Figure 6) and t = 10 (Figure 7) we observe an ongoing decrease in the subunit concentration, resulting from an ongoing increase in the polymer length (Table 2). At t = 3, at the extreme boundaries of the half plane, the subunit concentration has dropped somewhat below 100 (Figure 6), indicating that a somewhat larger "infinite half plane" should have been used in the simulation study. This effect is even more pronounced for t = 10 (Figure 7). However, the general effect of increasing polymer length and subunit concentration decrease is validly demonstrated here, within the limits of the approximations being made.

The effect of increasing the subunit diffusion rate constant is to "fill in the hole" created by the growing polymer ends (Figures 8 - 10; Table 2). The ends of the polymer grow more rapidly, but leave behind less of a local subunit concentration deficit. One could equally say that the ends of the polymer grow more rapidly, *because* they leave behind

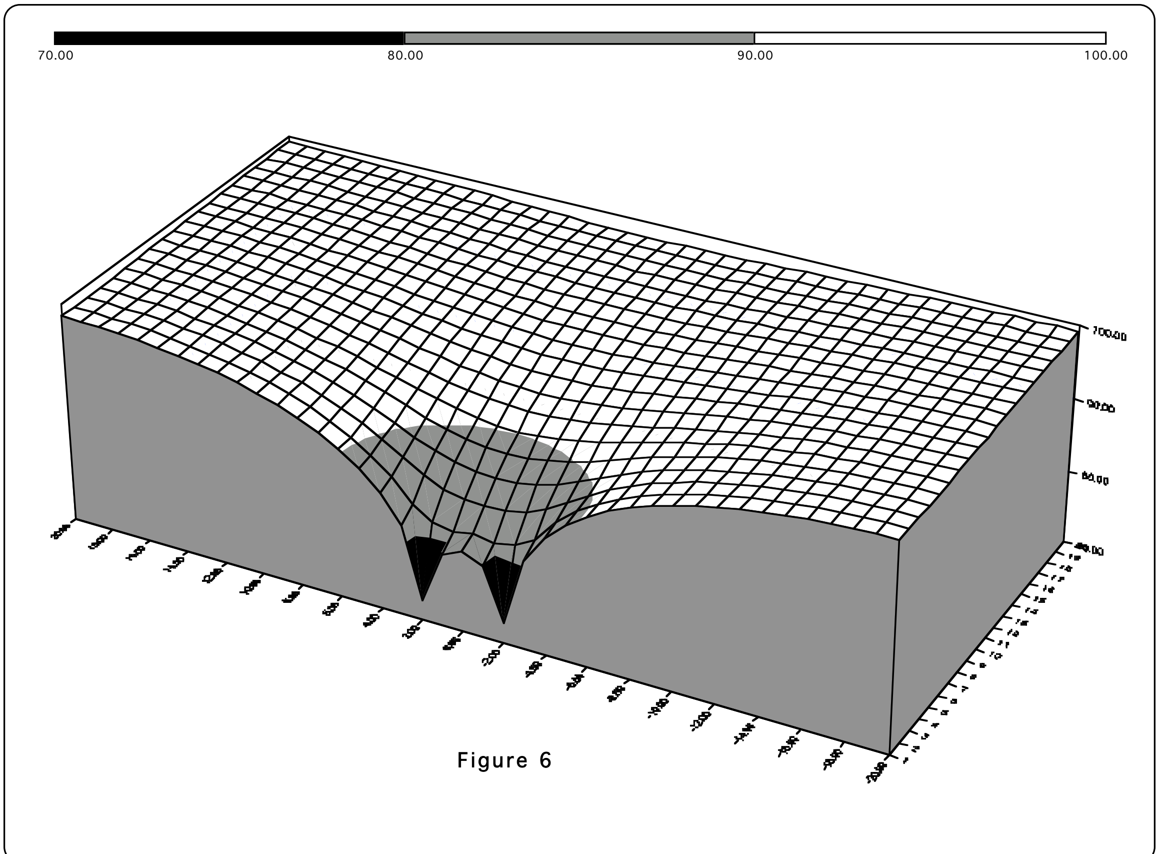


Figure 6

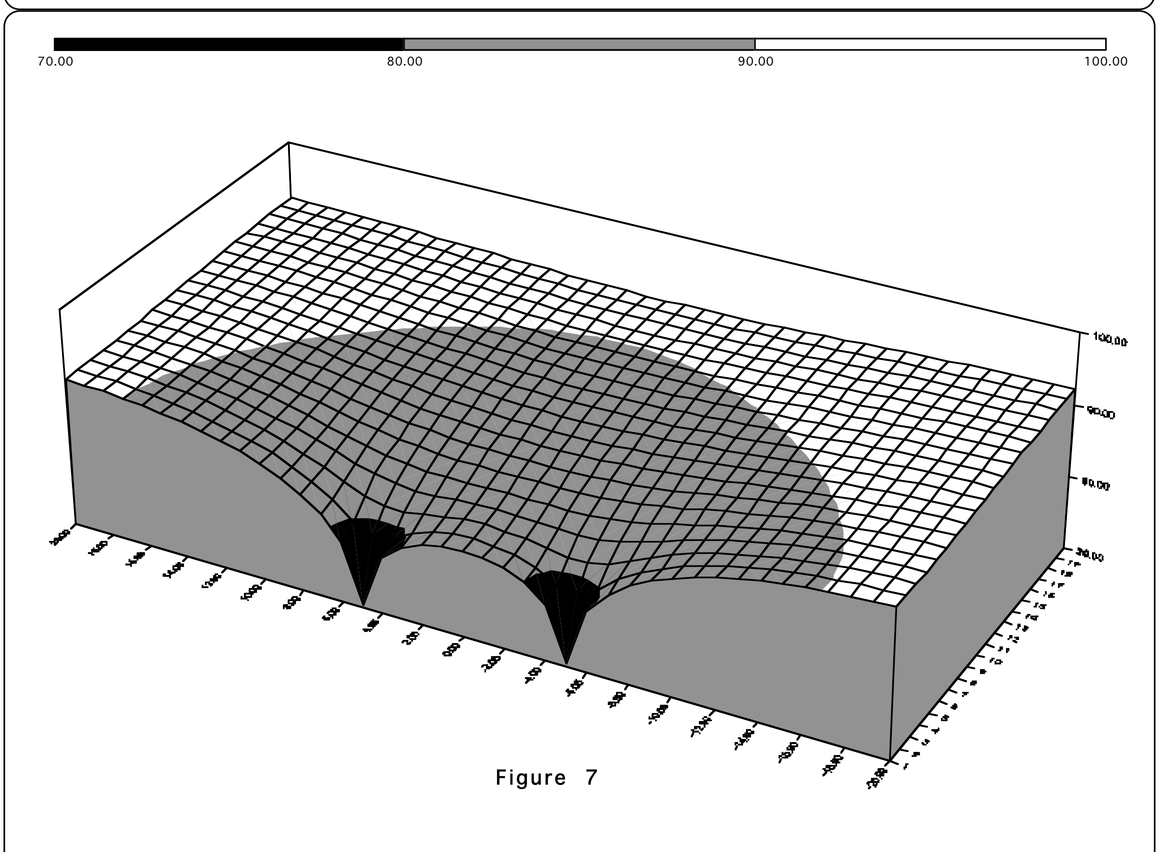


Figure 7

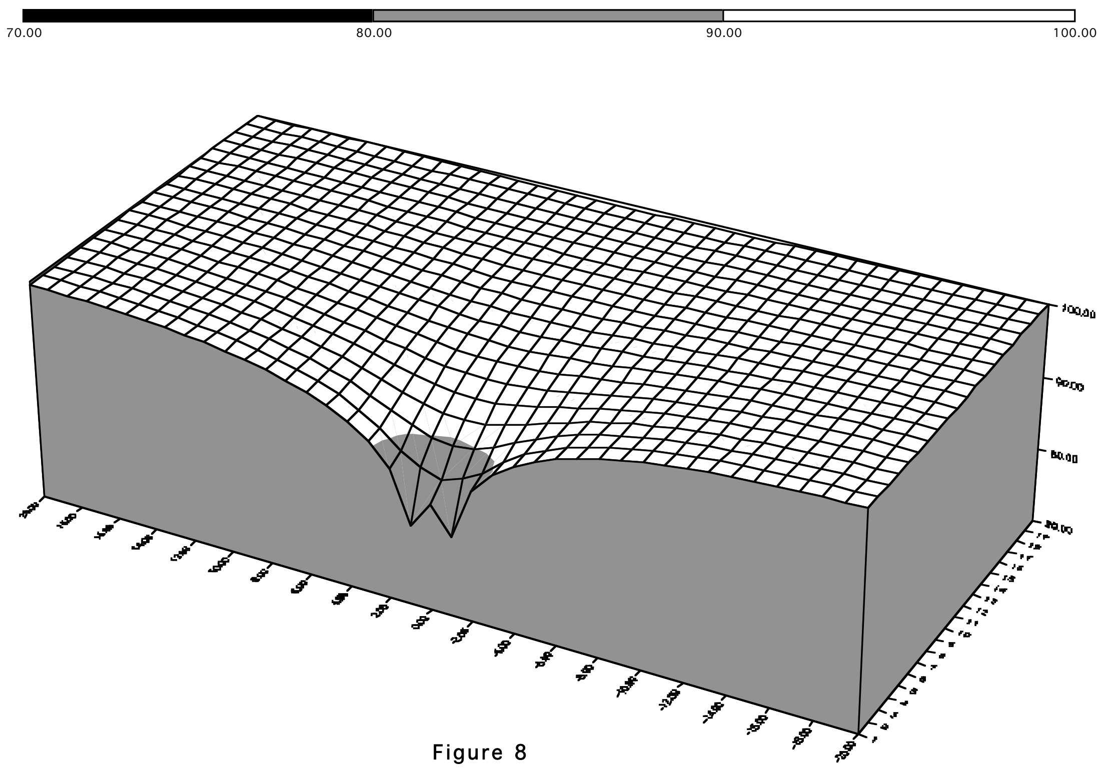


Figure 8

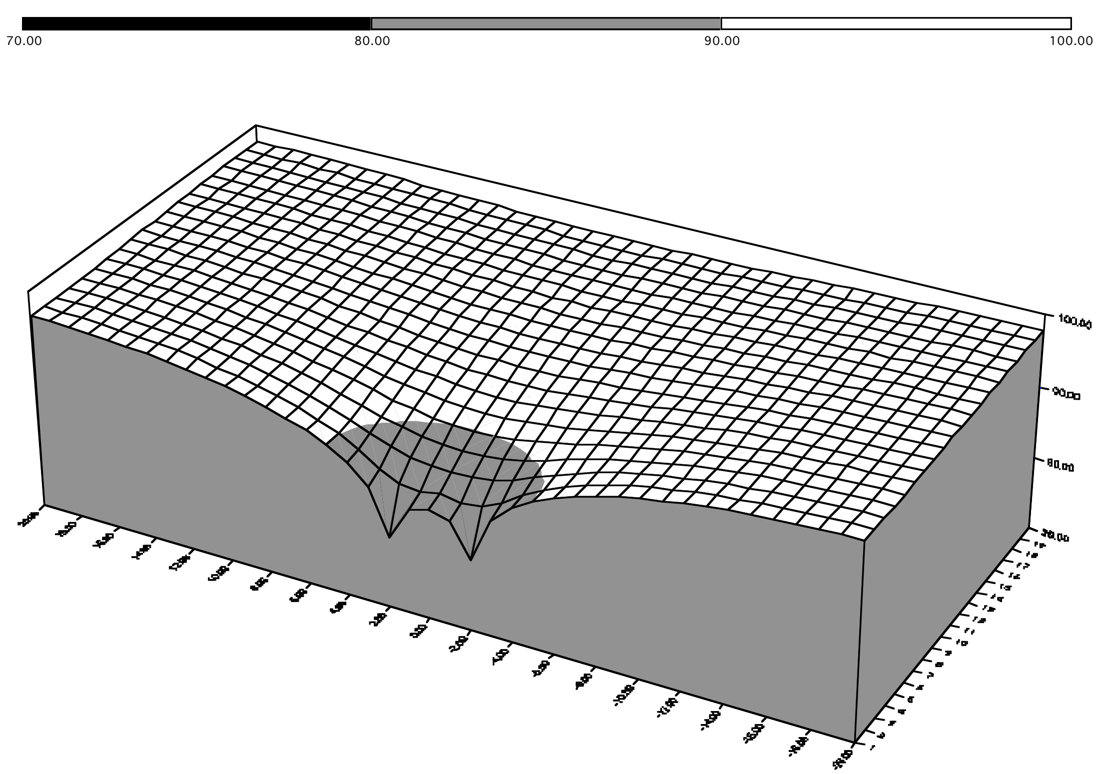


Figure 9

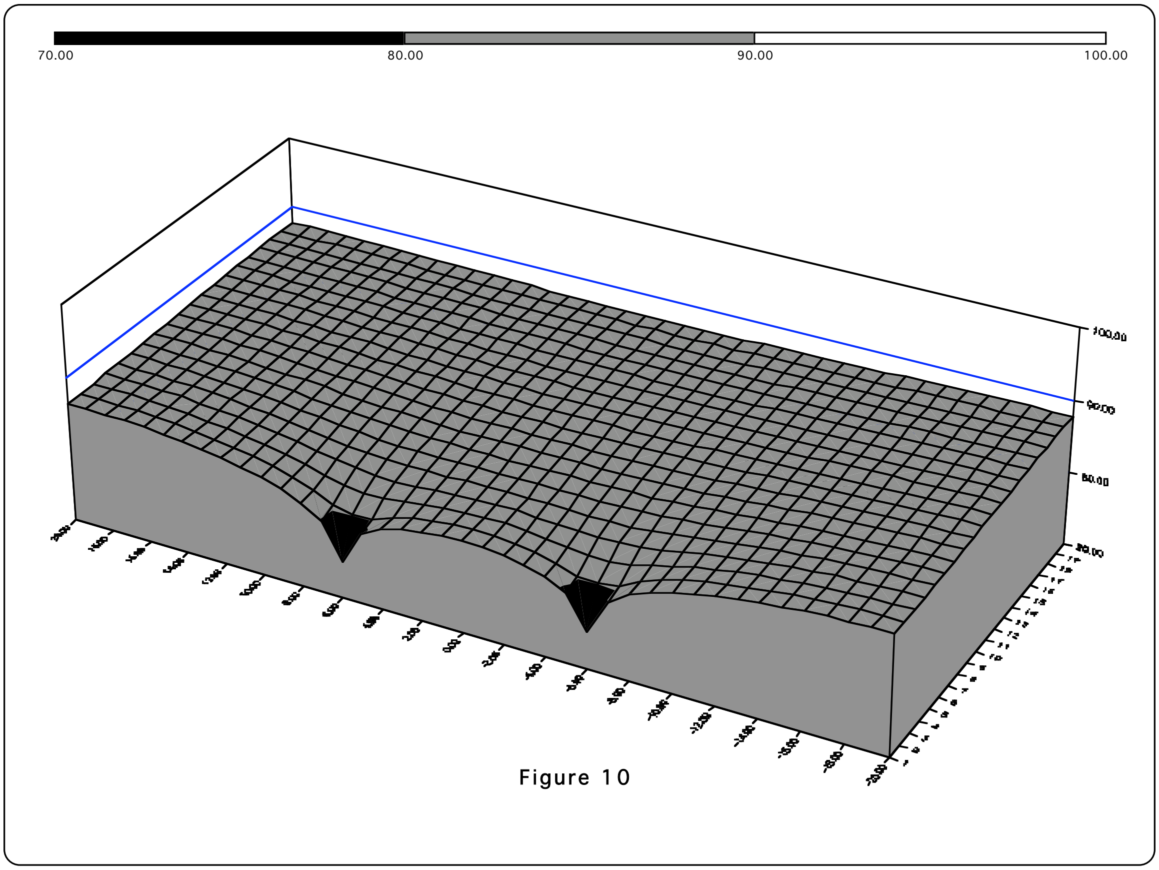


Figure 10

less of a local subunit concentration deficit. The total subunit concentration deficit is amoratized somewhat more equitably over the entire infinite half plane. In the extreme case of a very rapid subunit diffusion rate constant (Figures 11, 12; Table 2), this effect is exhibited even more dramatically.

At the other extreme is a subunit diffusion rate constant of zero. In this case, two different situations might arise depending upon the numerical value of the subunit concentration per polymer length. For moderately high values of the subunit concentration per polymer length, the subunit concentration adjacent to the polymer ends achieves an equilibrium value given by the expression $[\text{subunit}]_{eq}$ = (dissociation rate constant)/(association rate constant), since when the subunit concentration is equal to $[\text{subunit}]_{eq}$, we have net polymer

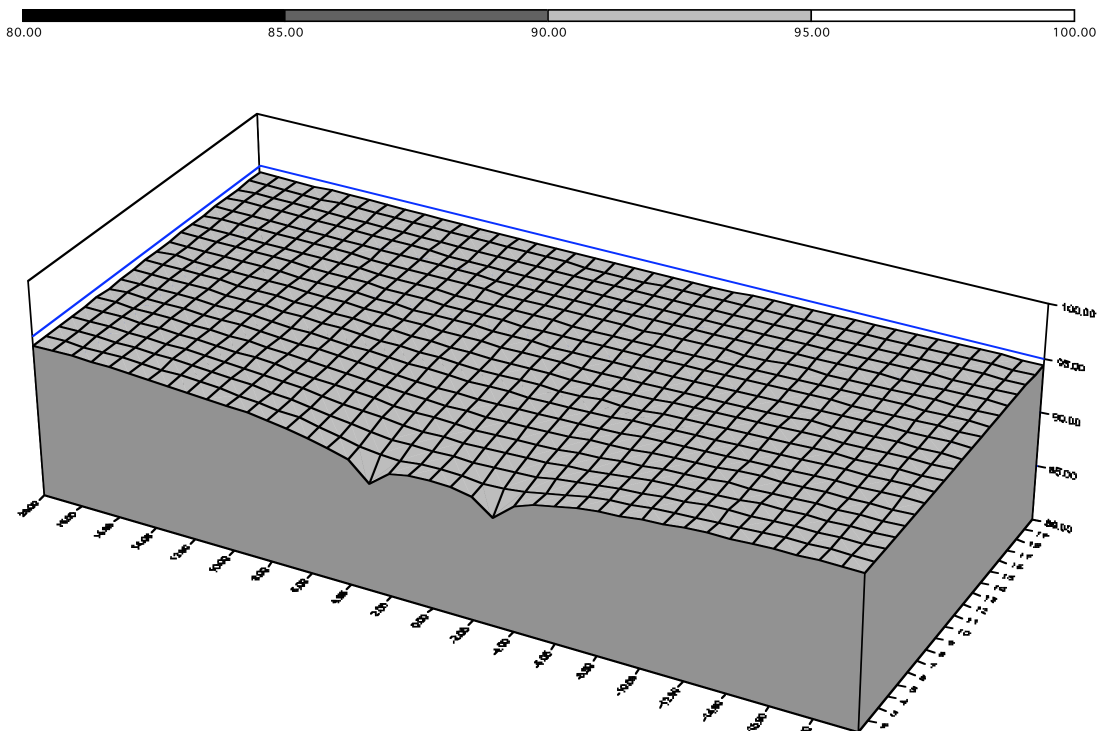


Figure 11

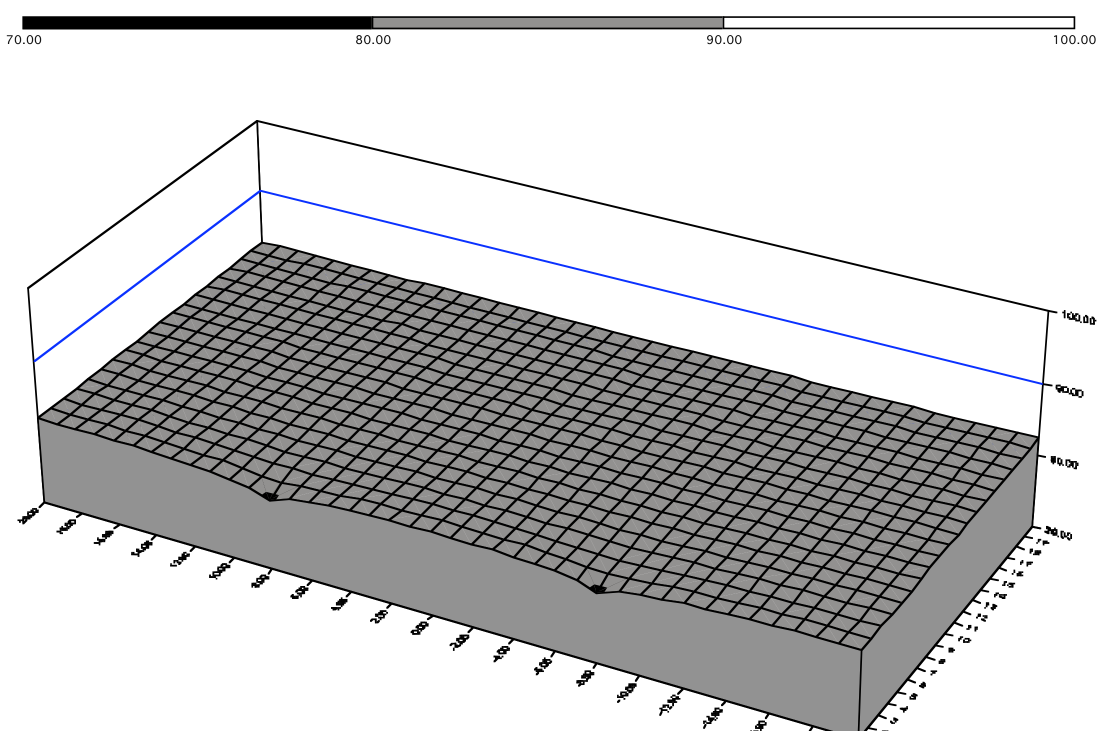


Figure 12

growth = (association rate constant) $[\text{subunit}]_{eq}$ - (dissociation rate constant) = 0. That is, the rates of adding and losing a subunit to the polymer end are equal, and there is no *net* addition or loss.

For the particular rate constants that were used in Figure 13 (Table 2), $[\text{subunit}]_{eq}$ = 1000/20 = 50. Since the initial subunit concentration was 100, 100 - 50 = 50 represents the change in subunit concentration adjacent to either one of the polymer ends. Since the subunit concentration per polymer length = 1000, we expect a growth in polymer length at each end of 50/1000 = 0.050, which was, in fact, observed (Table 2). Since this amount of growth was not adequate to move the polymer end to the next square and there is zero diffusion of subunits, the subunits in the next square were not available to add to the polymer ends. Therefore, the polymer ends remain indefinitely at this same position, trapped in a region containing an equilibrium concentration of the subunits.

For lower values of the subunit concentration per polymer length, the subunit concentration adjacent to the polymer ends may not achieve the equilibrium value given by the expression for $[\text{subunit}]_{eq}$, since the polymer might then grow fast enough to reach the next square before the concentration of the subunits in the current square is reduced to the value $[\text{subunit}]_{eq}$. In fact, when the subunit concentration per polymer length is equal to 49 rather than 1000, a subunit concentration of 51 will be left in each square when the polymer end has just added enough subunits to extend to the next square (Figure 14; Table 2). That is, the addition of a polymer concentration equal to 49 will extend the polymer length by exactly 1. This has two effects: the polymer end will now be in the next square,

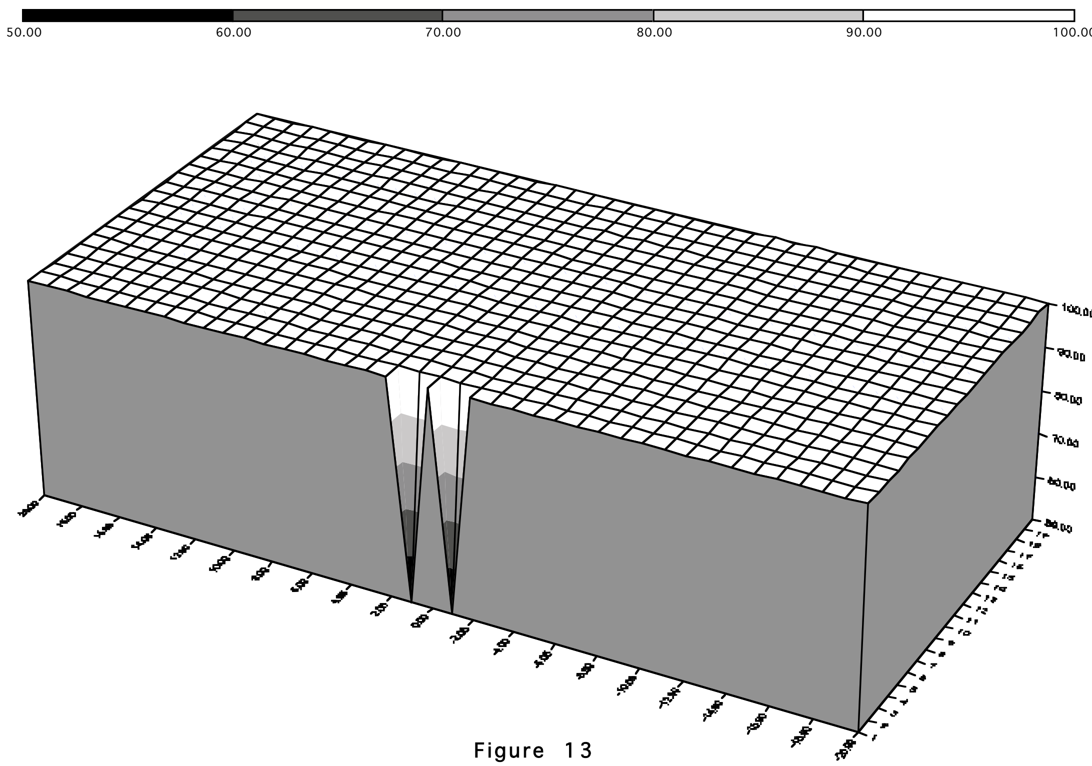


Figure 13

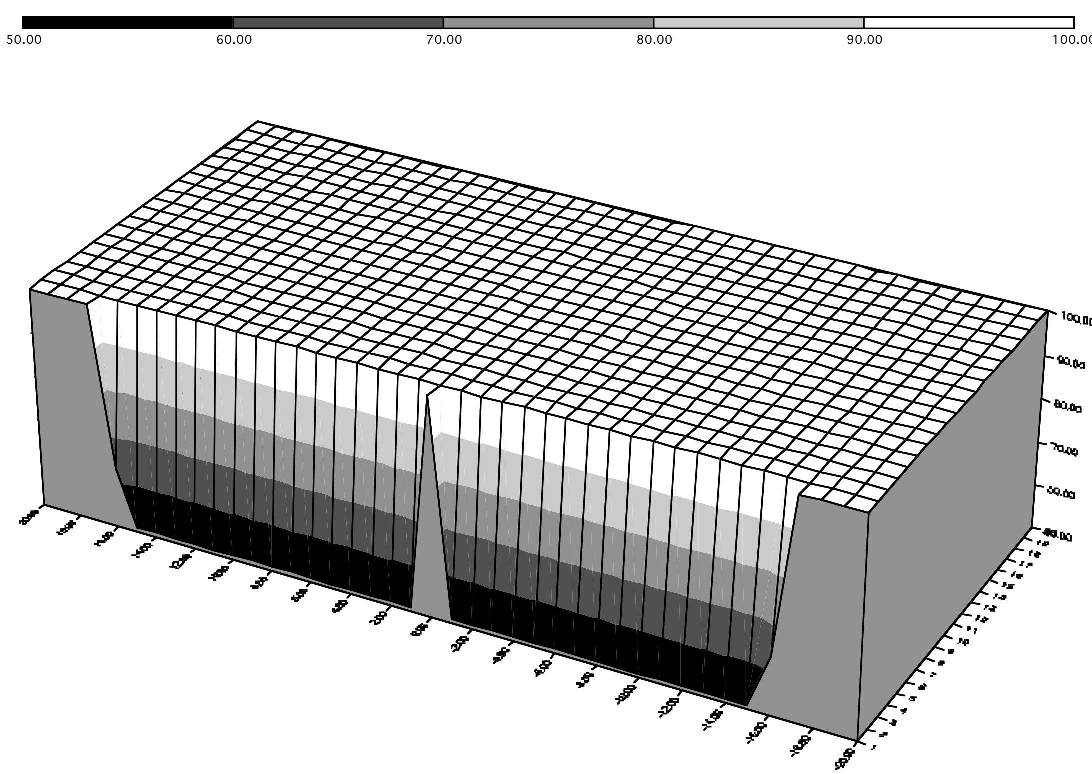


Figure 14

and the remaining subunit concentration in the current square will be 100 - 49 = 51. Thus, in contrast to the stalled growth (Figure 13), there will be robust growth (Figure 14).

All of the cases discussed so far involved symmetrical rate constants at both polymer ends. Now let us "resuscitate" subunit diffusion and consider what might be referred to as "treadmilling," which can result from asymmetrical rate constants at the polymer ends. In treadmilling the possibility exists for both ends to grow but at unequal rates, or for one end to grow and the other to shrink. In the latter case, there may be locations in the subunit gradient where the subunit concentration is greater than 100 (Figures 15 - 17; Table 2). This "excess" subunit concentration arose from the loss of subunits from the shrinking polymer end. An illusion can be created wherein the polymer appears to be "moving" through the solution (Figures 15 - 17; Table 2).

An extreme case of treadmilling would be "unidirectional" polymer growth, in which one end had an association rate constant of zero, and the other end had a dissociation rate constant of zero. The representation of the results of this simulation study is somewhat different from the preceding representations: since I am more interested here in illustrating the time course of events, I have simply put together the first x (axial) row from each of seven times (t = 1 to t = 7) in a single figure (Figure 18; Table 2). The y direction is no longer the radial direction; the y direction is now time (the radial direction has been "thrown away").

This same information is illustrated in an alternative way in Figure 19. In the bottom half of the figure, the polymer position and

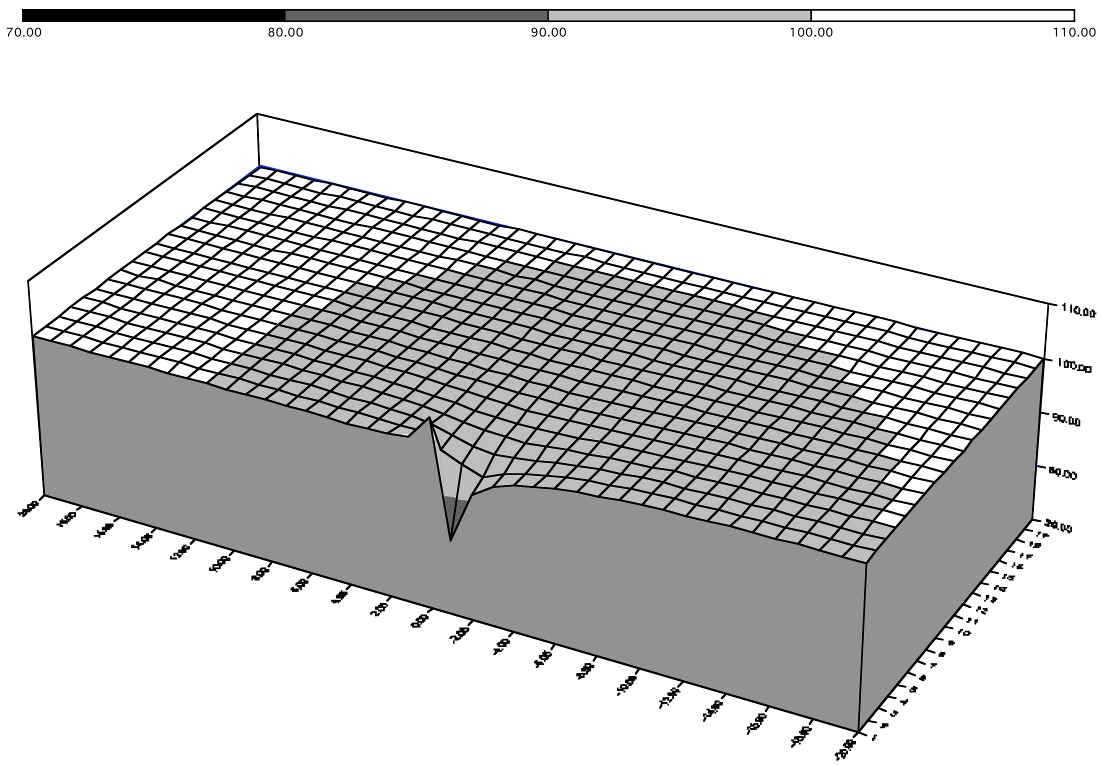


Figure 15

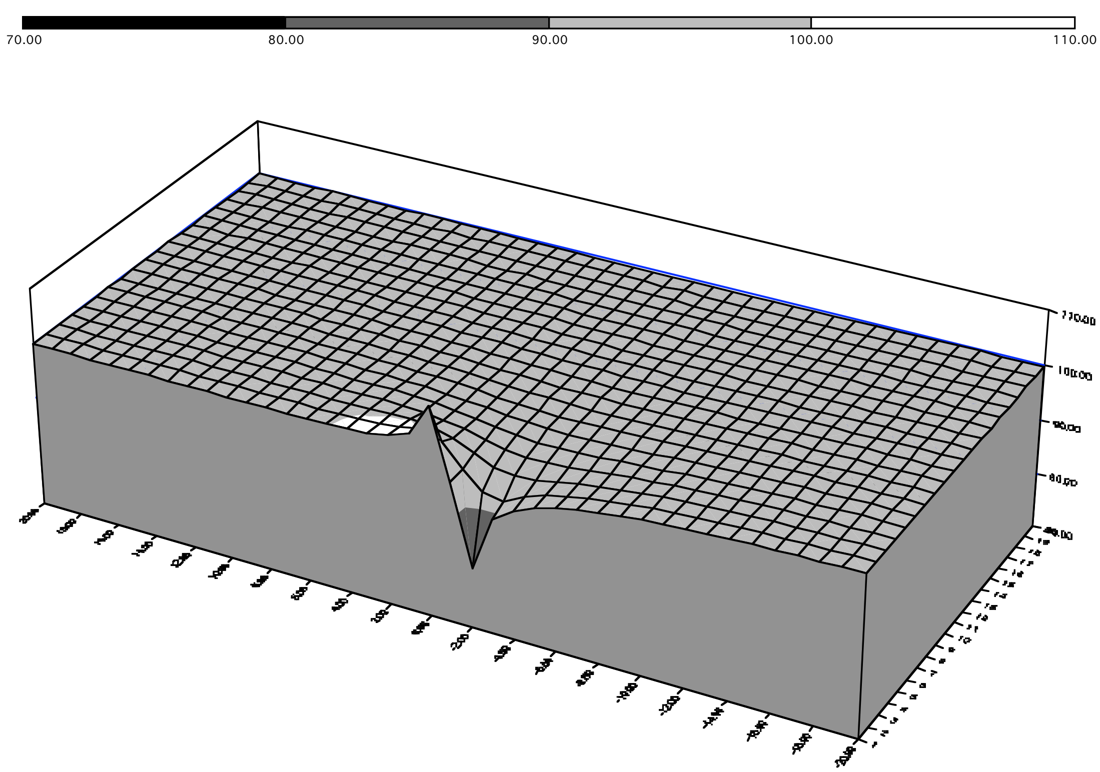


Figure 16

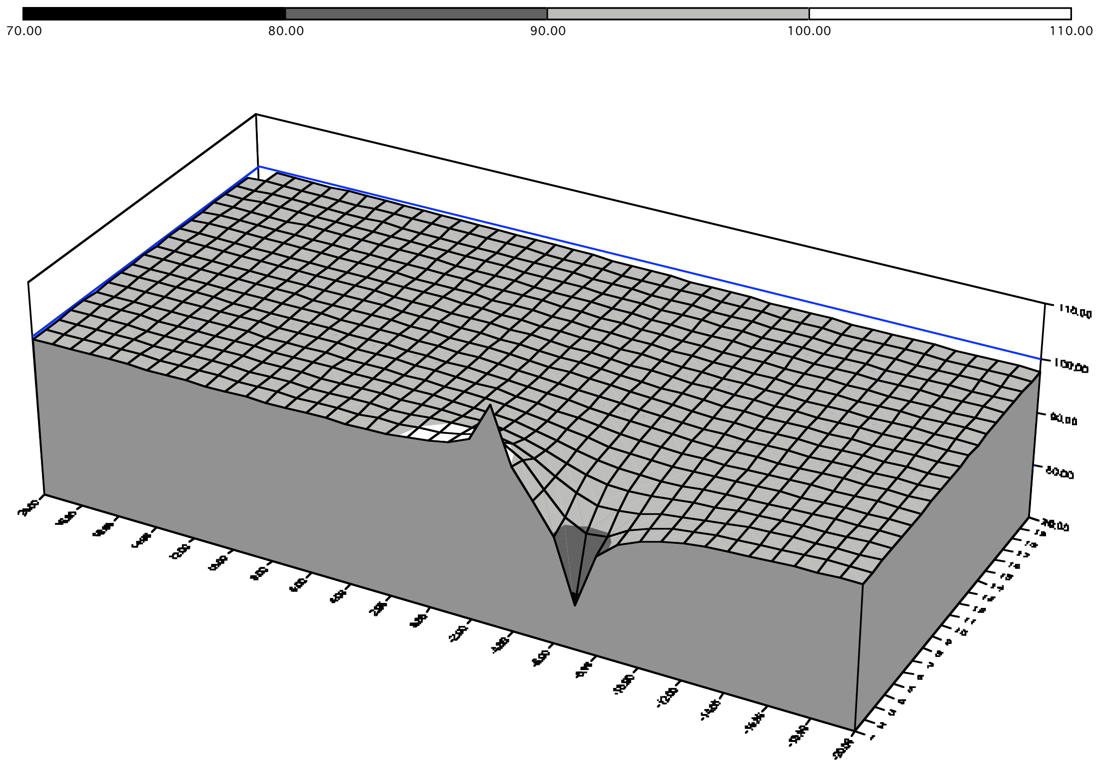


Figure 17

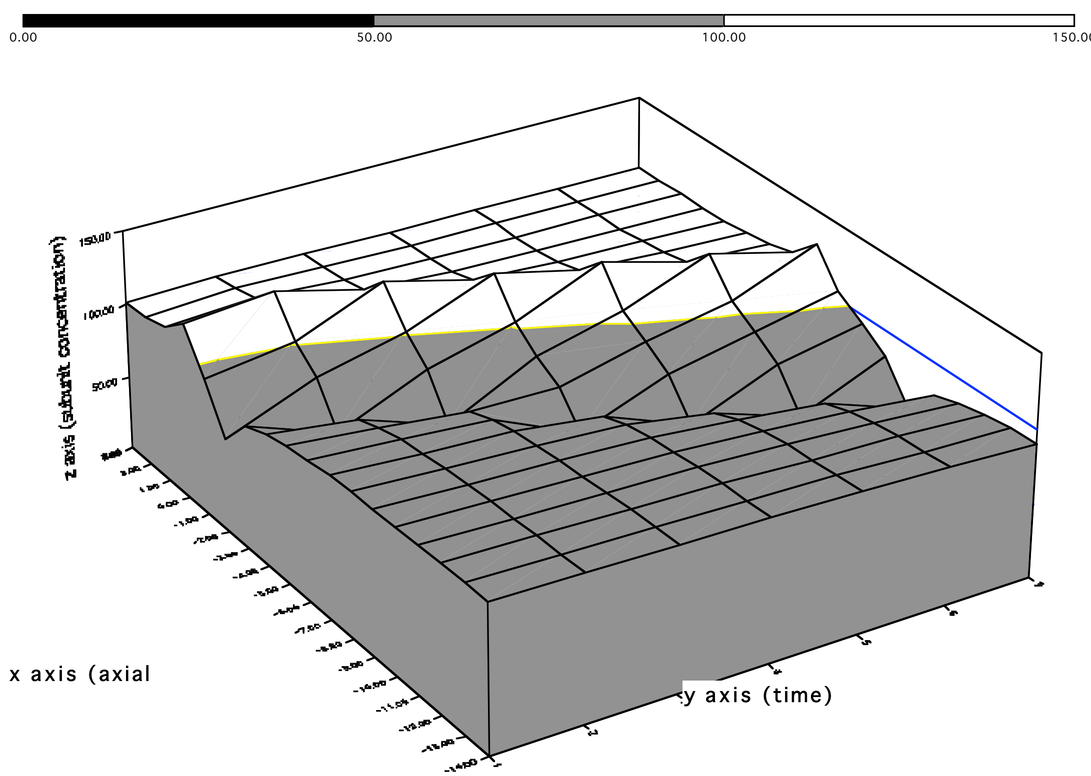
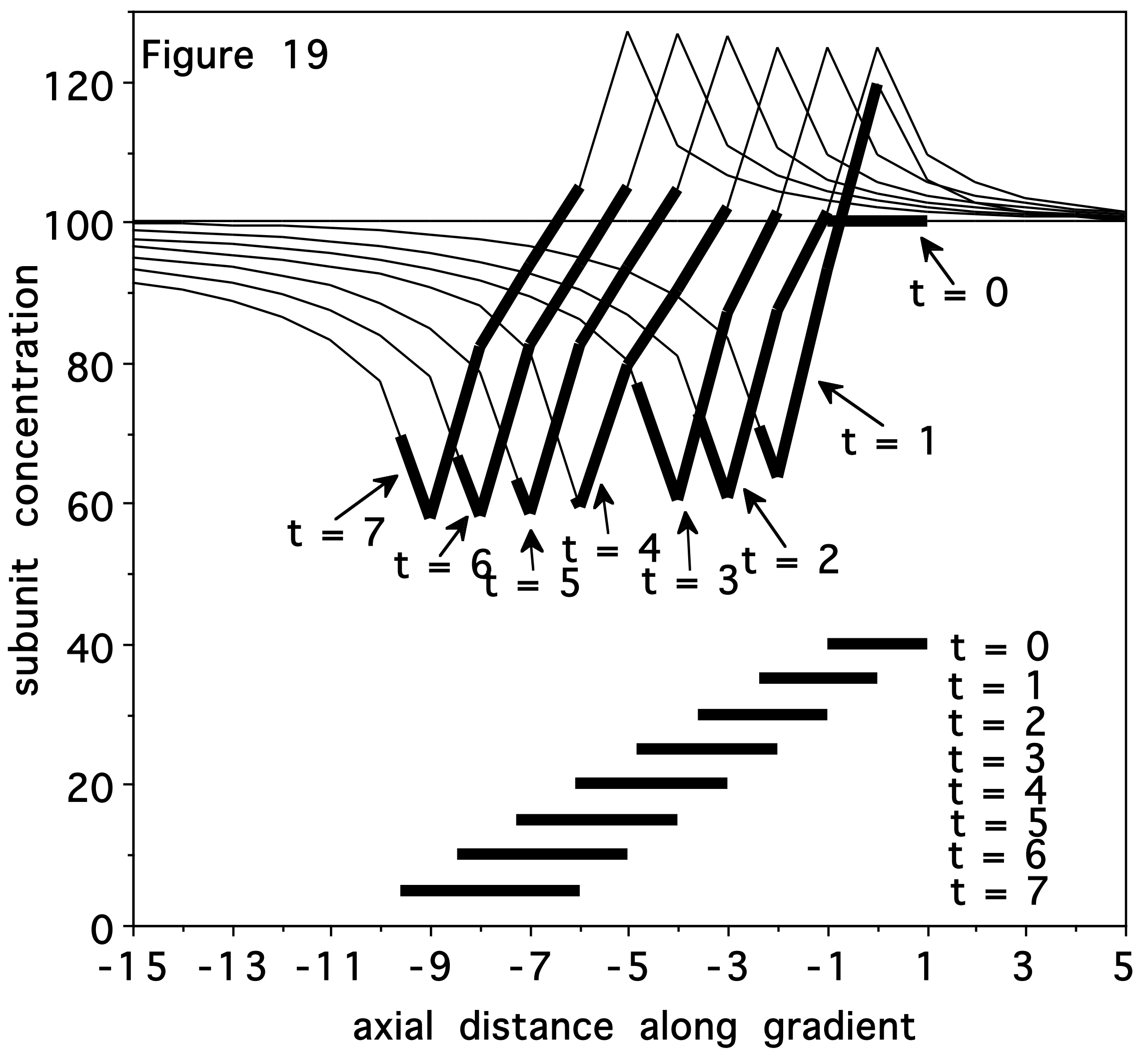


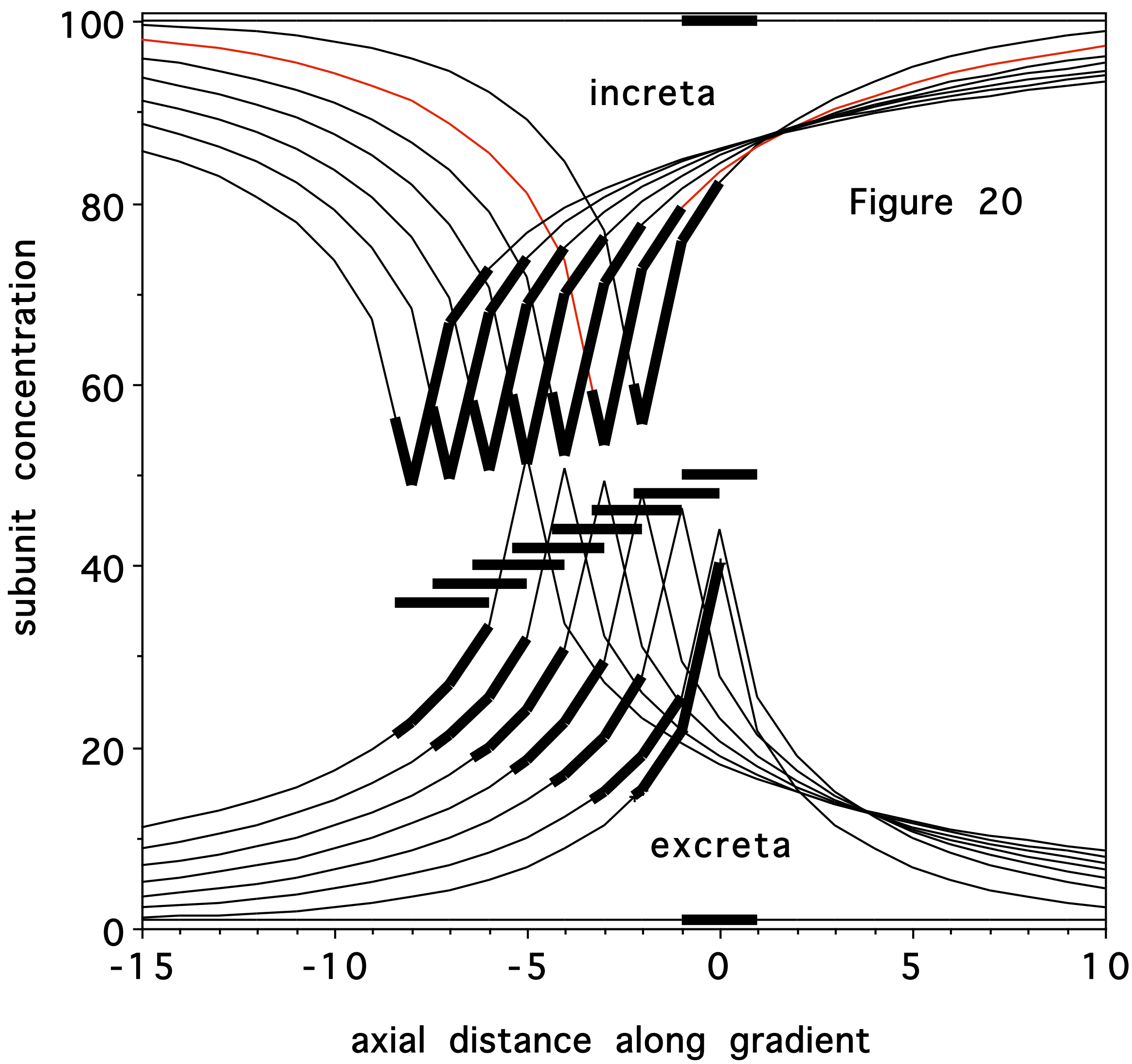
Figure 18



length are shown at each time from t = 0 to t = 7 along the x (axial) direction. In the top half of the figure, the projection of the polymer at each time is indicated by darkened lines within the subunit gradient. It is important to realize that the projection of the polymer does not directly indicate the shape or length of the polymer (remember that the z direction does not indicate spatial extent); it simply indicates which part of the subunit gradient the polymer is lying in. The apparent movement of the polymer creates the illusion that the growing end is moving away from the subunit deficit that it has created (by "consuming" subunits) and is moving up the subunit concentration gradient towards virgin subunit territory.

In the particular case of unidirectional polymer growth, it is not hard to imagine that the subunits that are lost from the shrinking end of the polymer are somewhat altered from the subunits that were initially in the solution (perhaps they have undergone a conformational change, or the chemical energy stored in a molecule of bound ATP has been expended by being converted to ADP and inorganic phosphate). In any event, let us call the altered subunits "excreta" under the assumption that they are no longer able to add to the polymer, even if they were adjacent to the growing end. Obviously, the original subunits would then be called "increta." For the sake of simplicity, let us assume that the subunit diffusion rate constant is the same for both increta and excreta. Then a simulation study that represents a very minor modification of the program for the other simulation studies will produce information about the excreta as well the increta. For example, Figure 20 can be interpreted in a manner identical with that for Figure 19, except that now the growing mounds of excreta (which had an initially uniform value of 1) are shown. This rather scatological example creates not only the illusion that the polymer is moving away from the increta deficit and is moving up the increta concentration gradient towards virgin increta territory, but also that the polymer is moving away from the growing mounds of its own excreta.

We thus have a purely prebiotic system that represents a transitional zone of the organization of matter that demarcates a border between the prebiotic and the biotic. This system behaves in a manner that creates an illusion of the type of behavior that would be



expected of living material. The polymer is a machine that is *remarkably* efficient at converting increta into mounds of excreta.

Because the genetic material of all currently living materials is composed of DNA or RNA, and because DNA and RNA are self-replicating molecules, much of the research and speculation about the origins of life have centered upon the prebiotic synthesis of self-replicating molecules. I feel that this approach is no more closely related to reality than is the approach used by Sherlock Holmes scholars in trying to date the Sherlock Holmes stories by seeking concordances between the fictional weather in a fictional Sherlock Holmes story and the weather recorded on some Tuesday in an almanac published in 1886 (Baring-Gould, 1967). While it is certainly true that contemporary life is dependent upon the self replicating (if you can call a process that requires rRNA, mRNA, tRNA, and about a thousand special enzymes as "self" replicating) features of DNA or RNA, this self replication is simply a reliable way to get a homogeneous supply of molecules bearing genetic information. However, a homogeneous supply of molecules bearing potential genetic information may pre-exist, as exemplified by the triad discussed above, and by discussions to follow below. The development of internal genetic information stored within a self replicating molecule may very well have occurred subsequent to the transition from prebiotic to biotic.

While we are on the subject of DNA and RNA, I would like to recall the fact that, in addition to being self-replicating, the other major property of DNA and RNA is to encode the sequence of amino acids in proteins. Since there are twenty one amino acids that occur in proteins and four nucleotide bases that occur in DNA or RNA, a common question in a biochemistry course is how many nucleotide bases within DNA or RNA are required to form a code word for an amino acid (under the assumption that the code words for the amino acids are of the same length). The common answer to this common question is three, with a consequence of high degree of redundancy. It is interesting to note that a computer uses only half as many code elements (two different bits) as does DNA or RNA (four different nuceotide bases). So how many bits would be required to form a code word for an amino acid? Five, still with some consequent redundancy, but not as much redundancy as with four nucleotide bases.

This raises the interesting questions (which I do not purport to have the answers to) as to why there are four nucleotide bases used in DNA or RNA rather than two or three or five; why there are twenty one amino acids in proteins rather than fewer or more; and whether the number of nucleotide bases used in DNA or RNA determined the number of amino acids in proteins, whether the number of amino acids in proteins determined the number of nucleotide bases used in DNA or RNA, whether these mutually determined each other in a manner analogous to the solution of two simultaneous equations in two unknowns, or whether other considerations within the cellular or external environment determined either or both of these numbers.

It is not known whether the pathway for the evolution of the (now) nearly universal triplet nucleotide genetic code involved organisms (for which evidence obviously no longer exists) whose genetic codes were based upon something other than DNA or RNA, whose genetic codes were based upon two or three or five "bits," or whose genetic codes were based upon code words that were of variable length, perhaps utilizing a Huffman or a Markov encoding (Hamming, 1980) for the sequence of whatever it was that it encoded. Apparently, the ease of decoding a fixed-length code won out over the storage efficiency of a variable-length code. This last statement presumes that there was the opportunity for the other alternatives to be tried out; that is, that the evolution of the genetic code did not get trapped in a "local energy minimum" (Kauffman, 1993).

An interesting experiment to study some of these speculations is based upon the observation that the genome of the nucleus of the yeast *S. cerevisiae* contains two amino acyl tRNA synthetases for each of several amino acids (based upon the MIPS database of December, 1996 in Table 1 of Goffeau *et al.*, 1996). This raises the possibility that a yeast can be found with the following (perhaps rather unlikely but not entirely impossible) properties: (1) the two synthetases partition the set of redundant tRNAs for the corresponding amino acid and (2) one of these two sets of tRNAs is potentially functional but is, in fact, not used. If such a yeast can be found, then recombinant DNA techniques should allow us to change its genetic code in the following manner, without harming the yeast: genetically engineer the appropriate synthetase of the yeast to mischarge an unused tRNA with a new amino acid, and then grow the yeast for a substantial number of generations either under normal conditions or under conditions containing some stress. The genetic engineering can be along the lines used by Riedel *et al.* (1986) to form a functional chimera of the insulin receptor and EGF receptor. This possibility is strengthened by studies with an artificially split gene for a tRNA synthetase (Burbaum and Schimmel, 1991). The question is whether a new protein molecule would evolve that contained the new amino acid. Of course, a large number of new amino acids might be screened in this manner. [Note added in March 2003 revision: In fact this prediction was explored experimentally in *E.coli* (Lei Wang, Zhiwen Zhang, Ansgar Brock, and Peter G. Schultz, Addition of the keto functional group to the genetic code of *Escherichia coli, Proc. Natl. Acad. Sci. USA*, Vol. 100, Issue 1, 56-61, January 7, 2003)].

As mentioned above, research and speculation about the origins of life centered upon the prebiotic synthesis of self-replicating molecules are not likely to be closely related to reality. DNA and RNA contain (at least) three forms of information. The abstract information content will be referred to as the "genotype." It would be incorrect to refer to the physical nucleotide sequence as the "genotype," since the physical nucleotide sequence may result in a particular three dimensional conformation of the DNA or RNA, or in a particular association of histones or other substances to portions of the DNA or RNA; such phenomena are clearly phenotypical, albeit at the molecular level. These types of properties of the DNA or RNA represent the first kind of utilization of the genotypic information. Second, the genotype encoded by the sequence of nucleotide bases provides the information for the production of another copy of the same sequence of nucleotide bases. Third, the genotype encoded by the sequence of nucleotide bases provides the information for the production of proteins. A fourth form of information, which would be unique to DNA, would be the production of RNA copies that are to be used not as intermediate encodings in the production of proteins (that is, not as mRNA), but rather as the production of RNA that has functions other than that of mRNA, such as tRNA or rRNA, etc.

Thus, the genotype can be thought of as several types of encodings of the phenotypic properties, with the remarkable characteristic of representing simultaneously three overlapping encodings: a fixed-length triplet code for protein synthesis, a fixed-length singlet code for replication, and a variable-length code for the molecular properties of the DNA or RNA molecule. Of course, the intrinsic information content is present within a context in which there is an intrinsic decoder (or perhaps multiple intrinsic decoders: one for protein synthesis, one for replication, etc.).

The question arises as to whether there would be any useable or understandable information content if there were not an intrinsic decoder. Before this question can be addressed, we must first ask whether the answer to this question can be obtained in an abstract general manner, or whether the answer to this question depends upon knowing some specifics about living things or at least about the environment in which we are embedded. For example, if we knew the intrinsic genotypic information in an organism, is it possible in theory that we could decode this information if we also knew that there is a fixed-length triplet code that coded for twenty one amino acids of known identity (but we did not know what the code is, we did not have samples of the protein product, and we could not conduct *in vitro* transcription experiments), but we could not decode this information if we did not know that there is a fixed-length triplet code? Given the knowledge that there is a fixed length triplet code that coded for twenty one amino acids of known identity, we then might perform computer simulations of all possible sets of sixty four triplet encodings.

How many of these encodings are there? A lower limit can be estimated by realizing that the first amino acid can be encoded by any of 64 code words, the second amino acid can be encoded by any of the 63 remaining code words, etc, yielding $64 \times 63 \times . . . \times 44 = 2.1 \times 10^{36}$. This is a lower limit, since the redundancy of the encoding will result in some amino acids being encoded by several code words. These

> 2.1 x 10^{36} simulations would involve, at a minimum, enumerating the entire set of protein products derived from the genotypic information in mRNA, simulating the three dimensional structures and interactions of the set of protein products, and determining (in a manner yet to be specified) whether this set of protein products would be capable of maintaining a life form. Presumably only the "correct" genetic code would have this capability, and could be uniquely identified by this property. Although this would be possible in principle, in practice it would not be feasible. The conclusion is that the genotypic information content is without value unless the intrinsic decoding mechanism is available and functional.

One of the characteristics of living materials (Dawkins, 1987; Dawkins, 1992), and, by extrapolation, of nonliving materials, is the propensity to "take advantage" of (that is, to parasitize) another material. Now, an apparently (or even truly) random sequence of 0's and 1's represents potential information if the appropriate decoding mechanism comes along. Anything as valuable as potential information will not be left to itself for long, but a decoding mechanism will inevitably utilize it (it just has to be the right decoding mechanism so that the information content in the random sequence of 0's and 1's represents or ultimately leads to a product that "works" in the external world). For example, let us consider polymers of the types described in Figures 5 - 20. A set of homogeneous polymers of identical lengths contains a limited (but not zero) amount of information. The information contained is the fact that they exist rather than not exist, that they are homogeneous, that they are of the same length, that they move at a certain rate in a direction away from regions of excreta and diminished subunit concentration, and that there is a certain concentration of them. Although this is not zero information content, it would probably not be sufficient to encode much more than (I don't mean this literally) a glycine molecule, even if a really good decoding mechanism came along.

However, it is not too difficult to add more information content. For example, perhaps the polymers are not all of the same length, but perhaps there is a length distribution ranging from a small number of very short or very long polymers, to a large number of polymers of intermediate length. Maybe something would come along that could decode the length distribution and utilize this information for its own purposes.

Perhaps more to the point would be to consider the excreta. The original subunits (increta) may be converted to excreta at any point along the length of the polymer. Furthermore, this may be a Markov process; that is, the rate constants for the conversion of increta to excreta within the polymer may depend upon the proximity of the increta and excreta forms of the subunit. A reproducible and predictable pattern of increta and excreta will then develop along the length of the polymer. Of course, a sequence of increta and excreta is logically equivalent to a sequence of 0's and 1's if we agree that increta are to be identified with 0's and excreta are to be identified with 1's, or *vice versa*. Now all we need is for the right decoder to come along, and this polymer will represent a rich source of information.

Another alternative is that there are two different types of subunits that can both become incorporated into the polymer (copolymer). Each subunit type has its own characteristic rate constants, etc. The rate constants may be a function of the composition of the copolymer (either the global composition or the composition near the end at which the association or dissociation event is taking place). Again, this would lead to a reproducible and predictable pattern of the two types of subunits, and the reasoning in the paragraph above will hold.

An objection to this possibility might be raised. We would like to have a reliable source of copolymers that are of high uniformity (that is, each copolymer has a subunit sequence that is the same as, or at least highly correlated statistically with, all the others). We would also like to have a source that is not depleted quickly, so that the process of producing these copolymers is relatively stable over some reasonable time period. We would also like the subunit sequence within a single copolymer not to be a simple repeating pattern such as "0101010101 . . . " But it would seem that conditions that would lead to reliability, uniformity, and stability would lead to simple repeating patterns. For example, one way to avoid the simple repeating patterns would be to have a depletion of one or the other of the subunits in the solution. However, this would violate the stability that we desire.

A possible solution to this would be for the kinetic properties of the subunits to vary differentially over time, but in a periodic manner. Thus, if the association and dissociation rate constants were dependent upon the flux of a particular wavelength of light (perhaps in the ultraviolet) then the changing flux between morning and night would lead to a variation in the simple repeating pattern, and the reliable periodicity of the day/night cycles would ensure that the same pattern of changing flux would occur repeatedly.

As an aside, there is much speculation about why naturally-occurring proteins are composed of L-amino acids rather than D-amino acids. Was there an actual source of asymmetry that influenced prebiotic or biological evolution, or was it random chance? Potential sources of asymmetry would be the direction of the Earth's orbit around the sun, the direction of the Earth's rotation about its own axis, and the direction of the moon's orbit around the Earth. If it is true that periodicity of the day/night cycles (or perhaps some other astronomical cycles) played a role in encoding information, then this could potentially provide a basis for a rational explanation of the L-amino acid *versus* D-amino acid question.

This is probably a good place to point out that there are two extreme forms of decoding information, and to make a subsequent analogy with biochemical phenomena. Make believe that there is a person holding a piece of wood with an apple resting on it. If the piece of wood is suddenly removed, the apple will fall to the ground. In a sense, by simply falling, the apple has decoded the fact that the piece of wood is no longer there. On the other hand, make believe that I write a computer program in the "C" programming language. The information in the program that I wrote is not useful until another program in the computer compiles my program into object code, and then this object code gets converted into machine code, and the machine code gets decoded by the electronic circuits of the hardware components of the computer. The point I am trying to make is that decoding can be something that seems very direct and simple and naturalistic, or it can appear to be rather contrived and indirect. Since most things are not simply black or white, there is likely to be a more or less continuous spectrum of decoding mechanisms ranging from the direct to the contrived.

A similar thought had occurred to me a while ago about biochemistry. The components of a biochemical system can be divided into two categories: those that are very direct and simple and naturalistic, and those that are contrived. For example, digestion of food by hydrochloric acid would appear to be direct and simple and naturalistic; after all, a strong acid will digest materials. On the other hand, the damage done by the intracellular release of calcium ions in a host of biomedical illnesses appears to be rather contrived and indirect. There is nothing inherently bad or threatening about calcium ions, except for the rather contrived and indirect effect it has upon activating various enzymes, etc.

Returning now to the idea of a polymer containing two types of subunits and the idea of a spectrum of decoding mechanisms, I would like to put forth two fairly specific decoding mechanisms for the polymer. The first decoding mechanism would, interestingly enough, be constituted of the three dimensional shape of the polymer itself, and its interaction with the aqueous environment in which it resides. That is, the polymer is, in a sense, decoding its own sequence. What I am about to suggest is fairly predictable, but I will say it anyway: A particular sequence might give the polymer the ability or the predisposition to assume a globular shape, in which the more hydrophilic surfaces are (surprisingly enough) on the "outside" and the more hydrophobic surfaces are on the "inside," forming a lipid-like "microenvironment" within an aqueous macroenvironment.

The second decoding mechanism would involve the polymer assuming a fairly linear conformation, and thus providing an extended surface upon which other materials might differentially align themselves (at approximately right angles to the polymer) depending upon which of the two subunit types the other materials were in contact with.

Biological evolution depends upon several factors. Certainly at the center lies the genotype; not the physical sequence of DNA or RNA, but the abstract concept of the information content inherent in that sequence. The DNA or RNA must have a high degree of stability, and an accurate mechanism must exist for the replication of the DNA or RNA. However, the stability must not be too great nor the replication too accurate, lest a low rate of genetic change not be feasible. There must be an interaction between the (abstract) genotype and the (real) external physical world. This interaction is most likely implemented by the relative ability of the (real) physical products of the genotype to exist in the ambient environment. Not only are there nonbiotic phenomena to cope with (ultraviolet radiation, fluctuating temperatures, fluctuating availability of water, etc.), not only are there other species to cope with (predators or prey that can run faster than you can), but there are also other members of your same species, whose needs for resources are identical to your own needs and who can therefore be your greatest threat when resources become scarce and limiting (Dawkins, 1987; Dawkins, 1992).

In contrast to a religious perspective of our existence, in strict Darwinian evolution all vestiges of teleology must be removed. That is, we are given a genotype embedded within a certain environment, and a change in the phenotype (that is, a mutation, duplication, etc. in the physical DNA or RNA sequence) results in a corresponding change in the genotype of an offspring (unless, of course, the change is such as to prevent conception or birth of the offspring). Given that the offspring is to some extent viable, it is then "tested" within the prevailing environment. The change may be catastrophic in that it does not permit conception or birth. The change may be directly self-limiting in that it leads to death before the offspring reaches reproductive age, or it leads to sterility. In the case of an organism that reproduces sexually, the change may be indirectly self-limiting in that it leads to a phenotype that members of the opposite sex find repulsive (unless the organism bearing the change is a rapist). The change may influence survival in an even less direct manner, adversely affecting nonreproductive functions such as deteriorating the acuteness of vision or lowering the speed of movement. In rare cases, the change may influence survival in a positive manner, such as enhanced attractiveness to the opposite sex, increased visual acuity, or more rapid speed of movement.

In environments where resources are not limiting, these indirect changes may have little or no effect upon survival and perpetuation of the new genotype. However, in environments where resources are limiting, these indirect changes may have a substantial effect upon survival and perpetuation of the new genotype.

The Lamarckian theory of inheritance, at one time a viable competitor of the Darwinian theory of inheritance, is now considered to be incorrect. A prototypical example of the Lamarckian theory would be a man whose occupation resulted in excessive muscular development, and whose son, therefore, inherited large muscles. The apparent difficulty with this theory is that there is no biological mechanism for this to occur. For example, performing muscle-building physical labor does not alter the genotype.

However, performing muscle-building physical labor certainly alters the phenotype, and a part of the phenotype is the physical DNA molecules in the genome. It is not inconceivable that some aspect of the physical DNA molecules is altered by muscle-building physical labor. For example, the levels or metabolic degradation of particular circulating hormones may be altered in a manner that alters the proteins that are associated with the physical DNA molecules. This could, in turn, alter the processes of sexual recombination relating to conception, leading to an altered probability of allelic inheritance in the offspring. Even more compelling would be the possibility that the pregnant woman's occupation leads to hormonal changes that differentially affect (or possibly even effect) the *expression* of embryonic genes within an unaltered embryonic genotype. This could lead to an irreversible "locking in" of neuronal or neuromuscular circuitry in the embryo for the duration of its lifetime.

Of course, the problem with the Lamarckian theory of inheritance is that there is no reason to believe that the hypothetical hormonal changes and their ensuing effects upon the physical DNA recombination or expression would connect the behaviors causing these changes to biochemical changes that would perpetuate these same behaviors in the offspring. It may happen, in some rare instance, that this correlation does exist. In that case, I suppose, one would consider this an example of the Lamarckian theory of inheritance. However, this would presumably not happen often enough to be considered as a viable *general* model for biological evolution.

This is not to say that coincidence does not have its role in Darwinian evolution. For example, we know that animals can adapt to be able to survive in environments that are different from those in which they originally had evolved. Also, we know that there is an amazing degree of plasticity in response to injury or tissue loss (for example, neuronal plasticity). Perhaps the most dramatic example is the ability of a small percentage of extant species to survive a mass extinction event (Stanley, 1987). However, since Darwinian evolution is devoid of any teleological component, it must be that the ability to adapt, plasticity, and survival of mass extinctions are coincidences. A genotypic change is tested only for its survival ability (presumably within an uninjured organism) within the environment which predominated at the time when the genotypic change occurred. This test did not take into account whether the organism would survive within a different environment, or would be able to compensate for an injury or for a catastrophe. Therefore, the ability to do so must be considered as coincidental. That is, the organism is sufficiently complex that certain "latent" interactions can occur that were not directly selected for by Darwinian evolution.

A third theory of evolution, which will not be described here in detail and which had an even shorter duration than the Lamarckian theory, is the little known Malarckian theory (Wallace, 1996).

Survival or existence at the macroscopic level (disregarding for the moment the possibility of the conversion between matter and energy and the transformation of one element into another by fusion or fission) refers to the fact that a particular atom (or collection of atoms) exist in one particular configuration rather than in another. For example, a salmon exists because certain atoms of carbon, hydrogen, nitrogen, etc. are in a particular configuration known as a "salmon." What is the nature of the existence of this salmon?

Well, if one were to inject a radiolabelled compound (for example, [^{14}C]acetate labelled in one particular carbon atom) into the salmon, within seconds the radioactivity from that compound (that is, the one particular carbon atom that was radioactive) would be found (for instance, by using two dimensional thin layer chromatography) to be an atom in some thousands of compounds within the salmon, not to mention that it would be part of the carbon dioxide gas that was released from the salmon's respiratory processes. Therefore, the salmon is not the same object that it was a moment before. It is now actually composed of different molecules than it was a moment ago, although these molecules would have the same chemical name.

But we don't ordinarily look at objects by examining the transformations that can be elucidated by "following" a particular atom or molecule within the object, and ordinary observation does not reveal these dynamic changes to our visual perception. Within a certain amount of time, that salmon may, in fact, be composed of entirely different molecules than it was composed of at an earlier time, yet we would still think of it as being the same salmon. If we were to eat this salmon, and if we could recover those atoms that (now digested) had made up the original salmon at the "moment" that it was eaten, we would no longer call this ensemble of atoms a "salmon."

Thus, we see that there is a very subjective component to the concept of survival or existence. The word "subjective" requires further explanation, and this is as a good a place as any to touch upon a rather difficult topic. I may refer to the human brain as having such and such a property or characteristic, with the idea being that this property or characteristic is unique to the *human* brain. I would not quibble with someone who took exception to this type of statement if they were to say that the brains of some great apes or dolphins, etc. also have this property or characteristic. I will simply use "human brain" as a shorthand notation to represent the class of all brains that might have the stated property or characteristic. There are two comments that I would like to make concerning this convention. First, it may, in some or even all cases, be that the statement happens to be true uniquely for (literally) human brains, even though it is stated simply as a shorthand for this class. It may also be that sometimes I will explicitly limit the class to (literally) human brains.

The expression "gene pool" has come into our language, presumably with the meaning of the set of all genes (alleles) that are in genomes. Presumably duplicates of genes are omitted, since just one copy or enumeration of each one is sufficient. The expression could be taken to mean all genes that currently exist in genomes in living organisms (that is, alive at the instant that the statement is made), or all genes that had existed throughout history, some of which are now extinct, or even all genes that potentially might exist in the future although they don't exist yet and haven't existed in the past.

I would like to use a similar expression which I suppose might be called the "human brain pool" (in this case literally human). The human brain pool would be the set of all human brains that exist at the moment within living humans. Since the very property of what is meant when we say that an object exists is what is currently under discussion here, it would perhaps be more prudent to define the components of the human brain pool not as the physical brains themselves, but the abstract properties of these brains. This presumes that, although the molecular components of a brain may change when examined at the level of radiolabelled tracers, the abstract properties of these brains are not affected (within relatively short time frames) by these changes of the molecular components. For the sake of clarity, at the risk of redundancy, when I speak of "changes of the molecular components" I do not mean that the statistical concentration and distribution of the NMDA receptor (just to be overbearingly specific) is changed, just that the atoms composing the NMDA receptor are not the same atoms that they had been a little while ago. Again, when I say "not the same atoms," I do not mean that a carbon atom is now a plutonium or technetium atom, but rather that it is a different carbon atom.

There are several purposes for using an expression like the "human brain pool." First, it provides a sense of thinking about human brains as objects removed from the bodies that they happen to reside in. Second, it emphasizes that there is not one single prototypical human brain, but a set of human brains that could be studied on a statistical basis. For example, I could imagine a property such as "abstract thinking" as being objectively evaluated and a numerical value for this property being attributed to each brain within the brain pool. Of course, "abstract thinking" would have to be defined fairly precisely: for instance, do we mean "having the (as yet untapped) potential for abstract thinking" or do we mean "actually does a lot of abstract thinking?" The natural thing to do at this point would be to construct a frequency distribution (histogram) of the numerical values for the particular property under consideration. Presumably some histograms (for instance, the abstract thinking one) would have a fairly normal (Gaussian, bell-shaped) distribution, whereas others (for instance, "degree of substantial contributions to Unified Field Theory") might be bimodal or sharply spiked.

So, moving right into the simplest living organisms, what did they need in order to get started? Well, I won't claim that the unidirectional polymers, with excretion, of Figure 20 are living (on the other hand, I have very cleverly managed to insert this disclaimer). They need a fairly reproducible source of molecules that would assemble in such a way that there would be an "inside" and an "outside." This sort of assembly would be of the type exemplified above by hydrochloric acid digesting food, rather than calcium ions regulating an enzyme. That is, the actual physical-chemical properties of the molecules would lead to this assembly, not some indirect machinery having nothing to do with the prebiotic physical-chemical properties of the molecules. For example, it could be the decoding of the information content of the sequence of "0's" and "1's" along the length of a copolymer that lead to this assembly. In this context "decoding" means, as described above, that either a particular sequence might give the polymer the ability or the predisposition to assume a globular shape, in which the more hydrophilic surfaces are on the "outside" and the more hydrophobic surfaces are on the "inside," forming a lipid-like "microenvironment" within an aqueous macroenvironment. A second possible decoding mechanism would involve the polymer assuming a fairly linear conformation, and thus providing an extended surface upon which other materials might differentially align themselves (at approximately right angles to the polymer) depending upon which of the two subunit types ("0's" or "1's") the other materials were in contact with.

We now know that reasonably stable genetic information that encoded the construction of machinery to use calcium ions to regulate enzymes, etc. eventually came about. Once this occurred, the dependence of the life form upon the physical-chemical properties of the molecules in the environment in which it is embedded did not disappear, but their dependence changed in a dramatic way. The original type of dependence was for the environment to supply a reliable and reproducible source of the molecules whose physical-chemical properties would lead to encoded information and a decoding mechanism. In this case the decoding was by "brute force," such as in the two decoding mechanism just reiterated above. The decoding depended entirely upon the prebiotic physical-chemical properties of the molecules. There would be no distinction between the decoding and a subsequent or consequent implementation of the decoding. In fact, the decoding *was* the implementation.

Again, at the risk of some redundancy, the decoding was like the example of an apple resting upon a square of wood, and the wood being suddenly removed. The apple "decoded" the effects of gravity by falling. The decoding *was* the implementation. To clarify this by distinction, a computer program simulating the effects of gravity upon the apple would, for instance, iteratively check to see if the wood had been removed since the last time this was checked, and if not it would check again. If so, it would then use the symbolic representation of the equations of motion of an object under the influence of gravity (perhaps performing a numerical estimate of the effects of frictional resistance to an object of the shape of the apple) and compute the relationship between the time at which the wood was removed and the distance the apple had fallen. It would also compute the numerical value for the time at which the apple came into contact with the Earth. Of course, an error, either subtle or not subtle, could be present in the program so that the decoding would not be accurate. In contrast, the real apple really falling would always fall the right way, because the way it actually falls is, by definition, the right way for it to fall. Before computers were invented (or at least before the equations of motion were discovered), the only way that the effects of gravity could be decoded was by implementation. This same principle is true for the assembly of the polymers before genetic material and its decoding machinery came about.

The existence of an internal microenvironment would provide the potential, at first, for natural ordinary physical-chemical interactions of materials occurring in the environment, but with several important implications. First, the microenvironment is likely to be different from the bulk environment, presumably being more hydrophobic. The consequence of this is that rates and equilibria of physical-chemical interactions are likely to be substantially different from rates and equilibria of the same interactions in the bulk environment. Secondly, the concentrations of the ordinary chemicals inside the microenvironment might be substantially different from the concentrations in the bulk environment, since materials might diffuse between the bulk environment and the microenvironment very slowly.

Living objects had the potential to gain independence from the requirement of a constant external environment as a result of the sequestration of the internal microenvironment. There are two important points to make here. First, it isn't necessarily true that this would be possible, but apparently the laws of physics and chemistry, as well as the materials that were available in the environment, were such as to render this possible. Secondly, the independence of the constancy of the environment is only relative, since extreme changes (such as the replacement of oxygen with carbon monoxide) would still lead to our death.

The key step, other than the formation of a selectively-permeable membrane, would be the formation of internal genetic components, including an encoding of genetic information, machinery for the decoding of the genetic information, and machinery for the implementation of the products of the decoding. A computer search on Medline in July, 1996, uncovered in excess of one hundred recent papers on theories of the origins and evolution of these components. For example, a number of related papers were published in one issue of *The Journal of Molecular Evolution* (May, 1995). I will put forth my own speculative hypothesis, without providing adequate details or support to be considered scientifically sound:

The origins of the genetic code were, in the terms used above, more like hydrochloric acid digesting food than like calcium ions activating enzymes. That is, it devolved from primitive interactions of hydrophobic regions of two molecules, primitive interactions of hydrophilic regions of two molecules, or primitive interactions of oppositely-charged regions of two molecules. These interactions had nothing to do with a genetic code, and only by the most unimaginable (or perhaps too diversely imaginable) route found itself the basis for a genetic code. Many intermediate steps are now omitted. My interpretation of our contemporary genetic code is that the beginning of the penultimate stage of its evolution involved a triplet code, but with the first and third nucleotides of each triplet having no information content. The first and third nucleotides were only a spacer, possibly because of the steric requirements or the binding energy of the tRNA molecule that coupled the amino acid to the mRNA triplet. That is, the code was really a singlet information wise, but a triplet for practical reasons unrelated to the information content. Thus, there were four amino acids that could be encoded, or, more likely, four classes of amino acids, since each singlet code may not have been specific for a particular amino acid, but rather for a set of amino acids with some particular features in common.

The beginning of the final stage of its evolution involved a triplet code, but with the third nucleotide of each triplet having no information content. Again, the third nucleotide was only a spacer, presumably because of the steric requirements of the tRNA molecule that coupled the amino acid to the mRNA triplet. That is, the code was really a doublet information wise, but a triplet for practical reasons unrelated to the information content.

The final stage of its evolution involved the utilization of the third nucleotide for its information content. Examination of the "universal" genetic code (Li and Grauer, 1991) indicates that the code still isn't really a triplet, but more of a "two-and-a-half-let."

Under the influence of increasing selective pressure, three properties will be crucial for continued existence. First, self-regulated control of a constant internal environment in the face of a changing bulk environment. Second, the ability to sense and interpret the conditions in the bulk environment. These conditions include the direction of impingement of sunlight, the location of molecules that are needed, and the location of other objects to whom *it* constitutes molecules that they need. Finally, the ability to respond to what is sensed and interpreted. I equate "response" with "directed movement."

Given that self-regulated control of a constant internal environment has already come about through the formation of semi-permeable membranes, the ability to sense and interpret the conditions in the bulk environment and the ability to respond to what is sensed and interpreted would at first most likely be "combined" into one single function. The ability to respond would be dependent upon a source of energy, which at first would most likely be direct physical-chemical interactions with the light from the sun. The "combined" function would most likely involve a photochemical reaction of a polymer that resulted in a conformation change of the polymer that re-oriented the polymer so as to alter its ability to interact subsequently with the light from the sun.

At least four types of possible effects of this alteration can be hypothesized. First, the polymer might bend slightly so as to more efficiently interact with light from the sun, and thereby be able to extract more of the sun's energy. Second, the polymer might bend slightly so as to less efficiently interact with light from the sun, and thereby be shielded from potentially harmful effects of the light from the sun. Third, the polymer might open up or close more compactly so as to alter the interaction of some other material embedded within the internal microenvironment with the light from the sun. Finally, the polymer might change its conformation in such a way that a more bulky structure that the polymer was attached to would be re-oriented in such a way as to alter its interaction with the light from the sun. These types of interactions are analogous to the digestion of food by hydrochloric acid, being the direct physical-chemical response of a material to exposure to sunlight.

A somewhat more sophisticated arrangement would involve polymers that are attached to a main "body." These polymers would be able to follow a gradient of food by directly coupling the energy from the food source to a mechanical movement of the polymers that resulted in the translocation of the entire assemblage in the direction in which there is more food. This would work if the food that provided energy for the mechanical movement of the polymers also happened to be the same kind of food that the rest of the organism needed.

Eventually the sensing and interpretation of the external environment would become a separate function from the implementation of a response to the sensing and interpretation. This would require an internal mechanism for co-coordinating these activities, which initially would involve a local reflex response governed by a rudimentary "peripheral nervous system."

The final stage would be the development of a central nervous system (CNS), which could integrate the information from all of the sensing systems and respond by activating an integrated response of the mechanical movement system. These latter stages would undoubtedly require the pre-existence of a fairly well-defined genetic apparatus.

In a sense, this is where our story really begins. The existence of a fairly well-defined genetic apparatus implies that the Darwinian theory of evolution would apply. In a very brief summary, the starting point is an organism that exists and that has a genome that encodes mRNA, tRNA, and rRNA. The mRNA, in turn, encodes proteins. The genome is subject to recombination and to mutation. Mutation can involve as little as one nucleotide in the genome being replaced by a different nucleotide, or as much as duplication of major portions of the genome or even the entire genome itself. These mutations are more or less random events, and they are then "tested." The first test is whether conception can still occur. If not, then that mutation was not "successful." The next test is whether the offspring can survive until it has achieved reproductive age. There are two general reasons why it might not. One reason is intrinsic, in the sense that the mutation has lead to a disease. The other reason is extrinsic, in the sense that, although the offspring is healthy, it does not do well in the external environment that it is embedded in. The organism having survived to reproductive age, the third test is very similar to the second test. The only difference is that failing the second test is a sort of absolute and immediate answer to the question; failing the third test may require an extended period of time. The important point is that the tests are made in a particular environment in which the organism is embedded, and the only *direct* effect of Darwinian evolution is the result of the testing of the new phenotype by that particular environment.

Several analogies may be helpful here. The first one that comes to mind is a person with the following hypothetical qualities. The person is a male, is very tall, is adept at playing basketball, and is not very intelligent or very well educated. In an environment in which a large segment of the population is willing to either pay a substantial amount of money to attend a basketball game, or is willing to subject itself to watching commercials in order to watch a basketball game on television, then such a person may be very successful. On the other hand, if there is little interest in basketball within that population, then that person may be unsuccessful. It is still the same person, but the environment in which he is embedded is different.

Another example would be an intellectually-gifted woman who is committed to a life or a career focused upon some particular intellectual pursuit. Within a certain culture at a particular time, this woman may be systematically (or systemically) denied the opportunity to do so. At a different time, even within the same culture, such a person might be encouraged to do so. The point is that it is often not the intrinsic properties of the individual that are important so much as it is the characteristics of the environment in which the individual is embedded that is important to success or failure.

In summary, Darwinian evolution directly tests mutations strictly in regard to the current environment in which the organism is embedded, without regard for what the environment used to be or what the environment will become.

The rudimentary CNS was probably deterministic, in the sense that it was little more than a mechanical linkage between a sensory input from the external world and a predetermined movement of the organism that changed its relationship to the external world. For example, this altered relationship could involve a rotation or a translation to a new position in space, presumably to a location containing a better source of food. An organism that had a different type of response (for instance, mechanically moving to a location containing a poor source of food) would not be likely to survive. In order to link this to the last several paragraphs, it is sufficient to note that the food, which is a property of the external environment, would need to be of a type that would generate a sensory input in the organism.

A slightly more sophisticated CNS would be a minor variant of the rudimentary CNS. In this case, the CNS would be able to "integrate" two sensory inputs, for example a gradient of food and a gradient of temperature. At this point, the response would still be deterministic and mechanical.

The exact details of the evolution of the human brain are not important here. What is important is that, once a rudimentary CNS developed, three nontrivial conditions held. First, there were materials available that were compatible with living organisms such that they could form a human brain (since we are hypothesizing that human brains do exist). Second, a sequence of mutations in the genome could occur that eventually might lead to the formation of a human brain. Finally, each mutation in this sequence was such that, given the environment in which it was embedded, it (or the organisms that it was in) survived.

Dawkins' point of view is that the gene is a *replicator* and the body is the gene's *survival machine* (Dawkins, 1992). Given this perspective, the gene is the master of the organism, and the various organs and body parts are clearly subservient, being present merely to ensure the survival of the gene in the particular environment in which it is embedded. My impression is that this is the major "take-home message" that Dawkins is proposing. I find it particularly interesting that Dawkins (1987; 1992) frequently references Cairns-Smith, and Cairns-Smith (1987) references Dawkins. Cairns-Smith's book is entitled *"Genetic Takeover and the Mineral Origins of Life."* The main thrust of the work of Cairns-Smith seems to be that the crystal structure of clays served as a prebiotic gene, and that there was a genetic takeover by DNA. This would be something like humans building computers, then all traces of humans vanish, and computers wondering how they (computers) came about, since there was no longer any fossil evidence of what built them.

Another critical concept in thinking about evolution is that of *coincidence*. As discussed above, evolution occurs as the result of a mutation being tested in the current environment in which it is embedded. However, a molecule or a structure or a network that forms through this "orthodox" route may coincidentally have an effect (perhaps much later in geological time) that is totally unrelated to the effect that it had that allowed it to pass the test. As mentioned above, generic examples of this phenomenon are the ability to adapt and plasticity. A specific example is the ancient enzyme lactate dehydrogenase B4 (LDH B4) in ducks. LDH B4 is also ε-crystallin, which is a structural protein of the lens of the eye (Piatigorsky and Wistow, 1989). As pointed out by the authors, "The recruitment of enzymes to a structural role in the lens must have involved modification at some level of gene expression - pretranscriptional, transcriptional, or posttranscriptional - leading to increased amounts of gene product in the lens. The mechanism of this change has not been elucidated . . ." It is concluded that "Gene sharing may occur widely in nature." Thus, the fact that LDH B4 is also ε-crystallin is a specific example of a coincidence in evolution.

The last two paragraphs have brought forth the ideas of genetic takeover and coincidence. These two ideas are not entirely unrelated, since genetic takeover in itself is a sort of coincidence. That is, given the hypothesis that, at one time, crystals were genes and DNA happened to be associated with the crystals, then DNA happened to be a better genetic material than crystals and thereby effected a genetic takeover. The coincidence here is that DNA was presumably associated with the crystals for reasons entirely unrelated to its (DNA's) potential as a genetic material.

Now I would like to put forth a speculative hypothesis, that the following sequence of events occurred. The original genetic material was something like the crystals of Cairns-Smith, the polymers discussed above, or some such entity. In accordance with Cairns-Smith's view, there was then a genetic takeover by DNA. In accordance with Dawkins' view, the DNA was the master, and the assemblage of the organs (be it kidney, liver, heart, muscle, brain, etc.) was a survival machine for the DNA. This view clearly implies a certain subservience of the organs and body to the DNA. In particular, the brain, which mainly serves to detect food and cause the parts of the body that participate in mechanical movement to move to this food, is a part of the survival machine for the DNA. Once the neuronal circuits that served one or both of these two activities had evolved so as to be sufficiently complex, a strange coincidence occurred. These neuronal circuits (either as is, or just a mutation or two away) served certain other activities. I hypothesize this as a coincidence, since these other activities did not enhance the survival of the organism, and yet the neuronal circuitry that performed these activities (and, by inference, the genetic encoding for this circuitry) exists.

To be somewhat more specific, these activities are what we might think of as the human brain's properties of abstract thinking and intellect. It is relatively easy to justify considering these two properties as coincidental rather than as selected for by Darwinian evolution. The test is whether, at the time that they came about and in the environment in which they were embedded, did they confer a survival advantage? The answer is negative, since other organisms (whose brains we would consider not to possess these two properties) were able to survive in the same environment.

There is a fact in mathematics known as Gödel's Incompleteness Theorem (Dauben, 1979), which shows that, given any system of mathematical logic that is sufficiently complex (that is, rich enough to contain elementary arithmetic), then there exist theorems that are true but that cannot be proven or disproven within that system. This is a surprising consequence of the generic nature of systems of mathematical logic, one which the inventors did not intend to put into their systems. I would speculate that, in a similar manner, given any system of neuronal circuits that is sufficiently complex, there will be unanticipated consequences of the circuitry that were not intentionally built into it. Of course, in the case of biological systems, "intentionally built into it" really means "selected for by Darwinian evolution."

These two new properties of the human brain were the basis of a second kind of takeover, one in which the human brain became predominant. The genes had lost control of the mastery of the organism, since the brain was no longer subservient to it. By this hypothesis I do not mean that genetic engineering and recombinant DNA technology allow the human brain to directly alter the gene to suit its own purposes. Rather, I mean this biologically, not technologically. The human brain was freed from the biological subservience that is required of components of the gene's survival machine. The human brain was free to do things (for better or for worse) that are unrelated to the gene's survival. Of course, this is somewhat of an exaggeration in two regards. First, the gene encodes the information that makes a human brain (but the gene has little choice in the matter once the decoding machinery takes over). Also, only a portion of the human brain is free, since those activities that a nonhuman brain contributes as part of a survival machine must also be made by some portion of the human brain.

An enormous amount of research has been done on the brain. As new technologies (such as tomographic imaging and molecular genetics) emerge, our knowledge of the brain will increasingly accelerate. This research has been performed at many different levels, ranging from the extremely reductionist (such as the molecular biology of a single ion channel) to the integrative (such as statistical correlations in the activation of multiple subcomponents of the brain). My hypothesis is that, in order to understand the nature of prejudice, we need to understand the nature of the functioning of the human brain. Although the nature of the functioning of the human brain is ultimately dependent upon the reductionist elements, it is of little use *at this time* to try to understand prejudice by studying details at the molecular level. It seems more reasonable to study prejudice as being a product of the human brain, while keeping in the front of our minds that the human brain is an object that was initially the product of biological evolution, but which "escaped" from being only a part of a survival machine.

Part of our difficulty is that we are using the human brain to try to understand the human brain, and it may be that one of the results of the biological evolution of the human brain is an inherent bias in our thinking about ourselves. This bias could very well have served to enhance survivability given the environment in which humans were embedded at the time that this evolved. This would raise the issue of whether enhanced survivability is equivalent to enhanced accuracy of perception and interpretation, or whether sometimes a certain degree of inaccuracy of perception and interpretation can be more closely correlated with enhanced survivability. As an alternative, this bias might not have evolved directly as a result of Darwinian selection, but might have been the result of a coincidence, as discussed above.

The best way that I have discovered so far to overcome the difficulty described above (that is, that we are using the human brain to try to understand the human brain, and it may be that the human brain has an inherent bias in thinking about itself) is to try to put yourself into the place either of an imaginary being who is "outside" of the system in which we are embedded, or of a nonimaginary nonhuman, such as a cow (I don't know why, but I have a definite preference for looking at things from the point of view of a cow).

Another technique that can be developed with sufficient practice is based upon my hypothesis that many instances of human prejudice (either towards other humans or towards nonhumans) is based upon a combination of two factors: that they are self-serving and that they are based upon a relationship in which the human has some form of power or control over the object. This hypothesis suggests a way to test for the possible presence of prejudice: examine whether a particular belief is self-serving and/or based upon a relationship in which the human has some form of power or control over the object.

A particular example of this technique is afforded by a quotation from Per Brodal (1992):

> Experiments on animals are often criticized from an ethical point of view. The question of whether such experiments are acceptable, however, cannot be entirely separated from the broader question of whether mankind has the right to determine the lives of animals by using them for food, by taking over their territories, and so forth. With regard to using animals for scientific purposes, one has to realize that a better understanding of man as a thinking, feeling, and acting being requires,

> among other things, further animal experiments. Even though cell cultures and computer models may replace some, in the foreseeable future we will still need animal experiments. Computer-based models of the neuronal interactions taking place in the cerebral cortex, for example, usually turn out to require further animal experiments to test their tenability. Improved knowledge and understanding of the human brain is, however, also mandatory if we want to improve the prospects for treatment of the many diseases affecting the nervous system. Until today, these diseases - most often leading to severe suffering and disability - have only occasionally been amenable to effective treatment. Modern neurobiological research nevertheless gives hope, and many promising results have appeared in the last few years. Again, this would not have been possible without animal experiments.
>
> Yet there are obviously limits to what can be defended ethically, even when the purpose is to alleviate human suffering. Strict rules have been made by the governmental authorities and the scientific community itself to ensure that only properly trained persons perform animal experiments and that the experiments are conducted so that discomfort and pain are reduced to a minimum. Most international neuroscience journals require that the experiments they publish have been conducted in accordance with such rules.

In this case, it is clear that both tests apply: using laboratory animals to study human diseases is self-serving and is based upon a relationship in which the human has complete control over the laboratory animals. Therefore, Brodal's attitude is in danger of being a prejudice, and should be analyzed carefully to determine whether this is the case.

In fact, some aspects of Brodal's statement make me rather uncomfortable. First, he recognizes that there is a question concerning the ethics of using animals for experiments related to alleviating human disease. In a way, this realization imposes upon him an obligation that is perhaps not imposed upon biologists that are not aware of this question (or at least that was not imposed upon us in the past before this became a widely known issue). His strategy for dealing with this issue is perhaps somewhat facetious (although there is some ambiguity here). Basically, he is obliquely saying that, since it is okay for humans to kill animals for food and to drive them out of their territories, it is also okay to use them for scientific research. He is thus ducking the issue, and it is not clear whether he is saying "it is really not okay for humans to kill animals for food and to drive them out of their territories, but as long as we are, it is no worse to use them for scientific research," or whether he is saying "it is okay for humans to kill animals for food and to drive them out of their territories, and it is no worse to use them for scientific research, so using them for scientific research is okay too." Perhaps even more disturbing is his notion that the existence of strict rules that supposedly ensure that only properly trained persons perform animal experiments, etc. is an ethically acceptable solution. He is confusing the methods of implementation of a policy with the question of whether the policy should exist under any conditions.

In a more general sense, the issue of disease has some important ramifications. The word "disease" is usually interpreted by humans as meaning a human disease, and the consequences or implications of this human disease upon the human having the disease or upon that person's family, etc. (of course, veterinarians and people with sick pets provide exceptions to this generalization). Another way to look at disease would be to take a more global view. For example, if a human dies of a disease, then animals or plants which that person would have eaten may be spared. Thus, from the perspective of the particular animal or plant, certain human diseases may be equated with their own survival.

A somewhat more contentious example would be the death or disability of a human who would otherwise be a contender in the human job pool. For example, there may be a job announcement for which there are several hundred applicants, most of whom are nearly equally qualified to perform the job, the final decision on the part of the employer being somewhat subjective. If the person who would have been the first choice suddenly dies (of natural causes) then the person who was second choice would likely be offered this position. It may be that in a very competitive job market, this person who was second choice would have applied for hundreds of other positions, and never been first choice for any of them. After several years of frustration, this person might have suffered ill effects ranging from using up his or her life savings, suffering a divorce, or committing suicide. On the other hand, because of the death of the person who would have been the first choice, the second choice person who actually got the job did very well and prospered in life. In this case, this person would have benefited enormously as the result of the other person's disease and decease. In fact, to this person (although I am sure he or she would not think of it this way) the disease was a blessing. The point that I am making is that the word "disease" is not an absolute, but should be thought of in terms of the effects that it has more globally.

After all, there are entire professions that are dedicated to studying and preventing diseases, and many people make their livings in these professions. If all of a sudden there were no more human diseases, these people would be out of a job. So, in a sense, they benefit from the existence of human disease. Earlier, I had spoken of decoders of potential information, in the context of molecular decoders of molecular information. Now we have a new phenomenon, human (and computer) decoders of molecular information (Konopka, 1994). There are presumably two purposes of having human decoders of molecular genetic information. One purpose is purely intellectual, a pursuit of knowledge or the inherent challenge. The second purpose, which is probably why there would be funding available, is to explore the decoded information as a potential approach for developing therapies for human genetic diseases. Ironically, for the human decoder of molecular genetic information, this activity of decoding could be thought of as a survival mechanism, since in a restricted job market this could be a lucrative profession that could make the difference between desperation and (career) success for the individual. In a sense, human decoders of molecular genetic information are "parasites" (I mean this technically, not judgmentally) since the DNA has no "purpose" in existing, in particular it does not have the purpose of allowing a human to "make a living" out of decoding its encoded information.

The other ramification that I want to address is concerned exclusively with human genetic diseases, and their relationship to biological evolution. For, during the billions of years of biological evolution, mutations have occurred that had the following three properties: (a) The individual with the mutation was recognized by the others of that species as "having a disease;" (b) that individual suffered some disadvantages as a result of that recognition; and (c) that mutation turned out to survive (that is, it was selected for in Darwinian evolution). The point is, that a particular mutation was viewed as a disease, but ended up by conferring a survival advantage. Mutations that lead to *perceived* diseases are the stuff of which evolution is made, and we would not be here today if this hadn't happened countless numbers of times over billions of years.

Now take the case of two actual contemporary human diseases, attention deficit disorder and clinical depression. A person with clinical depression is unhappy (to say the least) because of the intrinsic nature of the disorder. Attention deficit disorder is a very different matter. A child with attention deficit disorder is intrinsically very happy, energetic, and uninhibited. Any negative aspects of attention deficit disorder are extrinsic; that is, the constant disapproval coming from someone else (parent, teacher, etc.). A person with clinical depression can feel intrinsically better by using proper medication.

On the other hand, the medication that is administered to a child with attention deficit disorder can make the child feel better extrinsically by lowering the amount of negative feedback, but this medication fails to make the child feel intrinsically better, since the child was already intrinsically very happy to begin with. Thus, there are two reasons that a child with attention deficit disorder is given medication. One reason is that the child would not otherwise be able to function well in school, and the parents and teachers know (or believe) that, given the environment in which we (including the child) are all embedded, functioning well in school is important. The second reason is that it is very unpleasant (to say the least) for a "normal" adult to spend any amount of time (say even several milliseconds) with an unmedicated child with attention deficit disorder. Thus, at least the second reason is clearly one of self-interest; that is, to use medication to control another person whose behavior is annoying. Even the first reason might be called into question, since who is to say that the intense energy of the unmedicated child with attention deficit disorder would not eventually lead to a degree of creativity that may very well be stifled by routing through traditional schooling?

One of the basic implicit assumptions that the human brain makes is that its perception of external reality is accurate. One of the major routes by which the human brain perceives external reality is through its visual system. The visual system does not work like a camera. A camera essentially forms a two dimensional representation of the projection emanating from a three dimensional scene. There is a direct point-by-point correspondence of the two dimensional representation and of the projection. This direct correspondence ensures a certain degree of accuracy in the depiction of the projection of the three dimensional scene.

The human brain forms an image in quite a different manner, as described by Posner and Raichle (1994):

> Recordings from individual cells have shown that the visual system is organized into maps, each of which represents a kind of "picture" of the scene viewed by the retina. The monkey brain contains, at last count, some 34 different maps of the monkey's visual world, each analyzing a different attribute such as color or orientation, and the human brain is believed to be similarly structured. . . .
>
> The maps are further organized in a hierarchical manner. Each map corresponds to a distinct area of the brain responsible for carrying out a particular type of analysis of the visual information sent to the brain. Thus our recognition of a particular visual pattern appears to involve a set of areas that go from the primary visual area in the back of the brain out into the temporal lobes in the lower middle of the brain. This set of areas has been called the "what" pathway. Within the "what" pathway there exist regions that are specialized to recognize such attributes of the stimulus as shape, color, and speed of movement. On the other hand, information about where an object resides in our visual environment seems to concern areas that, again, go from the primary visual area, but in this instance ascend into the parietal lobes in the upper middle of the brain.

Posner and Raichle further explain that

> The primary visual cortex, which responds optimally to bars and edges, is joined in the analysis of this complexity [the scenery of our everyday visual world] by a large number of so-called extrastriate visual areas. These areas are situated adjacent to the primary visual cortex in the occipital, parietal, and temporal lobes.

Our visual world results from the differential activation and interactions between these multiple visual maps, each of which is specialized to a particular type of visual feature. This mechanism is entirely different from the formation of an image by a camera. The complexity of this mechanism might lead us to question the objectivity with which we visually experience our external reality. This mechanism is the product of biological evolution. As such, it came about because it conferred a survival advantage. Thus, the visual system might present us with a view of external reality which is (or at least at one time was) optimal for survival, which is not necessarily equivalent to being accurate or complete. In fact, the visual system selectively allows us to see light that falls within the visible portion of the electromagnetic spectrum. Thus, infra red, ultraviolet, x-rays, gamma rays, etc. are transparent to us. If they were not transparent to us, then we might have a more complete visual picture, but one which would be too complex to be useful for survival.

A subject that is very contentious is the question of "nature *versus* nurture" or genetic predisposition *versus* cultural influence. Related to this issue is the question of the existence or nonexistence of free will. In a sense, the issue of free will is very similar to that of God. That is, free will either exists or it does not exist, and a given individual may correctly or incorrectly believe either. It may seem strange, but if free will, in fact, does not exist, then someone who believes that free will exists will believe so incorrectly. Yet, since free will, by hypothesis, does not exist, this person had no choice in having this incorrect belief.

The question sometimes comes up as to whether certain behaviors of nonhuman animals (often one of the great apes) represents the ability to learn and to think, or is the result of "genetic imprinting." Human behavior is usually used as the "gold standard" for attempting to answer such questions, since the nonhuman animal's behavior is evaluated by its degree of similarity to analogous human behavior. Dawkins (1985; 1992) gives many examples of animal behavior which could be interpreted by a human in an anthropomorphic manner (that is, by the human "projecting" or attributing human-like motivations, for example, altruism, to the behavior), but which in reality are examples of genetically-imprinted Darwinian survival. The use of human behavior as the gold standard for evaluating nonhuman animal behavior, may, in a sense, be begging the question. This paradigm makes the assumptions that human behavior is not the result of its own genetic imprinting and that there is such a thing as free will (otherwise, human behavior would only apparently, but not really, involve making decisions or choosing alternatives, etc.).

Now I have in mind an example that I will introduce by reluctantly but honestly classifying as a sexist example. It is not inherently sexist, and it could have been told the other way, but since I am a male and the story is about me, it turns out to be sexist. When I was growing up, I was told by someone (probably an older relative) that a girl's (the sexist part is not that I use the word "girl" instead of "woman," because a female of my age at that time would have been a girl) appearance wasn't as important as her personality. Let us translate this statement from English into the language of the gene, in an admittedly oversimplified manner, but one which captures the essentials.

First, realize that there are three different "genome products" here: the perceiver (I suppose this is me), the perceived (the girl), and the person giving advice (the person giving advice). The two relevant parts of the genome of the girl are the gene for appearance and the gene for personality. Assume that in a particular position of the gene for appearance (let us call this position 1), of the four possible nucleotides (A, C, G, T), an A renders the girl unattractive and a G renders the girl attractive to me. Likewise, assume that in a particular position of the gene for personality (let us call this position 2), a C renders the girl not personable and a T renders the girl personable to me. It is unimportant (in this example) whether these traits of attractiveness or personableness are universally true for all perceivers, for all male perceivers, or just for me. The translation of the statement (of the person giving advice) from English into the language of the gene would then be "It is more important for a girl to contain a T in such and such a position in her genome than it is for her to contain a G in such and such another position in her genome."

Now, I also have a genome, and let us say, for the sake of argument, that my genome contains a gene that has the following property: If a certain position within that gene (let us call this position 3) contains a T, then it will be more important to me that the girl be attractive. On the other hand, if that position contains a C, then it will be more important that the girl be personable. Then we have the following strange (although probably extremely common) situation: If my genome has a T in position 3, I will have a preference for a girl who has a G in position 1; if my genome has a C in position 3, I will have a preference for a girl who has a T in position 2.

We generally value a person for his or her inner qualities more than for his or her superficial qualities, but this really amounts to a bias that having a particular nucleotide sequence in a particular location within the person's genome is more valued than having another particular nucleotide sequence in another particular location within the person's genome.

To summarize the most important concepts thus far, and to set the stage for the statement of my own explicit hypotheses about the functioning of the human brain:

1. Biological evolution by means of Darwinian selection means that a given atom exists in a particular spatial and temporal relationship with a set of other atoms, rather than doing something else. By "spatial" I mean that the atom is part of a structure arranged in three dimensional space. By "temporal" I mean that this structure is really a dynamic infrastructure.

2. Direct Darwinian selection results in those dynamic infrastructures that are able to survive or exist in the environment in which they are embedded at the time of the selection.

3. Coincidences can occur as a result of direct Darwinian selection that were not themselves directly selected for.

4. The human brain is itself the product of Darwinian selection. It was directly selected for not on the basis of whether it could perceive external reality accurately and interpret its perceptions accurately, but rather because it was able to survive or exist at the time of its selection.

5. The human brain's tendency to categorize people on the basis of easily observable phenotypical characteristics was directly selected for by Darwinian evolution.

6. Coincidence resulted in the human brain's partial takeover of control from the genome.

There is no doubt that the human brain tends to categorize people, and this type of activity can be roughly thought of as a "horizontal" form of bias. I hypothesize that a much more important "vertical" form of bias occurs. In addition to forming categories, the human brain creates a *hierarchy*. I hypothesize that the hierarchy came about in the following way: The human brain (or the direct precursor to it) was selected by direct Darwinian evolution. The neuronal circuitry at this point was so "rich" that coincidence was able to result in the human brain's partial takeover of control from the genome. Coincidence again gave rise to self-awareness, intellect, and the ability for abstraction.

It is with the rise of these three attributes - self-awareness, intellect, and the ability for abstraction - that two unique needs of the human brain came about: a need for it to think of itself as having a purpose, and a need for it to think of itself as being special. I do not believe that these two needs could come about unless all three of these attributes were present. Conversely, I could imagine that these three attributes, in themselves, are not sufficient to produce these two needs. However, any object with these three attributes that did not produce these two needs would probably not survive, for then the prospect of the negative aspects of existence (disease, dying, labor, competition for survival, needing to stalk prey, needing to avoid predators, etc.) would certainly make it seem not worthwhile to do those things that are necessary for survival.

To restate this in a somewhat different manner, when the organism had been simply a survival machine for the gene, then, even though there may have been no objective reason to try to survive, the survival machines survived (or sometimes failed to survive) anyway, because they had no choice. They simply mechanically acted out what they were programmed to do. After the takeover of the human brain, the situation changed. Then the human brain, with its self-awareness, intellect, and the ability for abstraction, could question whether it was worthwhile to survive. Unless it thought of itself as having a purpose or as being something special, the answer to that question would have been that it wasn't worth the effort to survive.

At this point, I would like to briefly re-examine Dawkins' concept of survival machine, since, in a very subtle, almost intangible way, this concept may be remotely tied in to a hierarchical bias. First, it is important to realize that the concept of a survival machine is just a concept or an interpretation. The fact that we have is that there are genes that encode body parts that the genes end up being embedded in. Another interpretation of this fact is that all of the body parts, including proteins, RNA, etc., are parasites taking advantage of the DNA. In other words, nobody left a taped message for the DNA saying "Your mission, DNA, is to code for body parts. Should you decide to accept this mission, . . . " After all, the DNA really had no mission, and it certainly had no decision as to whether it was going to be used as a code for something else. So, in a subtle way, Dawkins' survival machine concept may be thought of as a "back door" through which a teleological purpose for DNA is postulated, but without explicity making it clear that there is a teleological element to this.

Related to this is the whole issue of the so-called C-value paradox: the amount of DNA in some genomes is much larger than required to encode proteins and functional RNAs (Konopka, 1994). This is a paradox only if there is an implicit teleological purpose for DNA. Otherwise, why can't DNA just be whatever size it happens to be, not expected to be optimized in some way for a "mission?"

Thus, the human brain has five unique properties: self-awareness, intellect, the ability for abstraction, thinking of itself as having a purpose, and thinking of itself as being special. These properties are necessary and sufficient for the human brain to create a hierarchy, with itself at or near the top. It is interesting to compare this concept of hierarchy with Allport's concept of category. The main difference is that category does not necessarily carry with it a value judgment, although it certainly can do so sometimes. As Allport has pointed out, categories can sometimes be rational. On the other hand, The concept of hierarchy is, in and of itself, a value judgment, totally devoid of any elements of rationality. The concept of hierarchy is the human brain's way of solving an unsolvable dilemma: the absurdity (in the sense of Theater of the Absurd) of the existence of an object with self-awareness, intellect, and the ability for abstraction in a situation where there is no objective meaning (at least none that we can be aware of or certain of). The only "solution" is for the intellect to invent an abstraction.

Recall at this point the earlier comments about what I termed the "brain pool." In particular, think of the brain pool as constituting the basis of a histogram or frequency distribution. Some, but certainly not all, of the brains in the brain pool subscribe to a "solution" that is both diabolical and ingenious. It is diabolical because it cannot be disproved on a rational basis, and it is ingenious because it simultaneously serves three purposes: It allows the human brain to think of itself as having a purpose and as being special; it allows the human brain to find an apparent answer to certain intellectual questions that are not answerable given the state of technology or the knowledge base at the time; and it allows the human brain to have beliefs that are in harmony with the beliefs of a vast number of other human brains (and thereby be "accepted" by these other brains). This particular "solution" is to invent the concept of God, to equate this invention with an external reality, and to forget that this was originally an invention and to think of the invention as actually being an external reality.

Of course, it is possible that there is such a God in external reality. If this is so, it is purely accidental that the invention and the external reality coincide. This situation is reminiscent of Gödel's Incompleteness Theorem, in that there is a true statement whose truth cannot be demonstrated from within the system. It is like a multiple choice test in which each question has five possible mutually-exclusive answers, nobody knows the correct answer, everybody randomly guesses one of the five answers, and one-fifth of the people are correct, but purely by accident.

This "solution" is ingenious because God is special and important, and man (for "man" substitute "the human brain") is in a special relationship with God, which then makes man special and important. But there are questions for which answers are not available, such as about the creation of the Universe, etc., and it is convenient to have God rather than man at the top of the hierarchy in order to feel that, although we do not know the answer, the answer is known one step further up the hierarchy. This mechanism relieves us of the anxiety of uncertainty. Finally, this "solution" is self-perpetuating, since the human brain is more comfortable when its beliefs are in agreement with the consensus.

An alternative that is used by some of the other brains in the human brain pool would be to dispense with the idea of a God entirely, and to put the human brain at the top rather than at the second position of the hierarchy. Although this alternative differs in detail from the first one, the compelling property that they both have in common is the existence of a hierarchy, and there is little or no reason for us to distinguish between them.

In a general sort of way (but with occasional exceptions that in no way invalidate the generalization) the rest of the hierarchical structure is formed on the basis of "similarity to human." From top to bottom, we have: self, family/close friends, other humans, nonhuman animals, plants, nonliving objects. Micro-organisms and insects probably are at about the same level as nonliving objects.

A particularly interesting example is that of a person who is a vegetarian, not for health or for ecological reasons, but for a moral reason. In particular, the moral reason is presumably that it is immoral to kill another living animal for food. This philosophy is an example of hierarchical prejudice, in that there is no moral reason to choose to consume a portion of a plant rather than a portion of a nonhuman animal. The only apparent reason is that, from the perspective of a human brain, based upon the hierarchy, the nonhuman animal is closer to the human than is the plant. That is, both share the property of being animals, while the plant does not.

One might even carry the concept underlying vegetarianism one step further: since plants and humans share the common property of being living objects, one might be tempted to argue on moral grounds that eating a plant is worse than eating a substance composed of nonliving materials (such as synthetic vitamins, etc.). Again, this is simply a hierarchical argument, and, looked at from a perspective other than that of the bias of the human brain, there is no difference between living and nonliving materials. This is probably a very good spot to reiterate that if the (generic) human brain that did not have this bias, then it would not have survived or existed, for then it would not value living *versus* nonliving, it would therefore not go to the trouble of doing what was needed in order to remain alive, and it would certainly cease to exist.

There is an old saying: **Matter can be neither created nor destroyed, only changed from one form to another.** In spite of the fact that this saying ignores transformations between energy and matter, it can be thought of as applying to "ordinary chemical reactions." This saying is so familiar that I doubt that I would even realize if someone had just said it to me, but it is probably the most useful (if not the only) approach to "correcting" the human brain's propensity towards hierarchical bias. The entirety of external reality can be perceived as matter being changed from one form to another. That which survives or exists is an infrastructure that provides a form to which matter is converted or a form from which matter is converted, in the sense of the dynamic infrastructure discussed earlier. The property of being "living" or "nonliving" is totally irrelevant and is purely an abstract conception of the human brain. Survival or existence of an object A necessarily means extinction or lack of existence of some other object B. Objects A and B may either or both be living or nonliving.

We usually think of Darwinian evolution as "survival of the fit." That is, a mutated gene that makes an animal run slower is likely to not survive in the gene pool. Because of the harsh reality of competition for limited resources, this one-sided view of Darwinian evolution may be adequate for most purposes. However, it is not too difficult to conceive of a mutated gene that makes an animal run faster. If this animal can run sufficiently fast, then it may happen that this animal will eventually (or quickly) consume all of its prey, and thereby doom itself to extinction. Now, it may be a matter of semantics (or a matter of circular reasoning), but then one could simply say that this animal wasn't fit since it didn't survive. On the other hand, one could say that the animal was too fit, and therefore didn't survive. Looked at from this perspective, one could say that there is a spectrum of fitness, and within this spectrum there is a certain window of survivability; an optimum, above which or below which survival is less probable.

The human brain most likely has survived and exists because it fits into a window of survivability. For example, consider the normal type of categorization discussed by Allport. This type of categorization is likely to be necessary for human survival (in a Darwinian sense), and yet it represents a "defect" (in an absolute sense), since it is subject to substantial error. At the risk of belaboring the point, someone might have an appearance that is often associated with a particular set of character traits, although this person's character traits happen to be very different.

As discussed in some detail above, God may or may not exist. Independently of whether God does exist or not, many human brains have the *need* to believe that God exists and to believe that the belief that God exists is not only a personal belief, but that God is an external reality. Now, it is fortunate for those human brains for which this is true that it is also true that they are *able* to have this set of beliefs. That is, a human brain could probably not survive if, on the one hand, it needed to believe that God exists and to believe that the belief that God exists is not only a personal belief, but is an external reality, but on the other hand, it was not able to have this set of beliefs (presumably because it was a very "rational" brain and could not accept the non rational, however desperately it wanted to). Thus, in this case, the "defect" of the human brain to be able to accept the non rational might be thought of as constituting a survival mechanism in a Darwinian sense.

Another example was discussed above, in which the possibility that it may be the case that there is no such thing as free will. It is probably a property of most human brains that a belief in free will is essential to survival. This belief would be equally essential for survival regardless of whether free will actually exists or not.

The hierarchical bias of the human brain being special is probably another belief that most human brains would require in order to survive. Yet, as I indicate here, there may very well be nothing special about the human brain *except* for the hierarchical bias of the human brain that it is special. Thus, the *ability* of the human brain to have this bias that it requires in order to survive is itself essential to the human brain's survival, even though in an "absolute" sense the ability to have an incorrect bias might be regarded as a defect.

These last few paragraphs can be summarized by noting that there are two different properties of the human brain that might be confused as being just one property. These two properties are: (a) A *need* to have certain non rational beliefs, and (b) the *ability* to have non rational beliefs. The combination of (a) and (b) might be termed "plausible denial." The ability to have a non rational belief in the face of a need to have it is probably essential for survival. What about the need to have a non rational belief? The need to have a non rational belief is probably also related to survival, although it may not be essential. For example, given a capitalistic economic system and a level of technology in which the "legitimate" rational needs of the entire population can be taken care of by a small fraction of the population, the existence of non rational needs provides the opportunity to lower the unemployment rate. Otherwise, the welfare burden would be unacceptable to those few who were employed "legitimately," and the competition for the few "legitimate" occupations would lead to a way of life that would not be conducive to survival.

An analogy that comes to my mind is that external reality is like a hand upon which the human brain superimposes a glove. The glove is related to the underlying reality of the hand in that it has five fingers and a palm, but it represents an overlay upon external reality in that, in essence, it completely masks the true external reality and creates a rendering of external reality that is conducive to human survival.

Another analogy is that of a writer, for example, who happens to not be "computer literate," and uses only a word processing program, which someone else had installed on a Macintosh computer. Because of the human-machine interface of a properly designed Macintosh application, to this person, the Macintosh computer appears to be not a computer, but to be a word processor. I could even conceive of such a person as being surprised to eventually find out that he or she had been using a computer all this time. In this example, the human-machine interface (corresponding to the glove in the paragraph above) is the overlay upon the underlying reality (corresponding to the hand in the paragraph above).

The concept of plausible denial leads to the following personal speculation, which I put forth with entirely inadequate evidence. Acquisition of this evidence (or evidence to the contrary) would constitute an interesting sociological study. My speculation is that the human brain has such a strong bias to think of itself as being something special, that a dilemma is faced by the brains of scientists who study DNA sequences, molecular evolution, and biological evolution. This dilemma is based upon a dichotomy between the conscious intellect and this bias, which is probably functioning at a level other than the conscious (even for those of us who do have a conscious awareness of it; there is probably no escaping it). The dichotomy is that, as a result of their field of study, these scientists are aware intellectually of how we came about as a result of Darwinian evolution totally devoid of teleological properties.

As a matter of fact, there are many published attempts to construct phylogenetic trees, that include humans, based upon the degree of molecular sequence similarities. On the other hand (and this is the point of my speculation) the same scientist, having just completed the construction of such a phylogenetic tree, is probably in a state of denial (at least subconsciously) about the implications of his or her own research. The implications, of course, being that there is nothing special about the human brain in general, and about his or her brain in particular. That is, in spite of evidence to the contrary, and in spite of publishing papers to the contrary, deep down we still believe that there was a teleological drive that lead to the existence of the human brain, and that the human brain was the final goal of evolution all along. Thus, we probably all have a version of that inconsistency which I have attributed to Eccles many pages ago. This inconsistency evolved in a Darwinian sense, for without it an object with consciousness or self awareness probably would not survive or exist.

It would be appropriate to end with one final word about the influence of biotechnology upon the type of categorization described by Allport. The property of the human mind to categorize is not the result of a conscious effort, but is the result of biological evolution over a period spanning millions of years. In a sense, it persisted because at the time that it first evolved it "worked," and it became "locked into" the system. That is, at one time it conferred a survival advantage. Now we seem to be at a transitional period, where it is still, to some extent, necessary to function by categorization. It may be that survival will depend upon whether we have the ability to consciously overcome or learn to control our propensity to inappropriate categorization.

In particular, up to the present time, categories have been based upon visible cues. Individuals in these categories then "inherit" properties that seem to be correlated with them. On the other hand, now that the human genome is being sequenced, a new form of categorization is possible. This form of categorization may differ from the traditional form in two respects. First, it will not be based upon visible cues, and it will therefore require more of a conscious effort. That is, the genotype (not visible), rather than the phenotype (visible) will form the categories. Secondly, the correlation of categorical attribution may be more precise when it is based upon a genotype than when it is based upon a phenotype. That is, a particular DNA sequence within the genotype may, in fact, be correlated very strongly with personal characteristics or behavior, as has been found to be the case for certain disease states. This would be in contrast to the inaccurate stereotyping that occurs based upon phenotypes.

Finally, it should be mentioned that not all categorizations based upon phenotype are inaccurate. For example, many individuals of a particular phenotypical category may have undergone similar experiences in the past, possibly because of the attitude of the in-group to that phenotype at that time. The result may be that these individuals now do share a common behavior or attitude, and thus constitute a category with a strong correlation to the phenotypical category.

BIBLIOGRAPHY

Allport G.W., *The Nature of Prejudice* (Addison Wesley, Reading, MA, 1954).

Atkins P.W., *Creation Revisited* (W. H. Freeman and Company Limited, New York, 1992).

Baring-Gould W.S., *The Annotated Sherlock Holmes (vols. I and II)* (Clarkson N. Potter, Inc., New York, 1967).

Barrow J.D. and Tipler F.J., *The Anthropic Cosmological Principle* (Oxford University Press, New York, 1994).

Brodal P., *The Central Nervous System Structure and Function* (Oxford University Press, New York, 1992) pp. 98, 99.

Burbaum J.J. and Schimmel P., Assembly of a Class I tRNA Synthetase from Products of an Artificially Split Gene, *Biochem.* vol. 30, pp. 319-324 (1991).

Cairns-Smith A.G., *Genetic Takeover and the Mineral Origins of Life* (Cambridge University Press, Cambridge, 1987).

Dauben J.W., *Georg Cantor His Mathematics and Philosophy of the Infinite* (Princeton University Press, Princeton, New Jersey, 1979).

Dawkins R., *The Blind Watchmaker* (W.W. Norton and Company, New York, 1987).

Dawkins R., *The Selfish Gene* (Oxford university Press, Oxford, 1992).

Eccles J.C., *Evolution of the Brain: Creation of the Self* (Routledge, London, 1989).

Goffeau A., Barrell B.G., Bussey H., Davis R.W., Dujon B., Feldman H., Galibert F., Hoheisel J.D., Jacq C., Johnston M., Louis E.J., Mewes

H.W., Murakami Y., Philippsen P., Tettelin H., and Oliver S.G., Life with 6000 Genes, *Science* vol. 274, pp. 546-567 (1996).

Hamming R.W., *Coding and Information Theory* (Prentice-Hall, Inc., Englewood Cliffs, N.J., 1980).

Hartmann W.K. and Miller R., *The History of Earth* (Workman Publishing, New York, 1991).

Kauffman S.A., *The Origins of Order* (Oxford University Press,Oxford, 1993).

Konopka A.K., Fundamentals of Biomolecular Cryptology, in D.W. Smith, Ed., *Biocomputing Informatics and Genome Projects* (Academic Press, Inc., San Diego, California, 1994), pp. 119-174.

Li W-H and Grauer D, *Fundamentals of Molecular Evolution* (Sinauer Associates, Inc., Sunderland, Massachusetts, 1991).

Marks J., Chromosomal Evolution in Primates, in S. Jones, R.D. Martin, and D.R. Pilbeam, Eds., *The Cambridge Encyclopedia of Human Evolution* (Cambridge University Press, Cambridge, 1992), pp. 298-302.

Piatigorsky J. and Wistow G.J., Enzyme/Crystallins: Gene Sharing as an Evolutionary Strategy, *Cell*, vol. 57, pp. 197-199 (1989).

Posner M.I. and Raichle M.E., *Images of Mind* (W.H. Freeman and Company, New York, 1994).

Press W.H., Teukolsky S.A., Vetterling W.T., and Flannery B.P. *Numerical Recipes in C The Art of Scientific Computing* (Second Edition) (Cambridge University Press, Cambridge, 1992), pp. 714-722.

Riedel H., Dull T.J., Schlessinger J., and Ullrich A., A Chimaeric Receptor Allows Insulin to Stimulate Tyrosine Kinase Activity of Epidermal Growth Factor Receptor, *Nature*, vol. 324, pp. 68-70 (1986).

Schlesinger W.H., *Biogeochemistry an Analysis of Global Change* (Academic Press, Inc., San Diego, California, 1991).

Stanley S.M., *Extinction* (W.H. Freeman and Company, New York, 1987).

Wallace T., unpublished communication (1996).

Zehavi A.M., *Handbook of the World's Religions* (Franklin Watts, New York, 1973), p. 163.