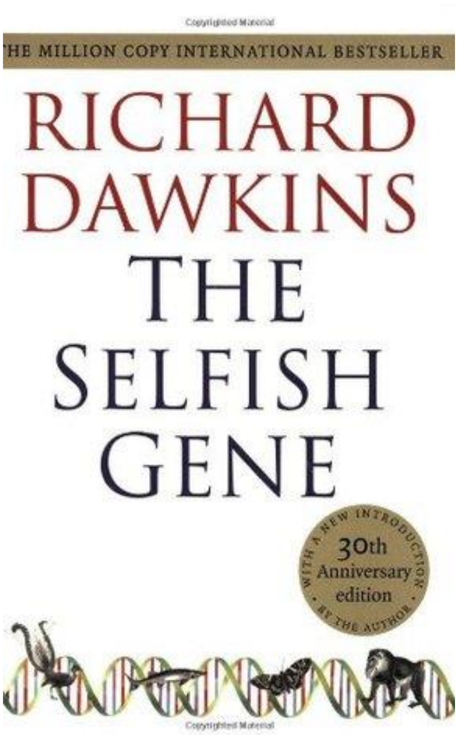
# Classic Text 10 - Nice Guys Finish First




Richard Dawkins' "The Selfish Gene," first published in 1976 is a book length argument in favour of a gene-centred view of evolution as opposed to the traditional organism or dissimilar group centred view. Although not a philosophical text in the familiar sense, "The Selfish Gene" has been included as recommended reading for an increasing number of introductory courses in Philosophy, for the simple reason that, if understood correctly, the gene's-eye view challenges many of our cherished ethical assumptions.

For copyright reasons we cannot reproduce the full text here, though we strongly recommend that the text be read in full so that the arguments presented below may be appreciated in context. Ch. 12 "Nice guys finish first" is however reproduced here because it provides some interesting insights into one game theoretical model of conflict and cooperation.

Although there are several chapter-by-chapter summaries of varying quality available on-line, nothing compares to actually reading the book for yourself and drawing your own conclusions. What follows is not a summary of the book but a reconstruction of its central arguments. Although a rudimentary grasp of genetics is necessary to understand the arguments from "The Selfish Gene" in this discussion, **appendix A** is intended as a rundown of relevant biological information, for the uninitiated. However, any senior secondary level Biology textbook will comprise of a more leisurely discussion. Note all page references to "The Selfish Gene" refer to the 30th anniversary edition, which is currently on sale.

On the first page of the first chapter of "The Selfish Gene" Dawkins laments: "Philosophy and the subjects known as 'humanities' are still taught almost as if Darwin had never lived. No doubt this will change in time." Hopefully this discussion will be part of that a change. In the same paragraph Dawkins sets out the unambiguous purpose of his book, namely "to examine the biology of selfishness and altruism." This is of enormous importance to philosophers of every stripe today because for millennia before Darwin or Dawkins, philosophers have been discussing just such issues under the rubric of Ethics and much of what has been said from Plato to Kant and beyond  has been taken on trust by many.

To begin we need a definition of altruism and selfishness (and welfare besides):

> An entity, such as a baboon, is said to be altruistic if it behaves in such a way as to increase another such entity's welfare at the expense of its own. Selfish behaviour has exactly the opposite effect. 'Welfare' is defined as 'chances of survival', even if the effect on actual life and death prospects is so small as to seem negligible...

> It is important to realize that the above definitions of altruism and selfishness are *behavioural,* not subjective... My definition is concerned only with whether the *effect* of an act is to lower or raise the survival prospects of the presumed altruist and the survival prospects of the presumed beneficiary. (p. 4)

Altruism in this sense is not to be confused with the product of the now discredited notion of group selection where one individual makes a sacrifice 'for the good of the species or 'for the good of the group' - the sort of stock in trade commentary that one has become used to in wildlife documentaries. The same is true of the individual-selectionist viewpoint in which people restrain their own selfish greed for the greater good of the group. The fact is that, although individuals may be possessed of foresight, "evolution is blind to the future." For Dawkins then, "…the fundamental unit of selection, and therefore of self-interest, is not the species, nor the group, nor even, strictly, the individual. It is the gene, the unit of heredity." (p. 11)

In the second chapter "The Replicators" Dawkins sketches some of the conditions necessary for the emergence of life. The details are admittedly speculative because we have no record of the earliest life forms, however we can be confident of some of their general characteristics:

> Darwin's 'survival of the fittest' is really a special case of a more general law of *survival of the stable*. The universe is populated by stable things. A stable thing is a collection of atoms that is permanent enough or common enough to deserve a name… [and] the earliest form of natural selection was simply a selection of stable forms and a rejection of unstable ones. There is no mystery about this. It had to happen by definition. (p. 12-13)

At some point, quite by chance, a stable molecule became capable of making copies of itself, perhaps by serving as a template for a duplicate molecule or perhaps via a process more akin to the growth of inorganic crystals. These were the first **replicators**. However improbable such a molecular accident, it only needed to happen once and given the hundreds of millions of years that intervened between the birth of this planet and the first glimmerings of life, it must have been a statistical inevitability.

We have no idea what such replicators looked like but we can be near certain that they were a far cry from the sophisticated DNA replicators of today. However, like all replicators they must have been imperfect copiers.

> [E]rratic copying in biological replicators can in a real sense give rise to improvement, and it was essential for the progressive evolution of life that some errors were made. …in any case we may be sure that mistakes were made, and these mistakes were cumulative. As mis-copyings were made and propagated, the primeval soup became filled by a population not of identical replicas, but of several varieties of replicating molecules, all "descended" from the same ancestor. (p. 16-17)

Some replicators would have been inherently more stable than others, hanging around longer and in so doing, making more copies of themselves than others, which were perhaps more prone to chemical or thermal instability or exposure to harsh UV radiation on an early world without an ozone layer, all of which would have been the source of differential survivability. Moreover no replicator survives alone: This was Darwin's point about competition.

> There was a struggle for existence among replicator varieties. They did not know they were struggling, or worry about it; the struggle was conducted without any hard feelings, indeed without feelings of any kind. But they were struggling, in the sense that any mis-copying that resulted in a new higher level of stability, or a new way of reducing the stability of rivals, was automatically preserved and multiplied. The process of improvement was cumulative. (p. 19)

To sum up, the properties that a successful replicator would have had would have been longevity, fecundity as well as high, but not perfect, copying fidelity. Some replicators, as Dawkins continues, would have gone on to evolve as "proto-carnivores," breaking apart the molecules of competitors, who would in turn, over time, evolve physical or chemical barriers by which to protect themselves. The upshot of which would have been that…

> "[those] replicators that survived were the ones that built *survival machines* for themselves to live in… Survival machines got bigger and more elaborate, and the process was cumulative and progressive… They are in you and in me; they created us, body and mind; and their preservation is the ultimate rationale for our existence. They have come a long way, those replicators. Now they go by the name of genes, and we are their survival machines." (p. 19-20)

Ch. 3, "Immortal Coils," explains the basic process of gene replication as well as the process of crossing over and "gene shuffling" that occurs during meiosis. (See Appendix A) Two unique insights however stand out from the customary textbook descriptions, namely:

1. A body is the genes' way of preserving the genes unaltered... [and that]… natural selection favours replicators that are good at building survival machines, [and] genes that are skilled in the art of controlling embryonic development. A survival machine is a vehicle containing not just one gene but many thousands. The manufacture of a body is a cooperative venture of such intricacy that it is almost impossible to disentangle the contribution of one gene from that of another.
2. The combination of genes that is any one individual may be short-lived, but the genes themselves are potentially very long-lived. Their paths constantly cross and recross down the generations. One gene maybe regarded as a unit that survives through a large number of successive individual bodies. (p. 23-24)

Although what Dawkins states are simply matters of biological fact, he has swung the focus around from the familiar organism centred view to the "gene's-eye" perspective. Instead of regarding organisms as the primary harbingers of genes, advantageous to their survival, Dawkins views organisms as transient constructs for the survival and transmission of genes down perpetual germlines.

Having defined the gene as "any [discrete] portion of chromosomal material that potentially lasts for enough generations to serve as a unit of natural selection," (p28) Dawkins reiterates:

> A gene is not indivisible, but it is seldom divided. It is either definitely present or definitely absent in the body of any given individual. A gene travels intact from grandparent to

> grandchild, passing straight through the intermediate generation without being merged with other genes…
>
> Another aspect of the particulateness of the gene is that it does not grow senile; it is no more likely to die when it is a million years old than when it is only a hundred. It leaps from body to body down the generations, manipulating body after body in its own way and for its own ends, abandoning a succession of mortal bodies before they sink in senility and death.
>
> The genes are the immortals, or rather, they are defined as genetic entities that come close to deserving the title… Moreover, just like the ancient replicators in the primeval soup, copies of a particular gene may be distributed all over the world. The difference is that the modern versions are all neatly packaged inside the bodies of survival machines. (p. 33-35)

Dawkins's idea of survival machines as constructs of, and in service of the genes they embody, if only for a reproductive generation, is not unorthodox. It is simply a novel of way of viewing the same biological information, that has been before us for well over 50 years, from a different perspective - that of the gene: "It is a different way of seeing, not a different theory." (xv. See the introduction to the second edition and the diagram of the Necker by way of analogy.)

Recall that in chapter 2 Dawkins identified the properties of a successful replicator as longevity, fecundity, and copying-fidelity. What universal properties would one expect to find in successful (*i.e.* long-lived) genes? Of those relevant to this book…

> at the gene level, altruism must be bad and selfishness good. This follows inexorably from our definitions of altruism and selfishness. Genes are competing directly with their alleles for survival, since their alleles in the gene pool are rivals for their slot on the chromosomes of future generations. Any gene that behaves in such a way as to increase its own survival chances in the gene pool at the expense of its alleles will, by definition, tautologously, tend to survive. The gene is the basic unit of selfishness. (p. 36)

Recall that from the reciprocal definition on p. 4 that "altruism and selfishness are *behavioural*, not subjective... [their] definition is concerned only with whether the *effect* of an act is to lower or raise the survival prospects of the presumed altruist and the survival prospects of the presumed beneficiary."

Two points that most of Dawkins' critics have either overlooked or never read are: a) that he is arguing at the level of the gene and b) in terms of *behaviour* not subjectivity. Of course genes are neither free nor independent agents however, from one generation to the next they must collaborate to influence embryonic development and interact with their external and internal environment. Never the less:

> As far as a gene is concerned, its alleles are its deadly rivals, but other genes are just a part of its environment, comparable to temperature, food, predators, or companions. The effect of the gene depends on its environment, and this includes other genes. Sometimes a gene has one effect in the presence of a particular other gene, and a completely different effect in the presence of another set of companion genes. The whole set of genes in a body constitutes a kind of genetic climate or background, modifying and influencing the effects of any particular gene. [Thus,] A gene that cooperates well with most of the other genes that it

is likely to meet in successive bodies, i.e. the genes in the whole of the rest of the gene pool, will tend to have an advantage. (p. 37, 39)

Here we have a paradox; or rather what Dawkins calls "a subtle, complicated idea". Genes must be both ruthlessly selfish in competing with their alleles and yet collaborate and interact with other genes in their internal environment, and with their external environment. Chapter 5 "Aggression: stability and the selfish machine" deals with human 'game theory' by way of analogy. If you have the book before you please do read it. Chapter 12 however, deals with one analogous game in particular that is both highly accessible and instructive and which serves as the classic text for the remainder of this study unit. If you haven't downloaded it yet you can do so here. There is also a BBC documentary: "Nice Guys Finish First" by Richard Dawkins available on YouTube; however we will be following the text version.

**Task**

1. What is an evolutionarily stable strategy or ESS?
2. What is the meaning of reciprocal altruism and how does it differ from altruism period? In practical terms does it really matter?
3. What is the "dilemma" in Prisoner's Dilemma?
4. Why is iterated Prisoner's Dilemma such a game changer?
5. What does Axelrod mean by his various adjectives: nice, nasty, forgiving, robust and envious?
6. What is a zero sum game? Are there any ethical lessons to be learned in identifying "nonzero sum" situations in real life?
7. What does Dawkins mean by the "shadow of the future" and how can it be a game changer?
8. How does Prisoner's Dilemma throw light on the paradox of the gene referred to above; or what Dawkins calls "a subtle, complicated idea?" (p. 39)
9. If there is one central idea that philosophers should take cognisance of form "The Selfish Gene," what is it?

**Feedback**

1. Informally, an evolutionarily stable strategy (ESS) is one that does well in a population dominated by the same strategy. (Dawkins, 1982 p. 286)
2. According to the *Oxford English Dictionary* altruism is defined as "regard for others as a principle of action." Biologists however use the word to mean the effect of promoting the welfare of another, at the expense of one's own. "In evolutionary biology, reciprocal altruism is a behaviour whereby an organism acts in a manner that temporarily reduces its fitness while increasing another organism's fitness, with the expectation that the other organism will act in a similar manner at a later time." (Wikipedia : Reciprocal altruism) In practical terms the difference does not matter. In a human analogue of the vampire bats in Dawkins' chapter 12, if you have a very rare blood type and are lying bleeding to death by the roadside and can only be saved by a transfusion of blood from someone else with a similar blood type, whether the altruistic act of someone saving you is reciprocal or otherwise is irrelevant to your survival.

3. A dilemma is a situation in which a difficult choice has to be made between equally undesirable alternatives. In Prisoner's Dilemma, two rational players will have worked out, in advance, that no matter what the other player does, he or she will have to defect, knowing full well that if only both had cooperated they would have obtained the relatively higher reward (or reduced punishment.)
4. Iterated Prisoner's Dilemma is just the same game played over and over again an indefinite number of times by the same players. While the only rational strategy to adopt in a single game of Prisoner's Dilemma is to defect, lest you cooperate and your opponent defects, leaving you with the sucker's payoff, an iterated game allows you to assess your opponent. If she is "trustworthy" you can both make a tidy profit by cooperating for several rounds until one of you succumbs to the "temptation" to defect and receive an even higher pay off, if for only one round. Such defections will likely lead to rounds of mutual recrimination, when both parties defect and therefore receive the much lower pay off. The point is that both of you are quite likely to develop some kind of *strategy* by which to maximise profit (or reduce punishment.)
5. Axelrod is using decidedly emotive terms to describe the *behavioural output* of computer programmes designed to simulate various strategies. This is an essentially blind, unconscious, mechanical process so that any talk of human kindness, nastiness, forgiveness *etc.* is out of place. Rather a "kind" strategy for Axelrod is one that is never the first to defect: Tit-for Tat is one such example. Conversely, a "nasty" strategy is one that does defect first: Naïve Prober and Remorseful Prober are two such strategies. A "forgiving" strategy is one that will retaliate against defection but will continue to cooperate if the other side does. Tit for Tat would therefore be classed as both "kind" and "forgiving." A "robust" strategy, according to Axelrod, is one that does well against a wide variety of other strategies. As it happened Tit for Tat did well against the strategies that it came up against, though given a different set of strategies, it might not have proved so robust. Finally an "envious" strategy is one that strives for a greater payoff than the other player, rather than for just an absolutely large quantity. A "non-envious" strategy, in other words, world accept an equal pay off as the other player, so long as each wins more from the banker.
6. A zero sum game is one in which one player's total gains (or losses) is exactly matched by the other player's losses (or gains.) *I.e.* a gain for me is a loss for you and *vice-versa*. Mathematically, if you added up all the gains and subtract all the losses they would sum to zero, hence the name. Some situations in real life do correspond to zero sum games such as one person taking a larger slice of cake, leaving less for everybody else. Quite often however, perceived conflicts of interest are not so: more for me need not necessarily mean less for you. Prisoner's Dilemma is one such example, where if we cooperate, we can both achieve a "win-win" outcome. Adversarial systems however discourage cooperation and foster unnecessary conflict. In the legal arena, conciliation and mediation can often result in an outcome favourable to both parties, at very little cost, compared to protracted adversarial litigation. On a more mundane level, if you perceive others adversaries rather than potential allies, or you are envious and strive to always do better than others, you are likely to regularly miss the numerous chances of deriving a mutual benefit.
7. The shadow of the future refers to the (perceived) length of dealings. If there are unlikely to be no, or very few, further rounds then, the rational strategy is to defect, lest you be left with the sucker's pay off. If however there is no end in sight, both players would benefit

from a protracted stretch of mutual cooperation. Dawkins thirty year old description of a harmless symbiont that turns nasty when the host is severely weakened, such as by septicaemia, has proven prophetic. Today tens of thousands of people die every year from opportunistic infections from highly virulent strains of otherwise symbiotic organisms when their bodies are so weakened by a combination of HIV / AIDS, TB and malnutrition. Normally, form the point of view of an endosymbiont, killing your host is suicidal, however if you sense that your host may be about to die anyway, it pays to defect and make as many copies of yourself as you can and get them out there to infect another host.

8. Unless one is prepared always to receive the sucker's pay off, Prisoner's Dilemma is a game that must be played motivated by self-interest. However, with very few rules and only two possible moves for each player in every round (cooperate or defect) ostensibly "nice" strategies can be favoured over "nasty" ones in repeated simulations. Of course such strategies are no more nice or altruistic than the computers on which they run, however Dawkins point is an analogous one: Just as "nice," "forgiving," "robust" strategies can evolve from a set of simple rules and self-interested input, so "selfish genes" can evolve "cooperative" behaviour with other genes in the gene pool in the vastly more complex "game" of life.
9. Yes. Philosophers should note that it is possible, if not inevitable, for altruistic behaviour to arise from an entirely blind process of natural selection, (without recourse to abstruse entities or immaterial forces.)

**Appendix A**

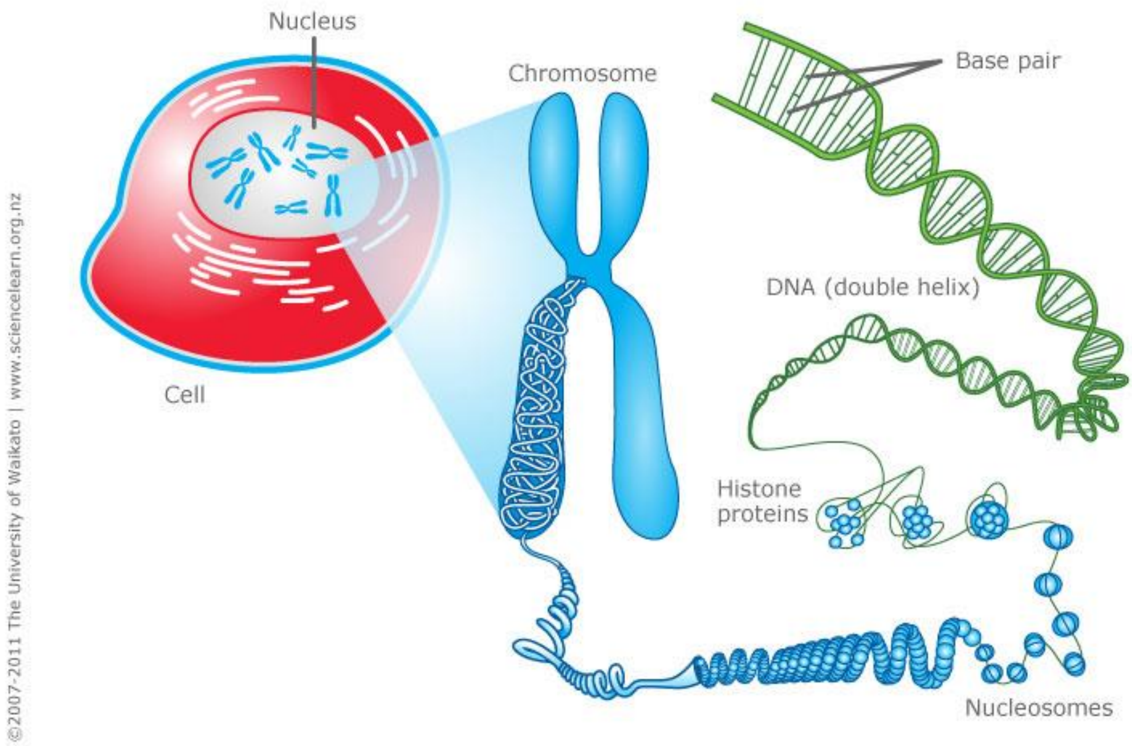



*Figure 1: The Structure of a Chromosome*

An **organism** is a living thing, whose growth and reproduction is orchestrated by a myriad of cellular processes. The molecules involved are typically the product of the activity of **enzymes**, which are a class of proteins that catalyse biochemical reactions. Many of the building blocks of cells are themselves proteins. Proteins in turn are made up of a sequence fifty or more **amino acids,** of which there are twenty that we find in nature. Shorter sequences of amino acids are known as **polypeptides**. It is the unique sequence of amino acids that comprise of a protein or polypeptide that differentiates one kind of protein or polypeptide from another.

The instructions for the manufacture of every type of protein or polypeptide that an organism can potentially produce are encoded in its **genetic material** (either **DNA** or **RNA** in the case of some viruses.) Not all of this genetic material actually encodes for proteins: some of it consists of instructions to either start or stop protein synthesis, while the function much else besides remains a mystery and has been dubbed "**junk-DNA**." Most organisms carry their DNA on one or more structures known as **chromosomes** and specific genetic information is held at specific locations or **loci** (sing. **locus**) on a chromosome.

**Genes** are the units of heredity and natural selection, which are transferred from parent to offspring. They consist of a certain sequence or sequences of DNA that determine the order of the amino acids that go to make up a protein or polypeptide. Genes therefore determine those characteristics of an offspring that are inherited from one's biological parents. Such characteristics are known as **traits** which may be visible, such as eye colour, or may only be detectable by a blood test, such as blood type: A, B or O. Biologists refer to an organism's full genetic makeup as the **genotype** *i.e.* all the hereditary information held within the genetic material. **Phenotype** on the other hand refers to all of organism's potentially observable properties. Not all of the heredity information within the genotype is necessarily expressed in the phenotype as traits; someone, may for example, have inherited a gene for O blood type, yet exhibit another blood type because another gene or genes .

Alternative forms of the same gene are known as **alleles** and typically occupy the same genetic locus. Generally, different alleles result in different observable phenotypic traits, such as blood group or eye colour. In the case of humans and most other multicellular organisms, we have two sets of chromosomes and so are termed **diploid**; organisms that have only one set are known as **haploid**. **Sex cells**, *i.e.* sperm and ova, also have only one set of chromosomes each and so are also said to be haploid. Ordinary, chromosomes other than sex chromosomes are known as **autosomes**. Because we inherit one set of chromosomes from each parent we will have one copy of each gene (and, therefore, one allele) from each parent on each chromosome. If the alleles are the same, then we and the alleles are said to be **homozygotic** (or **dizygotic**) and if they are different, **heterozygotic** (or **monozygotic**.) Of course we may be homozygotic for one set of alleles and heterozygotic for another. If we are homozygotic for one pair of alleles, then we inherited the same copy from each parent and if we are heterozygotic for the same, then we inherited a different copy from each parent.

Very often one allele is expressed over that of another allele at the same gene locus, in which case the first allele is said to be **dominant** and the second one **recessive**. The blood types A and B, for example, are dominant over O. Consider a gene that exists in two allelic forms R and r such that R is dominant and r is recessive, then there are three possible genotypes that we might inherit from our biological parents: Homozygotic (dizygotic)  RR, heterozygotic (monozygotic) Rr or homozygotic (dizygotic) rr. In the case that R stands for the *Rhesus* blood factor Rh+ and r for Rh-, having either a genotype RR or Rr will result in phenotype Rh+ because R is always dominant over little r, while being homozygotic for little r *i.e.* rr will result in the phenotype Rh- because only little r alleles will be present at a particular gene locus. Alleles that are neither dominant nor recessive with respect to one another are known as **codominant**. Note that dominance is a relative relationship between alleles *e.g.* among multiple alleles one may be dominant over another, recessive with respect to a third or codominant with respect to a fourth.

The duplication of genetic material occurs during simple cell division or **mitosis**. DNA condenses onto chromosomes which are separated into identical pairs. The contents of the parent cell are then parcelled up into two identical daughter cells, each with a single set of identical chromosomes. Mitosis is the ordinary process by which bodies grow. **Meiosis** on the other hand is the process by which usually diploid cells give rise to haploid **gametes** or sex cells (sperm and ova.) Often referred to as "reduction division," meiosis necessarily halves the number of chromosomes before gametes fuse to form a **zygote** with the original number of chromosomes. During meiosis **homologous chromosomes** (one each from mother and father) pair up and exchange genetic material in a process known as **crossing over**. This ensures that each daughter haploid cell contains a unique combination of genes from each parent, so that each and every sperm and ovum is unique. In creatures that reproduce sexually, crossing over is the mechanism that ensures that progeny vary from their parents genetically, allowing for adaptation to environmental changes. Species that reproduce asexually, because they are clones, are much more prone to population crashes when their environment changes suddenly and adversely. Occasionally a **point mutation** (or single base substitution) results in the replacement, insertion or deletion of a single "letter" of genetic code with that of another "letter" of the genetic alphabet. If the mutation occurs in the **germ line** (cells that contain genetic material that is passed to progeny) the alteration is usually **deleterious** (bad,) neutral but once in a while adaptive.

Dawkins' (1982) book, *The Extended Phenotype* contains an excellent glossary written for the non-specialist. Many terms related to genetics that are not explained above can be found there.

**References:**

DAWKINS, R. (1982) *The Extended Phenotype.* Oxford University Press : London

DAWKINS, R. (2006) *The Selfish Gene.* 30th anniversary edition. Oxford University Press : New York

The next Classic Text concerns physicalism as it relates to the Philosophy of Mind.