ALTRUISM: ANALYSIS OF A PARADOX

ALTRUISM: ANALYSIS OF A PARADOX

By
YUSSIF YAKUBU, B. Sc. (Hon), Dip. Ed., M. Sc.

A Thesis Submitted to the School of Graduate Studies in Partial Fulfilment of the
Requirements for the Degree Master of Arts

McMaster University MASTER OF ARTS (2012) Hamilton, Ontario (Philosophy)

TITLE: Altruism: Analysis Of A Paradox

AUTHOR: Yussif Yakubu, B.Sc. (Hon), Dip. Ed. (Cape Coast), M. Sc. (Guelph)

SUPERVISOR: Dr. Richard T. W. Arthur

NUMBER OF PAGES: vii, 136

Abstract

Theories that engender fundamental transformations in our world view seldom come perfect from the outset for two reasons. First, the empirical discoveries and theoretical framework necessary for their full explanatory efficacy are often not yet in place. Secondly, as a consequence of the first, some of the auxiliary theories and assumptions they rely upon are often antiquated and erroneous. For these reasons, anomalies are frequent in scientific theories. In this thesis, I discuss some of the major scientific anomalies, including particularly, the paradox of altruism. I suggest that the paradox of altruism arises because one of the most fundamental Mendelian genetic principles is misapplied. I show that today's explanatory models err in supposing altruism and selfishness to be genetic allelomorphs. The supposition is inconsistent with the field data on altruism, and entails a logical inconsistency in accounting for the evolution of altruism. Largely, the models that purport to resolve the paradox hinge on the conditional expression of the altruistic gene, a move which I argue contradicts the theoretical assumption that engenders the paradox in the first place. I demonstrate from the empirical data that altruism and selfishness are rather plastic phenotypic expressions of a single genotype. And by supplanting the standard neo-Darwinian assumptions with the principle of phenotypic plasticity, I provide a parsimonious account of the evolution and maintenance of altruism which entails no paradox.

Acknowledgements

Table of Contents

List of Illustrations

List of Tables

List of Abbreviations

ASA - Altruism Selfishness Allelomorphism
ASP - Altruism Selfishness Plasticity

General Introduction

Darwin's theory of Natural Selection is arguably the best known scientific theory today. It is invoked to explain an immense array of problems not only in biology, but also in diverse academic disciplines such as anthropology, sociology, psychology and economics. Notwithstanding its explanatory power the theory still suffers from some lingering doubts engendered by the perennial problem of altruism.[1] One of the logical consequences of natural selection is that traits that put their bearers at a fitness disadvantage should perish from populations. On the contrary, such maladaptive traits thrive in populations. Chief among them is altruism, which Darwin (1859; 1871) himself acknowledged to challenge his theory.

---

[1] The word altruism as used in evolutionary biology is slightly different from its ordinary general usage, just as in physics where terms such as power, force, energy etc. are very precisely defined and different from the layperson's use of them. The word altruism in evolutionary biology is used to describe any situation in which a trait (often behavioural) represents a cost to its bearer but has a beneficial effect on some other individual(s). It does not have to be conscious or deliberate. Biologists also coined terms such as symbiosis/mutualism and parasitism to describe other relations between organisms. In symbiosis the association between two organisms is mutually beneficial. *Lichen* for example is a symbiotic association between a fungus and an alga. The fungus provides nutrients for the alga and the alga provides photosynthates for the fungus. In parasitism, one organism just sneaks in and steals from another. The resource the parasite steals is not seen as designed for the benefit of the parasite. So these two latter cases are seen to be different from altruism. Some people outside evolutionary biology have tended to see altruism simply as mutualism, but that is not necessarily the case. The closest thing in the discussion of altruism to mutualism is reciprocal altruism. But evolutionary biologists prefer to keep the two separate.

To date there is no single theory that we can point to as a complete and satisfactory solution to the problem. Instead, there is a patchwork of explanatory models, each of which explains some form of altruism on the basis of assumptions that are utterly inconsistent with other forms of altruism. Apart from this limitation in range, the models are also limited in depth, in that they predicate upon an already existing population of altruists, and offer no persuasive mechanism as to how a single altruistic mutant gene could possibly spread in a population to the threshold frequency of altruists at which such models can take hold. We should not be content with this state of affairs.

In this thesis I show through extensive empirical data that the genetic model used in extant explanations of altruism is not consistent with the altruistic behaviour as observed in field studies. And that is the source of the difficulty those models have in giving a complete account for altruism. Under extant modeling of altruism, the altruistic and selfish traits are assumed to be genetic allelomorphs. On the contrary, observations in fact suggest that in almost all cases, the same individual may express the phenotypic condition we call "altruism" or the phenotypic condition we call "selfishness" contingent upon cues to which it is exposed in the (social) environment. That indicates unmistakably that altruism and selfishness are plastic expressions of a single genotype.

Consequently, this thesis puts forward *phenotypic plasticity*[1] as the genetic model for explaining altruism. The model has the advantage of parsimony because it precludes the existence of the putative exclusive altruistic genotype of lower evolutionary fitness, whose "imagined" prevalence gives rise to the paradox. The new model also provides a mechanism by which a single mutant that has the capacity to express altruism can spread in a population, another explanation that has eluded all existing theories of altruism. In fact, this thesis rejects altogether the existence of any allele or genotype dedicated to the expression of altruism. Instead, the thesis postulates a pleiotropic plastic "sociality trait" that can express either altruism or selfishness depending upon the circumstances. The model presented here provides the general principle for explaining the persistence in any population of any behaviour that lowers the reproductive fitness of its bearers. This will include psychiatric disorders, homosexuality and psychopathic.

The thesis is organized into three articles designated as chapters one, two and three. In chapter one I render a contextual analysis of the problem of altruism, in which I identify similar theoretical anomalies in the history of science, and how they are often resolved. I assert in that chapter that such anomalies are often engendered by one or more erroneous auxiliary assumptions or hypotheses, which are often so axiomatic that there is often not even a thought to question them. I support this assertion through a detailed review of the nascence

---

[1] Phenotypic plasticity is defined on pages 60 and 65

and evolution of the Copernican model of planetary motion, and how received assumptions that were taken for granted were ultimately the errors that held back the theory. I then put the Darwinian problem of altruism under that analytical lens and revealed a similar pattern of erroneous auxiliary assumptions in its history, including that which renders altruism a paradox today.

In chapter two I discuss in detail the assumption of altruism and selfishness as genetic allelomorphs in extant models that seek to explain altruism. I demonstrate through extensive empirical data why that is wrong, and point out that the empirical data suggest rather that altruism and selfishness are dimorphic phenotypes of a single genotype. In chapter three I use resources from behavioural and psychiatric genetics to show how maladaptive complex traits (which include altruism) may be maintained in populations through *polygenic plasticity* (i.e. traits that are influenced by multiple gene loci and have a plastic phenotypic expression). In the second part of chapter three, I demonstrate how sociality, rather than just altruism, may evolve under the phenotypic plasticity model.

Chapter One

THE EQUANTS AND EPICYCLES OF MODERN DARWINISM
A Contextual Analysis of the Problem of Altruism

I. Introduction

In the *Origin of Species*, Darwin (1859, 352) considered some difficulties of his theory of natural selection. Among them was "one special difficulty," he said, "which at first appeared to me insuperable, and actually fatal to my whole theory." He continues: "I allude to the neuters or sterile females in insect-communities." Darwin suggested a twist to natural selection, which he believed resolved that potentially "fatal" problem; and it was that "selection may be applied to the family as well as to the individual." He explains that if a community has some members who are sterile, but have "slight modifications of structure and instinct" that are advantageous to the community, that community would thrive. Problem solved? Well, unfortunately not. A century and a half after Darwin, not only is that potentially "fatal" little problem still alive and strong, but now has armies of scholars working to resolve it. In its grand manifestation today as altruism, it encompasses more than sterile insects.

Darwin tackles the problem of altruism in mammals, particularly humans, separately in a later work (The Descent of Man, 1871), in which he suggests the same group benefit explanation for the spread of what he refers to as "social and moral qualities" (1871, 162-3). Darwin easily explains how groups with more self-

5

sacrificing individuals would edge other groups and spread. He notes however, that within any group, such self-sacrificing individuals are not any fitter than the selfish ones, and admits therefore that it is "scarcely possible" that within any group, the altruistic virtue could evolve through natural selection (1871, 163). He only suggests what he calls "probable steps" of how altruistic individuals night increase in a tribe. Again, that was not quite satisfactory, as there has been a proliferation of theories after Darwin that are trying to resolve the same problem.

In fact, to date no one can point to any explanation of altruism that is completely satisfactory. Instead, there is a patchwork of explanatory models, none of which is universal in potency. Such models include kin selection (Hamilton, 1964; Maynard Smith, 1964), group selection (Wynne-Edwards 1962, 1986; Wade 1977), reciprocal altruism (Trivers 1971), cooperation (Axelrod and Hamilton 1981), etc. Each of the models explains some forms of altruism but not others. Apart from this limitation in range, the models are also limited in depth, in that they predicate upon an already existing population of altruists, and offer no mechanism as to how a single altruistic mutant could possibly spread in a population to the threshold frequency of altruists at which such models can take hold; what Gildenhuys (2003, 28) refers to as "the familiar objection,". We cannot be content with this state of affairs. In this discussion, I offer a diagnosis of this Darwinian theoretical malaise by identifying parallels in the history of science, which reveal the general nature of such problems and how they are often

resolved. My thesis is that anomalies in major theories are a consequence of erroneous auxiliary assumptions; and often the longest enduring difficulties are those resulting from the error of our most trusted assumptions. It is just as in crime stories, where the most difficult crime mysteries to crack are those in which a very unlikely person is the culprit, for example, somebody within law enforcement.

The situation (as it is now with altruism) in which different scientific theories explain different instances of the same phenomenon, or some parts of the data and not others, suggests that *ad hoc hypothesizing* may be at play here. Any model which has only a parochial explanatory efficacy of a problem is likely based on superficial considerations and not anchored to any deep underlying principles. A re-examination of our basic assumptions about altruism is thus warranted given that a complete and satisfactory theory remains elusive in spite of the protracted effort the subject has received.

II. Contextual Analogy

Theories that engender fundamental transformations in our world view seldom come perfect from the outset for two reasons. First, the empirical discoveries and theoretical framework necessary for their full explanatory efficacy are often not yet in place. Secondly, as a consequence of the first, some of the auxiliary theories and assumptions they rely on are often antiquated and erroneous. The valid scientific theories usually improve later with the elimination of some of the

carried over assumptions from their antecedent world view. Science aims at enabling understanding of our world. Its progress is measured by how better we understand and explain our world today relative to times past. That understanding is in turn measured by how effectively our theories, beliefs and assumptions about the world explain our observations and predict phenomena. In effect therefore, scientific progress is made when we alter these theories, beliefs and assumptions in a way that expands our understanding of our world and our capacity to explain observations and predict phenomena in that world. What I have just outlined actually is what scientific theories do for the human being. It is much more difficult however, to describe precisely what a scientific theory is and when a scientific theory is a good one or a bad one.

A good starting point for this is Karl Popper's conception of a scientific theory, because it is the most stringent and if followed to the letter, would disqualify many respectable theories today. That is why there has been no shortage of criticism of Popper, and a variety of modifications particularly to his falsificationist thesis have been proposed. So we start from Popper at the extreme and then work our way back to a reasonably realistic position at which today's major scientific theories can be accommodated. After outlining what he says are the features of a scientific theory (1963, 36-7), Popper (1963, 37) writes: "One can sum up all this by saying that *the criterion of the scientific status of a theory is its falsifiability, or refutability, or testability*." The suggestion is that if a

8

scientific theory is thus falsified either through direct testing or by observation, it ought to be rejected.

Popper's "conjectures and refutations" conception of scientific hypotheses was not original. Before him was C. S. Peirce (1931), and then William Whewell (1847) earlier on. Whewell (1847, 56) here describes the obligations of one who propounds a scientific hypothesis thus: "it is indispensably requisite that he be diligent and careful in comparing his hypotheses with the facts and ready to abandon his invention as soon as it appears that it does not agree with the course of actual occurrences." Having said that, Whewell (1847, 60) nevertheless recognizes that "hypotheses may often be of service to science, [even] when they involve a certain portion of incompleteness, and even error." He notes (1847, 60) that "the object of such inventions is to bind together facts which without them are loose and detached" (60). Whereas Whewell offers reasons why it may be justifiable to maintain a scientific hypothesis which faces some contrary observations, Popper frowns upon such moves; moves that would have mitigated the devastating impact of falsificationism on scientific theories. He writes (37):

> "Some genuinely testable theories, when found to be false, are still upheld by their admirers—for example by introducing *ad hoc* some auxiliary assumption, or by reinterpreting the theory *ad hoc* in such a way that it escapes refutation. Such a procedure is always possible, but it rescues the theory from refutation only at the price of destroying, or at least lowering, its scientific status. (I later describe such a rescuing operation as a "conventionalist twist" or a "conventionalist stratagem")."

9

Such is one of the reasons why I characterize Popper's criterion for a scientific theory as stringent; because this practice of introducing auxiliary hypotheses to counter anomalies in theories is quite common, and many respectable scientific theories today would not have survived without such measures. Kitcher (1982, 45) notes for example, that Newton's celestial mechanics appeared "falsified" by strict Popperian standards with the observed behaviour of Uranus when it was first discovered in 1781. However, in a typical infraction of the Popperian tenets, Newton's theory was retained whilst *ad hoc* hypotheses were proposed, suggesting the presence of another planet nearby that was disturbing the orbit of Uranus. In another "conventionalist stratagem," as Popper would call it, Einstein postulated the "cosmological constant" to keep the universe static when his field equations for general relativity suggested an expanding universe at a time that physicists believed the universe was static. Also, anomalies in the extant theoretical paradigm of astrophysics have led to the postulation of *dark matter* and *dark energy* as *ad hoc* measures to keep the core theoretical principles buoyant. *Dark matter* and *dark energy* as they stand now are today's equivalents of the epicycles and equants of Ptolemaic astronomy. They may be confirmed or may be found to be unnecessary when the real source of the anomalies is discovered. This was the case with Einstein's cosmological constant, which he dropped after it was revealed that the assumption of a static universe upon which

10

he relied was wrong. On the other hand, the postulation of another planet (Neptune) on account of the anomalies in Uranus' orbit was confirmed.

There are other ways that the stringency of Popper's theory becomes its bane. For example, Popper says above that when a scientific theory is "found to be false," it would be unscientific for its admirers to hold on to it through "conventionalist twists" rather than abandoning it. But finding a major scientific theory to be false is not that simple. In fact, that is one of the popular criticisms of Popper's falsificationist doctrine. It could be the theory itself or any of the auxiliary assumptions that is responsible for an incorrect prediction (Chalmers, 1999, 89), for scientific theories are never isolated explanatory tools. They usually dwell in networks with auxiliary beliefs and assumptions. Kuhn (1962) suggested there is usually a particular dominant world view that defines the framework under which such practices and auxiliary theories operate; and he called that a *paradigm*. Kuhn cites examples of such paradigms to include: "Ptolemaic astronomy, Copernican astronomy, Aristotelian dynamics, and Newtonian dynamics" etc. (1962, 10). None of these can easily be falsified by a single observation. According to Kuhn, a practitioner of a particular paradigm "will seldom evoke overt disagreement over fundamentals" (1962, 11). This suggests a set of core assumptions and practices that are to be followed and not revised by adherents of the paradigm. Imre Lakatos (1978, 4) describes such frameworks of scientific hypotheses and assumptions (Akin to Kuhn's "paradigm") as "Scientific Research

Programs," each of which consists of a "hard core" of an immutable set of beliefs and assumptions surrounded by a "protective belt" of auxiliary theories and assumptions.

Lakatos suggested that blame for the difficulties of a "research program" is often charged to the peripheral assumptions that constitute the "protective belt" rather than the "hard core." However, a "research program" usually has multiple peripheral assumptions, and the anomalies will persist if we are unable to identify the particular peripheral assumption responsible. It is in such situations, in which a very successful theory is blemished by one or more anomalies that adherents resort to *ad hoc* hypotheses, a practice Popper rejects as unscientific. What we come away with from Kuhn and Lakatos however, is that due to the protective belt of peripheral hypotheses and assumptions, scientific paradigms or research programs hardly ever fail cataclysmically. The core beliefs of a paradigm fade away gradually whereas there is a higher turnover rate for the peripheral hypotheses and assumptions, which are more responsive to observations and tests. This is illustrated in the Ptolemaic-Copernican transition discussed below.

III. The Copernican Revolution

Though popularly referred to as a revolution, the so-called Copernican revolution was actually not so revolutionary, nor was it intended by Copernicus to be when he proposed it. Copernicus was an accidental rebel. In the first place the most radical element of his planetary model was the idea that the earth moved, rather

than the immobility of the sun. However, Copernicus, in his *Revolutions* (1543), cites the writings of Plutarch and Cicero in which references were made to earlier planetary models in which the earth moved. Secondly, the model did not become fully convincing until a century after Copernicus proposed it. In fact, in proposing his heliocentric hypothesis, Copernicus did not see himself as toppling the dominant scientific paradigm of the day. He had such an abiding devotion to tradition and orthodoxy that he would have viewed a move like that almost as heresy. Consider the diatribe he visited upon the Nuremberg mathematician Johannes Werner, who dared to question the reliability of certain observations by Ptolemy and Timocharis. Copernicus writes:

> "It is fitting for us to follow the methods of the ancients strictly and to hold fast to their methods which have been handed down to us like a testament. And to him who thinks that they are not to be entirely trusted in this respect, the gates of our science are certainly closed. He will lie before that gate and spin the dreams of the deranged about the motion of the eighth sphere; and he will get what he deserved for believing that he can lend support to his own hallucinations by slandering the ancients." (Koestler.1963, 200)

Clearly, this does not sound like someone who will subvert a *scientific research program*. In fairness to Copernicus he did present in the same document, an extensive and detailed argument why Werner's claims could not be right. However, to go beyond that and condemn Werner's attempts at innovation and his audacity to entertain any ideas that are contrary to what the ancients handed down reveals how deeply traditionalist Copernicus was with regards to scientific knowledge. He was therefore not the kind to challenge existing scientific dogma.

13

So when Copernicus proposed his alternative arrangement of the heavenly spheres, it was with the view to preserving the orthodoxy of the day by eliminating an anomaly that undermined it. He was operating within the "protective belt" of the central scientific doctrine of the day, which was Aristotle's *principle of absolute motion*. Copernicus was troubled by the fact that the celestial spheres as arranged by Ptolemy and others did not give the planets a consistently uniform motion as the paradigmatic *principle of absolute motion* required. He writes in the *Commentariolus* (57):

> "Our ancestors assumed, I observe, a large number of celestial spheres for this reason especially, to explain the apparent motion of the planets by the principle of regularity. For they thought it altogether absurd that a heavenly body, which is a perfect sphere, should not always move uniformly. They saw that by connecting and combining regular motions in various ways they could make any body appear to move to any position."

Copernicus observes that the existing planetary theories of the time, including Ptolemy's, were not consistent with the principle of uniform motion. His concern as he puts it was that "these theories were not adequate unless certain equants were also conceived; it then appeared that a planet moved with a uniform velocity neither on its deferent nor about the centre of its epicycle" (57). An equant is a position inside a planet's orbit from which the planet would appear to have a uniform motion. Such equants, in fact as well as the epicycles, are what Karl Popper would call "*ad hoc* hypotheses," because they were arbitrarily invoked solely for the purpose of propping up a theory that was faltering. Copernicus was

14

troubled by this, saying that "a system of this sort seemed neither sufficiently absolute nor sufficiently pleasing to the mind" (57). Upon this, he tells us why he did what he did (Commentariolus, 57-58):

> "Having become aware of these defects, I often considered whether there could perhaps be found a more reasonable arrangement of circles, from which every apparent inequality would be derived and in which everything would move uniformly about its proper center, as the rule of absolute motion requires."

Even though Ptolemy's system may have been dominant, it does appear that there were several other theories with alternative arrangements of the celestial spheres, in fact, including some that involved a rotating and revolving earth, some of which Copernicus alluded to. So the arrangement of the planets at that time was a domain of theoretical flux within the paradigm of the principle of absolute motion. So Copernicus cites these examples and goes on to plead in his letter to Pope Paul III saying (1543, 4):

> "I knew that others before me had been granted the freedom to imagine any circles whatever for the purpose of explaining the heavenly phenomena. Hence I thought that I too would be readily permitted to ascertain whether explanations sounder than those of my predecessors could be found for the revolution of the celestial spheres on the assumption of some motion of the earth."

Thus, if we assume Lakatos' *scientific research programs* model, the arrangement of the planets would be in the protective belt surrounding the hard-core principle of absolute motion.

Copernicus' heliocentric proposal happened to be in the right direction, but it was like a square peg in a round hole, since it was embedded in Aristotelian physics and Ptolemaic astronomy, whose "ancient and medieval" assumptions could not support such a modernistic scientific hypothesis. Consequently, Copernicus' theory itself had to be propped up by ad hoc hypotheses. He introduced epicycles wherever necessary in order to preserve the standing dogma of "the principle of absolute motion." In the end, his system had almost as many epicycles as the Ptolemaic system.

There were other difficulties of the Copernican system, including the objection that a moving earth would lose its atmosphere and any other objects in suspension; and that falling objects should land some distance away. Again, Copernicus had to work with the only resources he had, and he invoked Aristotle's *earth, air, water and fire theory* of elementary matter to explain that falling objects stick with the earth's motion because they consist of the same "earth" substance as earth.

Needless to say, Copernicus' heliocentric arrangements of the celestial spheres didn't quite work out. It was not quite convincing, and remained one of many fringe hypotheses around the dominant Ptolemaic astronomy. This is evident in Galileo's correspondence with Kepler over half a century after Copernicus' proclamation of the heliocentric model. In this particular epistle of August 4th 1597, Galileo tells Kepler he has been working with the Copernican

system which he finds to explain many phenomena that "remain inexplicable according to the more current hypotheses." He continues:

> "I have written many arguments in support of him [Copernicus] and in refutation of the opposite view which, however, so far I have not dared to bring into the public light, frightened by the fate of Copernicus himself, our teacher, who, though he acquired immortal fame with some, is yet to an infinite multitude of others (for such is the number of fools) an object of ridicule and derision." (Koestler, 1963, 356)

Copernicus could not get his heliocentric model to work because there were just too many key assumptions of the day that were erroneous. In his book *The Sleepwalkers*, Arthur Koestler observed that "Copernicus carried orthodoxy regarding circles and spheres even further than Aristotle and Ptolemy" (195). Copernicus thought all the received astronomical data, theories and assumptions were sacrosanct, and all that needed to be done was find the right arrangement of heavenly spheres that would fit the received wisdom. As Koestler (1963, 199) notes, Copernicus' "absolute reliance, not only on the physical dogmata, but on the astronomical observations of the ancients was the main reason for the errors and absurdities of the Copernican system."

Ironically, the central dogma, "the principle of absolute motion," which Copernicus took for granted and worked so hard to preserve, was itself in need of adjustment, if not total rejection, in order to make sense of the astronomical data. That was what Johannes Kepler did in the first two of his three laws. He overturned two key assumptions which Copernicus took as given and sacrosanct,

and which paralyzed his heliocentric model, namely, (1) that the planets moved in circular orbits, and (2) that they moved with uniform velocities.

Kepler too had to come to this realization of the fallibility of fundamental axioms the hard way. He spent five years trying in vain to describe the orbit of Mars using Tycho Brahe's excellent astronomical data. In the end he found himself compelled by the data to discard those received assumptions and theories one after the other. For example, in order to preserve Mars' circular orbit a reciprocating equant was required. Consequently, that whole idea of a circular orbit had to go, and following which Kepler remarked: "The blame for this discrepancy among the different ways of finding the eccentricity … falls entirely upon the faulty assumption studiously entertained by me" (Kepler, 1609, 292). But there were other ways that he was set back for relying on received assumptions that were erroneous. Models he labouriously built would later be found to contradict some other data and would then be torn down. He reportedly made up to nineteen such failed attempts before he got the right description of the motion of Mars (Whewell, 1847, 41-2). Kepler states the nature of the failures and pinpoints their source thus: "Consequently, what we had previously constructed from the Brahe observations we have later in turn destroyed using other observations of his. This was the necessary consequence of our having observed (in imitation of previous theories) several things that were plausible but really false" (Kepler, 1609, 300-1).

The experiences of Copernicus and Kepler are examples of how erroneous auxiliary assumptions can cripple an otherwise good scientific theory. There are many other such examples in the history of science, but in order not to beat up on a dead horse, I will just give one more example here and then proceed to show how that ailment afflicts the Darwinian paradigm today. In the other example, there seemed to be compelling evidence for the theory of continental drift as proposed by the German scientist Alfred Wegener. However, the theory had difficulty gaining wider acceptance because it "seemed to involve a geophysical impossibility" (Parsons, 2006, 37). As Parsons explains, it was hard to imagine the silicates that the continents are composed of, to plough through the denser basalt rock of the ocean floors. This view of the continental crust was however erroneous. The theory of plate tectonics came to the rescue. By providing a mechanism for the movement of the continents, plate tectonics eliminated the main difficulty of the theory of continental drift.

IV. The Problem of Altruism

Darwin needs a similar rescue as in the examples above, from the anomaly of altruism. This problem has beset Darwin's theory of natural selection for over a century and a half now. Unfortunately, the rescue attempts so far have been *ad hoc.* And that is evident in the fact that each of the rescuing auxiliary theories applies to only a very narrow set of the data on the problem and even so, they are still beset with exceptions. The *inclusive fitness* hypothesis for example can

only apply to altruism in which kinship is involved. That is why it was dubbed kin selection by Maynard-Smith (1964). It ceases to be kin selection if the principle is expanded to include any non-kin altruism as Queller (1985) proposed. Group selection explains only the cases of altruism that involve group benefit. Another hypothesis explains altruism that involves two unrelated individuals, and so on. Thus, the theories fit what Karl Popper describes as "a conventionalist twist" by admirers of a theory that is faltering. Auxiliary theories such as kin selection, group selection, reciprocal altruism and cooperation constitute what Chalmers (1999, 141) describes as "contrived" as opposed to "natural" predictions or explanations. These theories were inspired by the anomaly of altruism, and were introduced specifically to make the Mendelian genetic model of natural selection work for altruism. They each act like one of the epicycles and equants that were introduced into the Ptolemaic model of planetary motion in order to make it work.

A hypothesis is *ad hoc* if it is proposed in order to resolve a specific difficulty of another hypothesis. In other words, a hypothesis is *ad hoc* or "contrived" if it is neither motivated nor suggested by any independent data other than the reason that a particular theoretical difficulty would be overcome if we were to adopt that auxiliary assumption. Thus, they are themselves often in need of proof. In other words, they beg the question.

When the theory of plate tectonics was proposed, it was not motivated by the difficulty of the theory of continental drift, neither was it aimed at solving that

problem. It was derived independently, based on a different set of observations. In this way, even though the theory of plate tectonics resolved the geophysical difficulty of the continental drift hypothesis, it was not an *ad hoc* hypothesis. However, if somebody had proposed plate tectonics as a way to deal with the geophysical problem of the theory of continental drift before there was any evidence that the earth's crust is split into giant plates that sit on a molten core, that would have fit the description of an ad hoc hypothesis or a "contrived prediction." In another example, Mendelian genetics provided the necessary mechanism to resolve the problem of inheritance that beset Darwinian natural selection from its inception up to the beginning of the 1930s. But it was neither *ad hoc* nor contrived since it emanated from independent data. If Mendelian particulate genetics were proposed prior to any supporting independent data such as Mendel's experiments, but rather only because it was seen as a theory of inheritance that would fit Darwin's theory of natural selection, then it would have been "contrived." It is in this sense that kin selection, group selection, ESS models, etc. are ad hoc hypotheses akin to equants and epicycles because they were proposed in response to the problem of altruism, and did not emanate from independent observation. On the other hand, the principle of phenotypic plasticity, which I shall propose in this work as the solution to the evolutionary problem of altruism, emanated independently from research in the biological discipline of genetics. In fact it is essentially a description of empirical facts rather

than abstract theoretical speculation. Other than encompassing an expanding array of empirical observations, no one before my current proposal has ever considered it to hold the key to resolving the problem of the evolution and maintenance of altruism and other persistent maladaptive traits in populations. So it is by no means *ad hoc* or *contrived*.

V. The Evolution of Darwinian Explanation

Evolution is understood technically today to be a change in gene frequencies of a population. Today's biology recognizes four ways in which evolutionary change may occur. They are: mutation, natural selection, random genetic drift, and migration. These modes are supposed to be the possible causes of any such changes in gene frequencies; whilst population genetics provides the tools for measuring such gene frequency changes. Yet when Darwin annunciated his theory of natural selection in 1859, the scientific discipline of genetics was yet unborn, even though natural selection relied on what Darwin (1859, 168) called "the strong principle of inheritance" for the preservation and gradual accumulation of favourable variations.

Thus, heritable traits are the mainstay of Darwin's theory of natural selection. In fact, Darwin makes it clear in discussing natural selection that: "Any variation which is not inherited is unimportant for us" (1859, 31). In spite of the cardinal importance of inheritance in his theory of evolution, Darwin lacked a robust theory of heredity. As Pigliucci (2009, 219) notes, "What the original

Darwinism was really missing was not a solid philosophical foundation but rather a theory of heredity." In Darwin's day, the *blending theory of inheritance* was extant, to which he subscribed (Fisher, 1929, 1). The blending theory of inheritance conceived of heritable factors as blendable or miscible, such that in sexual reproduction, the hereditary factor that controls any particular trait, say height, from one parent blends with the corresponding factor for height in the other parent to produce a single new median factor for height in the offspring. Thus, if in sexual reproduction every pair of hereditary factors coming from a pair of parents to an offspring is fused in this manner, the heritable variation would tend to halve with each new generation (See illustration in figure 1). A rapid diminution of heritable variation is therefore logically implied with the blending theory of inheritance. This posed a difficulty of no small measure for Darwin and his theory of natural selection, which required stable variations to work on.

Blending Inheritance   Particulate Inheritance
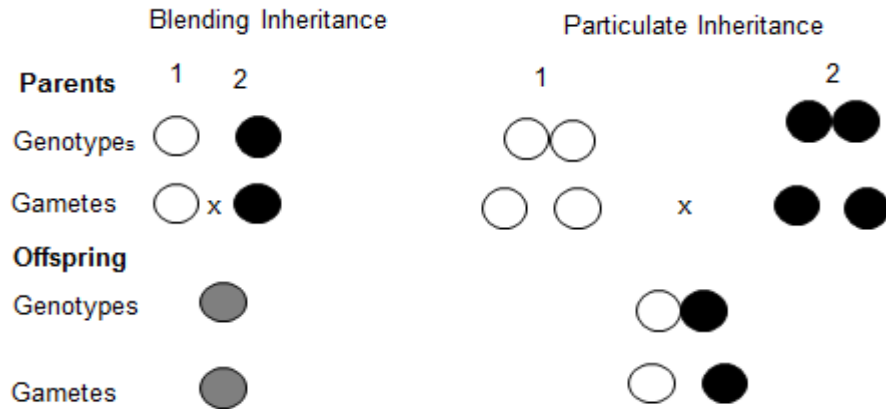
Figure 1. Stability of variations under blending inheritance vs. particulate genetics. Contrasting traits are blended into one under blending inheritance, but remain segregated under particulate genetics.

In fact, this erroneous view of heritable variations was a drag on Darwin's theory of natural selection and undermined its ability to edge out competing theories of evolution. Citing Kellogg (1907), Gould (2002, 506) identifies four major evolutionary explanations that vied for ascendancy in the early 20th century to include Darwinian gradualism, saltationism, Lamarckism and orthogenesis. This theoretical pluralism for evolution would persist till the advent of population genetics in the early 1930s, when the *particulate genetics* of Mendel was substituted for *blending inheritance*. The persistence of the rival evolutionary explanations was the result of the difficulties brought on Darwinism by its reliance on an incorrect auxiliary hypothesis of inheritance.

The difficulties arising from blending inheritance would mount at the turn of the twentieth century as more and more empirical discoveries in biology went

contrary to the predictive consequences of Darwinism. And as Darwinists of the time were unwilling or unable to discard the offending auxiliary assumption, they tended rather to dismiss such contrary observations. In his day, even Darwin himself, on account of blending inheritance, was sceptical of empirical observations that suggested that when domesticated animals are returned to wild living, they tend to regain some of the wild traits that were lost in the process of domestication; a process referred to as *reversion* (Darwin, 1859, 33), which Mendelian segregation later proved to be real and explainable. More such Darwinian views stemming from the blending theory of inheritance would be refuted in the first two decades of the twentieth century by experimental biology. It was the advent of population genetics that would put Darwinism on the path to harmony with the rest of biology.

Population genetics, which is "often regarded as the theoretical cornerstone of modern Darwinism," is a mathematical modeling of "changes in allelic frequencies through the integration of the principles of Mendelian genetics with Darwinian natural selection" (Okasha 2008, 1). This union however was not love at first sight. As Gould (2002, 507) noted, "Darwinians before the synthesis had generally downplayed, ignored, or actively rejected Mendelism." The Darwinians perceived Mendelian "particulate genetics" to be incompatible with the gradualist evolution that Darwinian Theory espoused. Instead they held on to Darwin's blending inheritance as the model that was more compatible with

gradualism and the mathematical modeling (biometry) they had at this time begun to incorporate into their evolutionary explanation. However, whatever progress the Darwinians thought they were making with biometry was pale in comparison with the plethora of empirical confirmations and explanatory successes of Mendelism. Consequently, as Huxley (1942, 24) noted, the Darwinians were "for a considerable time rendered sterile by their refusal to acknowledge the genetic facts discovered by the mendelians". Darwinism was becoming increasingly marginalized by its inability to integrate with the vibrant nascent biological disciplines such as genetics, cytology, comparative physiology etc. During this period a number of spectacular refutations of some of its claims by experimental biology, particularly genetics, had some biologists sounding the death knell of Darwinism (Haldane, 1932, 32). For example, because of the rapid depletion of heritable variations under his blending inheritance, Darwin had to identify ways by which heritable variations are replenished. He suggested changing environments (1859, 124, 174) and natural selection (1859, 156, 169). In the early twentieth century however, experimental biologists demonstrated through pure-line experiments that variations induced by the environment are not heritable, nor does selection *per se* give rise to new variations (Haldane, 1932, 18). During this period, it required devotion beyond the preponderance of the scientific evidence to continue to subscribe to the Darwinian explanation. So as

progress in scientific explanation surged on, Darwinism was shearing off under the drag of the erroneous auxiliary hypothesis of blending inheritance.

Fortunately for all of us today, some lateral thinking Darwinians managed to break free from the ideological shackles and saw reason. Once they acknowledged the errors of their school of thought, they had an easier time convincing members of their camp than did the research data from other biological disciplines. One of the architects of this harmonization, R. A. Fisher (1929) for example, argued that Mendelism could actually be a friend rather than a foe to Darwinian gradualist evolution. In his seminal work on population genetics, Fisher (1929) highlighted the difficulties of blending inheritance and demonstrated how Mendelian particulate inheritance actually resolved those difficulties. He demonstrated for example, that the mutation rates expected under blending inheritance were many thousand fold greater than those suggested by the empirical data, and that the mutations rates required under Mendelian particulate genetics were more realistic. Consequently, he urged the abandonment of the blending theory of inheritance and the adoption of Mendelism as a move that would disencumber the Darwinian explanation of all the untenable and wildly speculative *ad hoc* auxiliary hypotheses. He urged, in other words, a crossing over to the enemy, arguing that: "The whole group of theories which ascribe to hypothetical physiological mechanisms, controlling the occurrence of mutations, the power of directing the course of evolution, must be

set aside, once the blending theory of inheritance is abandoned" (1929, 20). He continues (1929, 21): "The sole surviving theory is that of natural selection, and it would appear impossible to avoid the conclusion that if any evolutionary phenomenon appears to be inexplicable on this theory, it must be accepted at present merely as one of the facts which in the present state of knowledge seems inexplicable." Fisher observes here that all the ad hoc auxiliary "hypothetical physiological mechanisms" that are being put forward to explain new variations (mutation) would not be necessary "once the blending theory of inheritance is abandoned." Similarly, I say that in today's difficulties with altruism, auxiliary hypotheses such as kin selection, group selection etc. would not be necessary once Mendelian allelic segregation of altruism and selfishness is abandoned.

Fisher's work was reinforced by other mathematical biologists such as Sewall Wright and J.B.S. Haldane together with Theodosius Dobzhansky and others. This upheaval within Darwinism, which was dubbed "The Modern Synthesis" by Julian Huxley (1942), was characterized most essentially by the adoption of Mendelian genetics as "the genetical theory of natural selection" as was suggested by Fisher in his 1929 book of the same title. With that, modern population genetics was born, and allowed Darwinian explanation to expand to include other biological disciplines which were thought to be inconsistent with Darwinism.

VI. Mendelism and the Problem of Altruism

Within half a century following the transition from blending inheritance to particulate genetics, modern genetics started to outgrow Mendelism. Today we are at a point again when Darwinism has to adjust in order to conform to advances in biology. Once again it is not an easy transition, and there are ideologues who will oppose any tinkering with Darwinian explanation as it stands today. However, Darwinism once again risks becoming sterile if it fails to incorporate new empirical discoveries in genetics and to adjust to novel genetical explanations.

Population genetics is the primary analytical tool of evolutionary biology. The principles and practice of population genetics are essentially Mendelian. That was what R. A. Fisher achieved for modern Darwinism. Fundamentally population genetics assumes every trait to be controlled by a pair of allelomorphs and provides the mathematical tools for describing their frequency dynamics in a population. The empirical evidence that altruism and selfishness are not Mendelian style contrasting alleles (or allelomorphs) is overwhelming. Yet all the best known theories to date that seek to explain the evolution and maintenance of altruism are based on the supposition that they are genetic allelomorphs. Today we know through ecological and genetic studies of the eusocial insects, that the behavioural and morphological characteristics we associate with altruism and selfishness are actually cases of polyphenism, i.e. differential phenotypic

expression of the same genes, rather than the corresponding products of different genes (Evans and Wheeler 1999). In fact this could be inferred from field studies going back to the 1830s, when it was observed that the same honey bee egg could produce a selfish queen or an altruistic worker depending upon what diet the larva was fed (Prete, 1990).

The fact is that altruism and selfishness as observed in natural populations do not fit a model of two competing genotypes as today's evolutionary models assume. I will refer henceforth to those models as the *altruism selfishness allelomorphism (ASA)* models (see details on p.44). However, as I argue in this work, there is nothing about the altruistic and selfish behaviours that suggest they are controlled by two contrasting Mendelian alleles. The assumption is driven largely by theoretical convenience. Empirical observations of altruism in nature seem to establish generally that the same individual that exhibits the altruistic phenotype would also exhibit the selfish phenotype under some different environmental circumstance. Incidentally, the ASA models call for this very feature (which they call conditional altruism) as a way out of the theoretical quagmires resulting from the ASA assumption. Yet they seem to be oblivious to the fact that you cannot assume two traits to be allelomorphs and then expect an individual to be able alternate between them in a lifetime. Haldane (1955, 38) explains that "two genes are said to be allelomorphic if a nucleus with a single chromosome set, for example that of a spermatozoon, can only contain one of

the two." Even where both the alleles are present in a single diploid organism (heterozygote), only one (the dominant allele) is expressed. In incomplete dominance, there is usually a blended simultaneous expression of the morphological phenotypes. There is never environmentally contingent expression of one or the other in an alternating fashion. Where the latter happens, in modern genetics parlance we describe that as phenotypic plasticity (Pigliucci, 2001)

The assumption of genotypic dichotomy between altruistic and selfish individuals does not fit altruism as observed in honeybee society (Prete, 1990; Winston 1987). Nor can we say it fits altruism as we observe it in eusocial wasps (Vespidae) (O'Donnell, 1998), vampire bats (Wilkinson, 1984), the Belding's ground squirrel (Sherman 1977), olive baboons (*Papio anubis*) (Packer (1977), vervet monkeys (*Cercopithecus aethiops*) (Seyfarth and Cheney 1984), the naked mole rat (*Heterocephalus glaber*) (Lacey and Sherman, 1991), humans (Zimbardo, 1971; Castro et al, 1998). In all these empirical examples, any individual in the population can behave altruistically or selfishly depending on the circumstances. Thus, within each individual in such social populations there is the genetic capacity to express both altruism and selfishness, contingent upon the circumstances of the (largely social) environment. Consequently, as opposed to the *Altruism Selfishness Allelomorphism* (ASA) models, I propose an *Altruism Selfishness Plasticity (ASP)* model, which construes the altruistic and selfish traits as dimorphic phenotypic expressions of a single genotype rather than

31

Mendelian genetic allelomorphs. I contend that altruism remains an evolutionary paradox only because of that genetic mischaracterization by today's models.

Every standard extant statement of the problem of altruism carries the implicit assumption that it is a Mendelian alternative allele to a selfish allele. It is the idea of altruistic phenotypes in competition with selfish phenotypes for the dissemination of their respective genotypes that engenders the paradox of altruism. It is hard to imagine how Altruism could be a paradox if we did not think of it in such Mendelian genetic terms. It is not immediately obvious that the assumption of contrasting altruistic and selfish genotypes is the source of the problem because such an assumption is well in consonance with the principles of Darwinian evolutionary theory. According to Darwin, evolution comes about by natural selection between heritable traits. Therefore, in order to explain the evolution of altruism, biologists have, in the modern synthetic fusion of Darwin and Mendel imagined a heritable altruistic trait (i.e. an altruistic gene) in competition with an alternative gene that is selfish. Darwin (1871, 163) noted however that such an approach is unlikely to work for altruism for reasons he gives in the next section below. Hamilton (1964) also notes this difficulty in the introduction to his article in which he proposes the inclusive fitness hypothesis. However, rather than discard the idea of distinct altruistic and selfish alleles, Hamilton (1964) leaves it in place and instead, redefines fitness in a way that he believes rectifies the anomaly.

Gene expression as demonstrated by modern genetics is often more complicated than the direct genotype to phenotype mapping entailed in the Mendelian model. There are gene-by-gene and gene-by-environment interactions which are indeed key to explaining altruism, but which extant theories of altruism have failed to incorporate. Unlike classical Mendelism, we cannot, in today's understanding of genetics, assume automatically any two distinct phenotypes to be correspondingly expressed by two separate genes. As science writer and reporter Stephen Hall (2010, 67) puts it: "What was once assumed to be a straightforward, one-way, point-to-point relation between genes and traits has now become the "genotype-phenotype problem," where knowing the protein-coding sequence of DNA tells only part of how a trait comes to be." Epigenetic factors loom large in gene expression, especially for complex traits such as behaviour, which of course include altruism.

The problem with explaining altruism today is that there is a growing chasm between experimental biology and theoretical evolutionary biology. Theoreticians have gone off on a tangent since the advent of population genetics when mathematical biologists such as R. A. Fisher, Sewall Wright, J. B. S. Haldane and others successfully established a quantitative account of natural selection based on Mendelian genetics. Since then, the theoretical biologist has led the way with evolutionary models such as inclusive fitness, group selection, reciprocal altruism, cooperation (ESS) etc. which experimental biologists then set

up experiments to confirm. Many of these studies have indeed reported the sought after confirmations, and relatedness for example is being reported to be linked to altruism in some studies. But it remains to be seen how many of these are driven by the psychology of "confirmation bias." For in recent years, indubitable empirical data which starkly contradict the theoretical claims of inclusive fitness has been building up. Some scholars (Nowak et al, 2010) have suggested this has sufficiently undermined inclusive fitness for alternative explanations to kin altruism to be sought.

Defenders of inclusive fitness however are infuriated by the suggestion that it be abandoned. As Michael Marshall of *New Scientist* (September, 2010, 8-9) reports, a rancorous debate ensued during a conference over that issue. Reportedly, Nowak et al's work was condemned as "unscholarly," "misguided" and "lunacy" among other choice words. Of course, Nowak et al's work does not deserve those labels. On the contrary, they presented a very cogent argument in which they cited research after research that refutes each and every one of the central claims of inclusive fitness. From the report, it appears that in opposing Nowak et al's work, the defenders of inclusive fitness appear to concede the refuting research against inclusive fitness, for Marshall writes (p.9): "They say theoretical biologists have always known that inclusive fitness was an approximation, though this seems not to have filtered through to experimental biologists, who have tended to take it as gospel." I am not sure what such

passionate adherents of inclusive fitness mean when they say it is "an approximation not to be taken as gospel." Perhaps it is a euphemism for an admission that it is not quite a satisfactory theory. For one, the empirical evidence that contradict the inclusive fitness hypothesis has been mounting. In fact, its most central claims have repeatedly been contradicted by research discoveries in the last couple of decades (see a compendium in Nowak et al., 2010).

All these difficulties notwithstanding, there are those who still view a call for a theoretical departure from inclusive fitness to be sacrilegious. The critics of Nowak et al reportedly claimed Nowak and his colleagues were "transparently wrong," and if they were able to convincingly prove that error, that would have been the extent to which their criticism was justifiable. However, to go further and collect signatures, as it is reported, for a letter to the journal that published the article of Nowak and his colleagues condemning the article, crosses the line from science to advocacy and an ideologically driven campaign to stifle dissenting views regarding the theory of inclusive fitness. That the adherents of inclusive fitness would resort to such unscientific tactics to oppose Nowak et al suggests that they realized they could not win a scientific argument against Nowak et al.

VII. History of the Altruism Problem

Darwin (1871) suggested that groups with more individuals who are prepared to sacrifice for the group would do better than groups with fewer such individuals and would therefore spread. He realized however, that "it seems scarcely possible," that within a group, the number of such self-sacrificing individuals could increase by natural selection, given as he puts it (1871, 163) that: "It is extremely doubtful whether the children of such [altruistic] individuals would be reared in greater number than the children of selfish and treacherous members of the same tribe." That problem of group selections remains unresolved today.

Then in 1964 W. D. Hamilton proposed the Inclusive Fitness Hypothesis. Hamilton suggests in that hypothesis that if altruists direct altruistic acts towards genetic relatives (who are more likely to carry the altruistic gene),and the fitness gain of the recipient more than compensates for the fitness loss of the altruist, then altruism may evolve by natural selection. It is expressed in the now famous inequality:

$$rb > c$$

Where r is the coefficient of relatedness between the altruist and the recipient; b is the fitness benefit to the recipient; and c is the fitness cost to the altruist.

Parsimony is a frequent casualty of theories that are premised on some assumptions that are false. A false assumption introduces contradictions which the theory has to meander around, thus lengthening its explanatory trajectory.

The paradox of altruism itself came about because Darwin in his era had a very rudimentary understanding of heredity. What he called "the strong principle of inheritance" (1859, 21, 168) was an emphasis of the simplistic notion of "like produces like," which he cites as the driving principle in artificial selection (1859, 31). Thus, Mendelism fused so well with Darwinism because of its supposition that evolved contrasting phenotypes, physical and behavioural, are underwritten by distinct corresponding hereditary factors. Even though this is still assumed in modeling the evolution of traits in today's population genetics, not all contrasting phenotypes fit that model. Altruism and selfishness are examples of contrasting phenotypes that, by every indication, are borne by the same genotype.[1] Behavioural traits in general seem to have this phenotypic flexibility, which we describe technically as *phenotypic plasticity*. It is a post-Mendelian empirical genetic discovery that was not available to Darwin in his day. Without the benefit of the knowledge of phenotypic plasticity, Darwin assumed altruism and selfishness to be controlled by separate hereditary factors, and that presented a difficulty for his theory of natural selection. Recall from the introduction above, that another case of phenotypic plasticity that Darwin encountered was in the

---

[1] The evidence for altruism and selfishness as plastic phenotypes of a single genotype is the field research that shows that individuals that express the altruistic behaviour/phenotype can also express the selfish behaviour/phenotype under other circumstances (citations discussed P.51-55). And that is exactly what the term phenotypic plasticity was coined to describe, even though the examples till now have largely been morphological traits. My work I believe is the first comprehensive demonstration that altruism and selfishness, also fit the phenotypic plasticity description; and that has a huge implication for the paradox of altruism. In fact it eliminates the paradox.

reproduction of neutered offspring with modified morphological features by queen ants.

When such erroneous auxiliary assumptions obstruct the direct explanatory path to the resolution of a problem, parsimony is immediately compromised because the way around the obstacle is always more complicated and uncertain. Thus, the first wrinkle in Darwin's otherwise simple theory was introduced when his direct path to explaining altruism was obstructed by his ignorance of phenotypic plasticity, as he had to introduce the *ad hoc* hypothesis of group selection to deal with the contradictory posture of altruism in the theory of natural selection. Darwin's group selection glue does not seem to be holding altruism firmly enough to the theory of natural selection; as a result we see parsimony recede further into the distance as even more complex *ad hoc* hypotheses are introduced to deal with the problem. In his *inclusive fitness* hypothesis, Hamilton (1964) suggests that altruism could be sustained in a population if its harmful effects on its bearer are more than compensated for by the beneficial effects it confers on related individuals who also carry the altruistic gene. The complication this introduces is that we now have to be concerned with not just what the trait does in the individual involved, but what it does for the relatives of the individual who exhibits the trait. Maynard-Smith consequently dubbed it "kin selection." Hamilton (1964a, 9) here explains one of the features of his hypothesis:

> "If, for example, the advantages conferred by a "classical" gene to its carriers are such that the gene spreads at a certain rate the present result tells us that in exactly similar circumstances another gene which conferred similar advantages to the sibs of the carriers would progress at exactly half this rate."

He then go through a series of mathematical manoeuvres to come up with an equation that determines whether an altruistic gene would increase or decrease in a population. Furthermore, according to Hamilton (1964b, 21), an altruistic individual who is dispensing altruism randomly in a population might just be breaking even, at best. So he introduces a further layer of complication into his hypothesis, which requires a mutation that allows for the altruistic individual to discriminate in favour of close relatives in its altruistic behaviour.

In spite of all the complications, Hamilton's theory, by design can only possibly account for altruism that exists between close relatives, and is especially tailored for eusocial behaviour. There is of course a lot more altruism outside that range, and we have to look up to other *ad hoc* hypotheses for their explanation. Even within the narrow range of kinship and eusociality that inclusive fitness purports to explain, many exceptions are being uncovered (Nowak et al, 2010). Also, nearly half a century after Hamilton's hypothesis, evolutionary biologists are still of the view that "the evolution of eusociality, especially how selection would favour sterility or subfertility of most individuals within a highly social colony, is an unresolved paradox" (Johns et al, 2009, 17452). Wilson (2005, 159) also laments that altruism remains "one of the enduring unsettled issues of evolutionary

biology" Even for a tiny sub category of altruistic behaviour such as human parochialism, Bernhard et al. (2006, 914) conclude from their research that "currently, no single theory seems to be able to explain the entire pattern of parochialism cross treatments, providing an opportunity for developing new theories or modifying existing ones." It is the same conclusion for which Nowak et al above came under so much flak from *inclusive fitness* enthusiasts.

The *inclusive fitness* hypothesis was hailed as a breakthrough when it was first proposed by W. D. Hamilton nearly half a century ago. Today we know that its central hypothesis that genetic relatedness (kinship) is the mitigator of the fitness depressing effect of altruism is not necessarily true. There are probably as many examples of altruism between genetically unrelated individuals as there are between genetic relatives. If that is so, then kinship is neither necessary nor sufficient for the evolution of altruism. In the work that won the 2010 Lakatos Award for an outstanding contribution to the philosophy of science, Peter Godfrey-Smith's (2009), *Darwinian Populations and Natural Selection*, we learn that in the current reformulation of Hamilton's principle, "relatedness in the normal sense of it is now optional" (p. 120) and the beneficiaries of altruistic behaviour need not be relatives for altruism to be maintained in populations (p. 121). If in what is called kin selection, kinship is no longer required for the evolution of altruism, do we still have kin selection? Is it not the exhaustive repudiation of such major theoretical claims of inclusive fitness that scholars such as Bernhard

40

et al. (2006) and Nowak et al (2010) are calling for alternatives to inclusive fitness?

Hamilton (1964b) also cites haplodiploidy as support for his inclusive fitness hypothesis, arguing that the unusually high relatedness it generates among sisters (r = 0.75) is what fosters eusocial organizations in the hymenoptera. However, the high relatedness is true only if the queen mates with only one male in which case all the sisters would be of a single paternity. Some evidence was available even to Hamilton that that is not always the case (Hamilton 1964b, 33). We now know more certainly, that it is the rule rather than the exception that the queen mates multiple times with as many as 20 to 39 different drones during the nuptial flight. Consequently, rather than the theory-boosting 75%, the relatedness between majority of the sisters is more likely around 25%. In fact, there is now research that demonstrates convincingly that honeybee colonies with queens that are inseminated by multiple drones are fitter than colonies with queens that are inseminated by a single drone (Mattila and Seeley, 2007). This is evidence that reveals a net evolutionary force in the direction of genetic diversity rather than higher relatedness in eusocial colonies. This effectively maims the claim of the inclusive fitness hypothesis that eusocial evolution is a consequence of high relatedness due to haplodiploidy.

Besides, eusociality is not limited to the haplodiploid hymenoptera. Johns et al (2009) report that there are at least five phylogenetically distinct diploid

41

animal taxa in which eusociality has evolved. They include termites, naked mole rats, beetles, shrimp and aphids. In their study of dampwood termites, Johns et al (2009) observed that two genetically unrelated adjacent colonies upon contact can merge into one eusocial colony. This led them to declare that (p.17455) "in general, theories emphasizing genetic relatedness to explain eusocial evolution have not been well supported by evidence from primitive termites," and that "above-average relatedness is not a prerequisite" for the evolution of eusociality. No wonder some scholars who are sympathetic to the inclusive fitness hypothesis are now deemphasizing relatedness as we saw in Peter Godfrey-Smith's (2009) statement above.

Even in situations where relatedness is associated with altruism, it could simply be a non-causal correlate with an underlying driver of the behaviour. For example, altruism cannot occur between any set of individuals unless they have the opportunity to interact. Individuals cannot interact unless they live in sufficiently close proximity for one's behaviour to impact the other's (i.e. within communication range). Such proximity between individual organisms very often correlates with their genetic relatedness, i.e. the relatedness between individual animals in any circumscribed physical location is likely to be higher than the relatedness between them and other individuals farther afield. This, Hamilton recognizes (1964b, 20, 22). The correlation of genetic relatedness with proximity has often been misconstrued to suggest genetic relatedness as the condition for

altruism, whereas it may be a factor that merely coincides with spatial proximity. Imprinting experiments as well as adoptions of genetically unrelated infants demonstrate that organisms by instinct would react a certain way towards another individuals whose presence coincides with a certain combination of factors. There is a high probability that the organism present at that moment when the instinct is triggered would be a genetic relative such as a mother, offspring or sibling. However, the fact that adoptions do occur between genetically unrelated individuals under those circumstances is an indication that relatedness might not be the underlying criterion. Also, explanations that suggest discrimination based on kinship often tend to suggest conscious and teleological mechanisms, which most of the organisms involved are not capable of.

VIII. Summary

The objective of this chapter was to identify the source of the difficulty of explaining altruism under Darwinian natural selection. I used a few examples in the history of science to illustrate the general nature of such theoretical anomalies and how solutions have often come about. Using the difficulties of Copernicus' heliocentric model of planetary motion, Alfred Wegener's continental drift hypothesis, and even Darwin's own difficulty accounting for variations, I showed that such difficulties often result from some erroneous auxiliary assumption or auxiliary hypothesis in the theory. Applying this insight to the problem of altruism I suggest that altruism remains a problem because we

assume in our models that it is a contrasting Mendelian genetic allele to selfishness, which is erroneous, and in fact, contrary to empirical evidence. In the next chapter I argue in detail why the assumption in our models that altruism and selfishness are contrasting Mendelian alleles is erroneous.

Chapter Two

ALTRUISM: THE PARADOX THAT NEVER WAS

Abstract

The notion of two competing alleles, one expressing the selfish phenotype and the other the altruistic phenotype, is fundamental to extant conceptualization and modelling of altruism. On the contrary, field observations in fact suggest that in almost all cases, the same individual may express the phenotypic condition we call "altruism" or the phenotypic condition we call "selfishness" contingent upon cues to which it is exposed in the (social) environment. The kind of flexibility that individuals express between the contrasting phenotypes of altruism and selfishness suggests the traits are unlikely to be controlled by contrasting Mendelian alleles (allelomorphs). In today's post Mendelian genetics, the two traits are better construed as alternative phenotypes of a single genotype. Yet, the best known explanatory models today presuppose separate alleles for the two traits. There is a glaring inconsistency between our theoretical assumptions and the empirical facts on the ground, which if addressed might make altruism much less of a paradox than it is now.

I. Introduction

The "Paradox of Altruism," is arguably the most enduring riddle in evolutionary biology. In extant conceptualization, it fits what William (1981, 164) describes as "the classic problem …of a mechanism by which a behavior can evolve (genetically) even though it lowers the fitness of the individual engaging in this behavior". But why is that a problem? Darwin's theory of natural selection predicts that a fitter trait would increase in frequency in a population, whilst the less fit alternative would be eliminated from the population. In line with that, we imagine a less fit "altruistic allele," which ought to be eliminated from populations by a fitter "selfish allele."

I think that has been a mistaken conceptualization; and in this chapter, I bring forth the evidence to bear upon that. I point out that the paradox of altruism emanates from the assumption of such an altruistic gene of lower evolutionary fitness in competition with a selfish allele. All the efforts put into developing explanatory theories such as kin selection, cooperation, group selection etc. would otherwise not be necessary if altruism and selfishness were not misconstrued as contrasting alleles (allelomorphs) in competition. The evidence herein presented is incontrovertible, that the best known models of altruism today do unambiguously assert distinct altruistic and selfish alleles. On the contrary, I demonstrate through overwhelming empirical evidence, that in fact, it is rather the case that the altruistic and selfish phenotypes are plastic expressions of a single

46

genotype under alternative environmental circumstances. With that conceptualization, a most effective and parsimonious account of altruism emerges, in which altruism presents no evolutionary anomaly in the first place. A plastic gene establishes by evolving ever more efficient criteria for expressing the alternative phenotypes in ways that are most favourable to its fitness. If we respect the observational data and construe altruism and selfishness as dimorphic phenotypes of a single genotype rather than as phenotypes of two competing genetic alleles, we would realize that we do not need any auxiliary hypotheses to explain how a flexible gene that expresses altruism under certain conditions and selfishness under other conditions can evolve by natural selection. Some may argue that the phenotypic flexibility I am pointing out here is recognized and supposed in many of today's models. I do not deny that. In fact I point to that and show how it is inconsistent with the assumption of altruism and selfishness as allelomorphs. I suggest that that kind of flexibility is more consistent with altruism and selfishness as the plastic expression of a common genotype.

II. A Castrati Aristocracy

Suppose that in a particular society there is a tradition backed by law, that requires every first born child, whether male or female, to be surrendered to the state. A couple can then keep and raise any subsequent children they have, who can go on to marry and procreate, and their first born children also taken by the

state and so on. Children so surrendered to the state are raised as wards of the state. They are neutered at puberty, and are not to marry or have families of their own. This as well as other safeguards, is to ensure that they will not be corrupt, since within the means of the state all their needs and wants are provided. Such wards of the state are trained to be soldiers, police officers, judges, legislators and other high state officials. Let us call them *The Guardians* (after Plato). Given what we know about such a *guardian caste*, would their persistence in any such society present an evolutionary anomaly that needs to be explained by any auxiliary theories? In this article, I contend that it is a similar situation happening in the caste systems of the eusocial organisms, where some individuals are allowed to develop into "selfish" reproductive adults, while a significant number of other individuals in each brood are made into neutered "altruistic" workers not by their genes but by their society. Yet, we suppose in the latter case a gene responsible for the altruistic trait and consequently a paradox there to explain with such theories as kin selection, group selection, multilevel section, etc. Further in the thesis, the reader will find illustrations of the same flexible expressions of altruism and selfishness in non-eusocial altruism such as reciprocity and cooperation.

III. Extant Genetic View

In his review of Sober and Wilson's (1998) book, *Onto Others*, Maynard Smith (1998, 639) notes that "there are two kinds of individual:" altruists, who benefit

others at a cost to themselves, and non-altruists who do not. A field example of these two distinct individuals is given by Okasha (2009)[1] who writes: "To see this, imagine that some members of a group of Vervet monkeys give alarm calls when they see predators, but others do not." In this example, those individuals who call the alarm are the altruists and those who do not call the alarm are selfish.

The question is: what is the genetic relationship between such altruistic and selfish individuals in a population? Is it a case as described by William, (1981, 165) that "differences among phenotypes are causally associated with genotypic differences (in other words) genetic differences underlie phenotypic differences?" In extant evolutionary thought, especially in population genetic models, the altruist and non-altruist are "presumed" to be distinguishable genetically by the possession or lack thereof, of "a gene for altruism." In explaining kin selection, Okasha (2009) invites us to "imagine a gene which causes its bearer to behave altruistically towards other organisms." He suggests "organisms without the gene are selfish;" and then goes on to say that "the altruists will be at a fitness disadvantage, so we should expect the altruistic gene to be eliminated from the population." That renders most concisely the problem of altruism as conceived by evolutionary biologists today. More so, the article cited describes "biological altruism" in the Stanford Encyclopaedia of Philosophy. So it must at least be the mainstream conception. Okasha (2009) vehemently denies

---

[1] All entries of Okasha (2009) *was* first published Tue Jun 3, 2003; substantive revision Tue Oct 28, 2008

"genetic determinism" in this kind of genetic supposition in evolutionary models. In other words, as he writes (2009): "Kin selection theory does not deny the truism that all traits are affected by both genes and environment."

In this work, I make no charge of the kind of genetic determinism Okasha denies. What I report about extant altruism models, which is very explicit in the relevant quotation above from Okasha (2009), and the other citations to follow, is that separate and contrasting genes for altruism and selfishness are supposed. This is further demonstrated by Trivers (1971, 36) when he sets the genetic assumption for his reciprocal altruism model thus: "Assume that the altruistic behaviour of an altruist is controlled by an allele (dominant or recessive), a_2, at a given locus and that (for simplicity) there is only one alternative allele, a_1, at that locus and that it does not lead to altruistic behaviour."

The general understanding amongst biologists, and which Okasha has here reminded us, that there is often some influence of the environment in gene expression, is never taken to ambiguate the terms "genetic trait" and "non-genetic trait," In this sense, the statements from Okasha, Trivers and others cited below suggest altruism is a genetic trait in the sense that it is for genetic reasons that an organism is altruistic rather than selfish. If this is not what those assumptions are intended to imply, then they should not be made at all. The genetic assumption of extant models of altruism that I have cited variously here carries certain logical

implications, and it would be disingenuous for us to escape those implications under the pretext of "the influence of the environment on genes."

As further evidence of the genetic stance of today's models of altruism, Bowles (2006, 1569) supposes in his group selection model that "(**A**) individuals are bearers of a hypothetical "altruistic allele"; those without the allele (**N**s) do not behave altruistically." In Haldane's (1932, 208) model, *aa* is the recessive character that causes altruistic behaviour. Similarly, Rousset and Roze (2007, 2321) engage in a very elaborate mathematical analysis of the possible evolutionary outcome of a "helping allele (H0)" versus a cheating allele (H1)." Sober (1984, 184) supposes an altruistic trait "A – one that causes individuals with the trait to benefit others at their own expense."

There appears to be a general academic consensus among evolutionary biologists, that we could draw a genotypic distinction between these two broad categories of individuals in a population. This common supposition emanates from the fundamental population genetic template for modeling evolution through gene frequency changes as shown here in figure 1. From Halliburton (2004, 133).

Table 1. Allele Frequencies Chart

Genotype	A_1A_1	A_1A_2	A_2A_2
Frequency	P^2	$2pq$	q^2
Fitness	w_{11}	w_{12}	w_{22}

In modeling the evolution of altruism, extant models build upon this template, in which they assume altruism and selfishness to be the contrasting alleles A_1 and A_2. That is why I have referred to this approach by extant evolutionary models of altruism as the *Altruism Selfishness Allelomorphism (ASA)* models (also see p. 26) as opposed to what I see the empirical data to suggest, i.e. altruism and selfishness as alternative phenotypes of a single plastic genotype, hence, the *Altruism Selfishness Plasticity (ASP)* model. In the latter model, we cannot represent altruism and selfishness separately as A_l and A_2 in table 1; a move which is consistent with the empirical data. The consequence of the ASA approach is that the "altruistic allele" would be of lower evolutionary fitness and therefore ought to decline in frequency. In fact the ASA assumption is taken as a given under the Darwinian-Mendelian population genetic paradigm prior to any auxiliary theories. It is its consequence of a declining frequency of the altruistic allele - which contradicts the empirical fact that altruism does evolve and persist in populations - that extant auxiliary theoretical models are designed to account for.

Extant models of altruism are generally inspired by this theoretically predicted attenuation of altruistic allele frequencies in populations. This Dawkins (1989, 184) describes in metaphorical terms as "cheat genes" spreading through the population while "sucker genes" are driven to extinction. Hamilton (1964, 1) introduces his *inclusive fitness* hypothesis by noting that "If natural selection

followed the classical models exclusively, species would not show any behaviour more positively social than the coming together of the sexes and parental care." Wilson (2005, 159) summarizes the problem thus: "How might such a behaviour evolve if the genes promoting it are at such a disadvantage in competition with genes that oppose it?" There is clearly a pervasive presumption of competing altruistic and selfish genotypes, and extant explanatory models hinge on that presumption. Under such a genetic view, the persistence of altruism in populations presents an anomaly. Hence, Haldane calls for mechanisms that will ensure that altruism benefits mostly other altruists. Possible mechanisms that have been suggested include: 1) altruists associating exclusively with other altruists (Maynard-Smith, 1998; Sober and Wilson, 1998) and 2) conditional deployment of the altruistic behaviour, i.e. only towards genetic relatives (Hamilton, 1964b) or towards other altruists (Trivers, 1971). Dawkins (1976, 89) makes the concept colourful for his popular science audience with his "green beard" metaphor, in which we are to imagine altruists identifying other altruists by a characteristic green beard. For group selection models, Godfrey-Smith (2009) explains that if social groups are formed randomly, "the **A** (altruistic) type is lost regardless of the details" (174). However, the altruist can be maintained, he explains further, "if groups are formed in a way that 'clumps' the two types, so like tends to interact with like [and] the benefits of having **A**s around tend to fall mainly on other **A**s" (174).

All of these safeguards that are sought extant (ASA based) models with the view to excluding the selfish individuals in the population from benefiting from altruistic acts would not be necessary if such models had a genetic conception of altruism in which the selfish individuals also carry and transmit the altruistic gene as the ASP model will show to be the case.

There is an interesting historical twist to this problem that might wake us up to empirical reality from the theoretical clouds in which we drift. In reproductive altruism for example, Darwin the naturalist saw a different problem, in fact, contrary to what the largely theoretical biologist is concerned about today in disregard of the field data. Darwin was befuddled by the field observation that the "selfish" queens reproduced the "altruistic" workers. He writes (1859, 353): "But with the working ant we have an insect differing greatly from its parents, yet absolutely sterile; so that it could never have transmitted successively acquired modifications of structure or instinct to its progeny." (see also Prete, 1990). So for Darwin the problem of such altruism was not a concern over the decline in the frequency of altruists under natural selection as we largely formulate the problem today. Rather, the enigma that he thought needed to be explained was the field observation in which there was a steady production of altruists (in fact, several fold more than selfish individuals), generation after generation by non-altruists. Now how does it come about that in spite of these field observations we think today that unless some special arrangements be made, as Godfrey-Smith (2009,

174) explains, "the **A** (altruistic) type is lost regardless of the details" (174). The problem for natural selection, Darwin thought, was how the peculiar features could evolve by natural selection if their bearers came from organisms that did not have them, and they, "being sterile, could never have transmitted successively acquired modifications of structure or instinct to [its] progeny." On the contrary, in today's explanatory paradigm, we seem to totally disregard this basic empirical fact of *the selfish begetting the altruist*, and instead approach the problem with the assumption of *like begets like*. So today we ask such questions as: "How might such a behaviour evolve if the genes promoting it are at such a disadvantage in competition with genes that oppose it" (Wilson, 2005, 159)? Statements such as these, regarding competition between the opposing genes of altruism and selfishness are inspired by the theoretical template in figure 1. However, that is not the problem we see if we are analysing the field data. If we are observing a eusocial colony, we would see that both the altruistic and selfish traits reside in the same individual, the selfish queen, who reproduces both the selfish and altruistic individuals. If we are observing other social organisms such as grooming baboons, alarm calling squirrels, or nest helping birds, we see the same individual behave selfishly at one time and altruistically at another; again, an indication that both traits reside in the same individual, and are not in competition. So concerns that "altruists will be a fitness disadvantage, [and] we should expect the altruistic gene to be eliminated from the population" (Okasha,

2009), emanate from theoretical assumptions, rather than promptings by direct empirical observation. It is because we ignore the empirical facts that we frantically search for mechanisms that will ensure that the altruistic acts benefit only other altruists to the exclusion of the selfish individuals. Godfrey-Smith (2009, 174) for example explains that altruists can be maintained, "if groups are formed in a way that 'clumps' the two types, so like tends to interact with like [and] the benefits of having **A**s around tend to fall mainly on other **A**s" (174).

I think that given what we observe especially in the eusocial organisms the least of our concerns should be the potential decline in the frequency of the altruistic allele, as our population genetic models predict. Neither do we need to contrive mechanisms as to how such alleles may be maintained in populations. For we know, through field observations (if we give a hoot about field observations), how the altruistic genes are generated in the population. The actual puzzle to explain then is how it comes about that the altruists come from the source they come from. That is, how such a reproductive arrangement can come about by natural selection; and that was exactly where Darwin saw the puzzle, based on the field observations.

If we subscribe to phenotypic plasticity as the genetic basis of altruism and selfishness, then the phenotypically selfish individuals in the population can also transmit the altruistic gene and need not be excluded from receiving altruistic acts. That is the case with the selfish queens in the eusocial organisms as we

56

shall see later. As I will also show later, even in non-eusocial altruistic societies, such as human, baboon or ground squirrel societies, selfish individuals have the capacity to reproduce altruistic individuals, and can even perform altruism themselves given the right social circumstances.

At least one scholar (unpublished) has argued that evolutionary models such as *Evolutionarily Stable Strategies* (ESS) and *game theories* only discuss the interaction of phenotypic strategies and indicate no genotypic relationships and therefore do not fall into the ASA class of models, and consequently immune to my critique of extant models. I think most scholars are rather of the view that those ESS theories generally imply, and expect us to understand each phenotype to represent a corresponding genotype. After all, isn't it the case, that Darwinian natural selection applies only to phenotypic traits that are heritable? Unless may be those ESS theories are not Darwinian. For if they are, Darwin, in talking about evolution by natural selection, makes it clear that: "Any variation which is not inherited is unimportant for us" (1859, 31). In any case, how is a behavioural strategy evolutionarily stable if we do not assume it to be transmittable by an underlying heritable factor? Also, Why do we co-opt such game theories into our efforts to explain altruism, if not because we want to explain the evolutionary sustainability of the genes that code for the different strategies, especially the cooperative strategy? In his seminal paper on *Evolutionarily Stable Strategies* (ESS), Maynard-Smith (1972, 21) explains that "a strategy qualifies as an ESS if,

in a population in which most individuals adopt it, there is no alternative strategy which will pay better." But what does "pay better" mean? Maynard-Smith explains: "The utility of an outcome is simply the contribution that outcome makes to the fitness of the individual – that is, to the expected number of future offspring born to that individual" (21). So the pay-off of each strategic phenotype is cashed out in terms of transmittable genetic fitness that underlies that phenotype. Maynard-Smith (1972, 20) writes for example in that specific case that "… any mutant individual adopting $E{\rightarrow}E$ will be favoured by selection." Of course, the strategic phenotype $E{\rightarrow}E$ is selected through the survival and reproduction of its corresponding genotype. We should note that if any of the evolutionary game models imagined any of the competing strategies to be expressed by a common genotype, the criteria for stability and elimination of the different strategies would be different from what they suggest in those models. For example, if a strategy X is fitter than a strategy Y, we cannot expect Y to be eliminated by X as the models predict if the two strategies are expressed by the same gene.

So it appears every evolutionary model today explicitly or implicitly assumes *genotypic dichotomy* or ASA. But how realistic is this notion of separate altruistic and selfish genotypes competing in social populations? Is it by virtue of a genotypic difference that some individuals in a population behave altruistically while others behave selfishly? If we say the worker bee is altruistic and the queen bee is selfish; and if by calling the alarm a ground squirrel is altruistic and the

58

adjacent squirrel that does not call the alarm is selfish; and if also by altruism we mean the vampire bat sharing blood with a roost mate, and the bat that refuses to share we call selfish; if these are the criteria by which we designate altruists and non-altruists in populations, it seems rather presumptuous that we should be seeking to explain the evolutionary sustainability of an "altruistic genotype of lower fitness." These examples, which are the most compelling cases of altruism, do not in themselves establish a genotypic difference between an individual that behaves altruistically and another that behaves selfishly. Rather, as I shall demonstrate in the course of this discussion, either behaviour is consistently associated with certain specific environmental cues, suggesting that any individual that is exposed to any such environmental stimuli would exhibit the behavioural phenotype that is associated with that environment. This point is lost in extant models of altruism because the altruistic and selfish individuals within a population are assumed to share an identical (social) environment, and that indeed is a grievous error.

IV. Genes and Environment

The interplay of genes and the environment in shaping phenotypes is not news in science today. We now know also, that there are many genes that would not express phenotypically unless triggered by some environmental cue. In such cases, two individuals carrying the same gene could nevertheless differ phenotypically with respect to that genotype due to differences in their

environmental experiences. Figure 2 (from Agrawal 2001) shows two clones of the water flea *Daphnia lumholtzi*. The individual on the left, with the spiny helmet and longer tail spine was raised in an environment in which chemical cues from a predacious fish were introduced. The other clone (on the right) was the control. The experiment demonstrates that any individual from that species can assume either phenotype depending on whether it is growing in an environment with predators or in one without predators. Hence, the two phenotypes, even though once thought to be separate species, actually do not differ with respect to the genotype for helmet.

Figure 2. Clones of *Daphnia* raised in different environments. (From Agrawal 2001).

The property of a given genotype to produce different phenotypes in response to distinct environmental conditions, defines the concept of *phenotypic plasticity* (Pigliucci, 2001, 1). In the case of *Daphnia* it is helmets that are expressed in response to the presence of a predatory fish. There are many other

examples of phenotypic features that are expressed only under certain environmental cues. Other such cues include: parasites (More, 1995); diet (Greene, 1989; Pfennig and Murphy, 2000); predators (Lively, 1986; Agrawal 2001,); competition (Harvey et al 2000); population density (Deno et al, 1992); temperature (Roth, 1986; Morreale et al, 1982). In all of these examples, as in *Daphnia*, a single genotype expresses one phenotype or the other, depending on the presence or absence of specific environmental cues. Clearly, any theory that assumes a genotypic difference between such dimorphic phenotypes would run into inconsistencies and other difficulties. For example, in table 1 it would be wrong to designate helmets as allele A_1 and non-helmets as allele A_2. But should we be doing that for altruism and selfishness?

In attempts to explain the evolutionary sustainability of altruism extant models assume a genotypic difference between the altruistic and selfish phenotypes. Could that be an error? Could it rather be the case that altruism and selfishness, like helmets and non-helmets in *Daphnia*, are dimorphic phenotypes of a single genotype? Could the *Daphnia* experiment be duplicated for altruism? In other words could we raise both altruistic and selfish individuals from a single set of clones?

Incidentally, evidence of this abounds in nature and we do not need to perform experiments to demonstrate this. In the polyembryonic wasp (*Copidosoma floridanun*), clones from a single embryo differentiate into altruistic

61

soldiers who do not reproduce but defend the selfish ones who reproduce (Donnell et al, 2004). Other cases of clones differentiating into altruistic and selfish individuals have been reported in gall aphids (Blackman, 1977). So altruism and selfishness can and do arise as plastic expressions of a single genotype.

Should there be any inclination to think that these are obscure anecdotal examples, consider some of the best known examples of altruism in the next section below. It would appear that there is a parallel between those today who assume a genotypic dichotomy between altruistic and selfish individuals, and those in the past who assumed a genotypic difference between the dimorphic phenotypes of *Daphnia* above.

V. The Hymenoptera/Daphnia Parallel

The discussion thus far has attempted to establish two views of the genetics of altruism. The traditional view which encapsulates models that suppose altruism and selfishness as allelomorphs (ASA) i.e. they assume a *genotypic dichotomy*, or an underlying allelic difference between the altruistic and selfish phenotypes; and then a new contrary view that suggests a *phenotypic plasticity explanation* for altruism, which I dub the *altruism selfishness plasticity* (ASP) model. Hamilton's (1964) *Inclusive Fitness Hypothesis*, the flagship of extant ASA models, is still touted by many as the best resolution of altruism, especially in the eusocial insects. Let us examine the detailed empirical observations of altruism

as expressed in the social hymenoptera, and see which of the two views (ASA and ASP) better explains the behaviour.

In a honey bee colony for example there are three castes consisting of a queen who does nothing but reproduce; a few hundred males called drones who also do not do much other than wait for an opportunity to mate with a queen; then there are thousands of non-reproductive females called workers, who toil all their lives taking care of the colony, including foraging and possibly laying down their lives when that is necessary in order to defend the colony. The reproductive queen has been designated as selfish whilst the non-reproductive workers are traditionally viewed to be the altruists. In fact, Shanahan (2004) regards the behaviour of the worker castes of the social insects as the epitome of altruism.

Now let us analyze some empirical evidence and see what we can learn about the underlying genetics of the two honeybee behavioural phenotypes. It has been known and documented since the 1830s that "a fertilized honeybee egg, which would normally yield a worker bee, will give rise to a queen bee if the ensuing larva is fed "royal jelly" (Prete, 1990, 273). Detailed modern studies have revealed further that whether a bee larva is raised a queen or a worker begins with the type of honeycomb cell into which the egg is laid (Winston, 1987). The workers will rear a larva as a queen if it is in a queen cell, by feeding it royal jelly. On the other hand they will rear it as a worker if it is in a worker cell, by feeding it worker food. The eggs and early larval stages are totipotent (i.e. can develop into

different functional entities). According to Winston (1987, 66) an egg or larva less than three days old that is moved from a worker cell into a queen cell will be fed royal jelly by the nursing workers and it will consequently develop into a queen. Conversely, an egg or larva transferred from a queen cell into a worker cell will be fed worker food and it will consequently develop into a worker. This is a very powerful indication that there is no genetic basis for initial placement of an egg in a queen cell or a worker cell, and whether a bee becomes a reproductive (selfish) queen or a non-reproductive worker (altruist), is determined by an environmental stimulus (i.e. diet) rather than genotype. We now know the specific genes in the honey bee for example, whose differential expression results in the selfish queen and the altruistic workers (Evans and Wheeler 1999), and they are plastic genes that are common to both the selfish and altruistic castes. Patel et al (2007) have detailed the signalling pathways by which different diet regimes activate or depress generic genes to yield different honeybee castes. In fact, intercaste individuals (i.e. individuals with both queen and worker features) have been artificially created by the experimental manipulation of larval diet (Winston, 1987, 68).

In other examples, experimental studies indicate that in the eusocial wasps (Vespidae), differences in nutrition during larval development are often the basis of caste determination (O'Donnell, 1998). In other species of social insects, it has been demonstrated that individuals can make a transition between altruistic

and selfish behaviour through experimental manipulation of their environments (Field et al, 2006). The observation, that each female honey bee has the potential to develop into a queen or a worker suggests both phenotypes are expressed by the same genotype in response to different environmental stimuli (e.g. diet regimes), rather than separate genotypes coding for the two phenotypes. Thus the genetics of altruism in this case is very much like that of helmets and non-helmets in *Daphnia*, and therefore an unmistakable case of *phenotypic plasticity*, which is defined as "the environmentally sensitive production of alternative phenotypes by given genotypes" (DeWitt and Scheiner, 2004, 2). This should be quite obvious in the social insects. West-Eberhard (1986) lists the queen-worker dimorphism in the social insects as one of the examples of alternative phenotypes that are produced by genes borne by all individuals of the population. Wilson (2008, 18) has also come to the understanding that "The different roles of the reproductive mother and her non-reproductive offspring are not genetically determined." Rather, he continues, "as the evidence from primitively eusocial species has shown, they represent different phenotypes of the same recently modified genome." A genetic switching mechanism then triggers such alternative phenotypes depending on the developmental stage or some environmental stimulus. Unless one rejects this entire catalogue of empirical data, it would be wrong to designate altruism and selfishness as the allelomorphs A_1 and A_2 in figure 1 above as the ASA models do. Also, it is clear that, with this kind of

genetics, the altruistic behaviour is at no fitness disadvantage relative to the selfish behaviour in the social insects. If so, why are evolutionary biologists as cited above, jumping hoops to provide circuitous explanations as to how an altruistic gene of lower evolutionary fitness against a selfish alternative can be sustainable, when that is totally not the case?

The answer becomes clear if we juxtapose the observations noted above against the genetic assumptions in the most widely accepted explanations of altruism today. Both Haldane (1932, 208) and Maynard-Smith (1964) assume altruism to be caused by a Mendelian recessive character aa as opposed to the characters AA and Aa for the non-altruistic condition. Hamilton (1963, 354) supposes "a pair of genes g and G such that G tends to cause some kind of altruistic behaviour and g is null." Trivers (1971) assumes an allele a_2 that controls altruistic behaviour and an alternative allele a_1 that does not lead to altruistic behaviour. These scholars clearly assume a genotypic dichotomy between the altruistic and selfish phenotypes, which in turn precipitates the concern over a declining altruistic allele and consequently the paradox of altruism and a scramble for explanations and auxiliary hypotheses. From these scholars emanated inclusive fitness, kin selection, reciprocal altruism and the ESS models. Of course, the story is no different for the proponents of group selection. The reason for this disconnect between extant theoretical assumptions and the observational data is found in Lakatos' (1978, 50) astute observation that

scientists working in a given research program or paradigm often have their attention so riveted in building their models according to the instructions laid down in the research program that they "ignore the actual counterexamples; the available data." In this case of altruism, the neo-Darwinian heuristic that is riveting the attention of evolutionary biologists to the neglect of the empirical data is the population genetics template (table 1) for evolutionary analysis that crystalized out of the union Darwinism and Mendelian genetics.

VI. Altruistic Expression and Social Cues

In section V, honeybee society was used to demonstrate how the altruistic and selfish phenotypes in the social insects are determined by different environmental cues rather than genetic differences. Now I turn to the non-eusocial social organisms. Starting with vampire bats, let us now examine the relative efficacies of the ASA and ASP models in explaining reciprocal altruism for which Trivers (1971) provided an explanation based on the ASA model. Vampire bats roost in dark places by day and go out at night to feed. For some species the diet is exclusively blood, usually from other mammals. There are occasions when some individuals will find very little to eat while others will be more fortunate. Researchers have observed that the hungry individuals would often solicit the individuals that are better fed for some food. Sometimes an individual would oblige and regurgitate some blood to a soliciting individual, while on other occasions individuals have also been observed to steadfastly refuse to share

food with a soliciting individual. It is traditionally held that those individuals observed to obligingly share their food with soliciting individuals are altruistic while those that refuse to share are selfish. Then under the ASA models we have to assume that the individuals that share blood carry the altruistic allele whilst those that refuse to share are under the influence of the selfish allele.

Closer observations reveal however, that whether a vampire bat shares blood or not in any situation would be determined largely by the circumstances at the time, such as, whether the solicitor has given the actor blood before (Wilkinson, 1984), or whether the solicitor is judged likely to give blood to the donor when he is in need. If so, we could suppose that the vampire bat that is seen today sharing blood with a neighbour and judged to be doing so under the expression of an "altruistic allele," could on another occasion be seen steadfastly refusing to give blood to a bat that is starving, possibly because the then solicitor may have refused to share previously. Hence, the bat that is characterized as the altruist today would be the selfish individual on some other occasion. Since organisms are not known to change genotypes in that manner, the difference between sharing then (altruism) and refusing to share now (selfishness) is not a matter of genes but largely the circumstances of the (social) environment. Thus the social environment, like chemical cues in *Daphnia* and diet in the honeybee, serves as a cue for the conditional expression of altruism and selfishness as dimorphic behavioural phenotypes.

Among Belding's ground squirrels, mostly adult females make alarm calls, and the frequency of the calls has been observed to correlate with the presence of relatives Sherman (1977). Thus, the alarm calling behaviour seems to be conditional, depending upon the presence of relatives. That is exactly what Hamilton (1964) suggested would enhance *inclusive fitness*, and many would celebrate this as a triumph for kin selection. Let me just point out here before I proceed with the current chain of thought, that evidence of kin motivated altruism here and there does not establish kinship as necessary or sufficient for altruism (see page 34). The point with this example is that the fact as it indicates, that the same individual can behave altruistically (i.e. call the alarm) at one instance and selfishly (i.e. refuse to call the alarm) at another, undermines the underlying genetic assumption upon which kin selection is established. Rather than two separate genotypes causing the two phenotypes as the architects of kin selection, Hamilton (1964) and Maynard-Smith (1964) suppose, the behaviour suggests a dimorphic phenotypic expression of a single genotype. Hence, it is consistent with the ASP model and contrary to the ASA models.

Among olive baboons (*Papio anubis*), Packer (1977) reports that an adult male will give aid to a soliciting troupe member based on whether he has received help from the solicitor before or whether the solicitor is deemed capable of giving meaningful help when it is needed. Thus an adult male is more likely to deny aid to soliciting juveniles and females during fights. It has similarly been

reported in vervet monkeys (*Cercopithecus aethiops*), that whether an individual responds to a solicitation or not depends on whether it has previously received grooming (or aid) from the solicitor, in addition to other social considerations (Seyfarth and Cheney, 1984). If giving aid is "altruism" and refusing to help is "selfishness," then it is evident here that external factors, rather than genotype, determine whether an individual behaves altruistically or selfishly.

A pattern thus seems to emerge from the key examples of altruism analyzed here: that an individual will respond altruistically only when certain environmental circumstances are present, and would respond selfishly if those environmental cues were lacking. It is no different from the arctic fox expressing white fur in the winter and brown in the summer. It is important to note that no evidence has yet been presented to date that demonstrates that under the same set of environmental circumstances only certain individuals (i.e. those who carry the altruistic allele) are capable of reacting altruistically, while others will always refuse to assist because they lack the altruistic gene. In other words, there is no empirical evidence of such an altruistic allele that serves as the underlying distinction between the altruistic and selfish phenotypes. To prove genotypic dichotomy we need to demonstrate for example, that certain honeybees will always mature into workers (altruists) irrespective of diet or any other external factor; that certain members of a vampire bat colony will always share food even when the solicitor is one who has consistently refused to share; that from one

external circumstance to another, only certain individuals will consistently call the alarm while others would never call the alarm under any circumstance. The ASA models presume these tests to be met. In reality there is no basis for such a presumption.

What is clear and consistent from the studies cited here is the association between certain environmental cues and the expression of the altruistic phenotype, while other circumstances trigger a selfish response. For example, a baboon gives aid (altruistic) under one circumstance and denies aid (selfish) under another. That strongly suggests a plastic behavioural response of a single genotype to different (social) environmental circumstances. Thus like the *Daphnia* example in section III, there cannot be separate genotypes for altruism and selfishness, since each individual in the population has the capacity to express both phenotypes.

VII. Reproductive Altruism and the Social Environment

Reduced fecundity in deference to others has often been cited among the examples of altruistic expression (Shanahan, 2004; Okasha, 2009). Reproductive suppression (or even exclusion) of subordinate females and males is a common feature of animal social organizations. In these cases of altruism it becomes ever more preposterous to imagine that a genotypic difference could be causing the behavioural difference between the individuals who reproduce and those who do not. Observations indicate very strongly that the "altruistic" behaviour is imposed

by external circumstances rather than by specific genotypes. In social mammals for example, it is often the dominant female or male that prevents the others from breeding, through a variety of schemes, including physical deterrence from mating. In the naked mole rat (*Heterocephalus glaber*), pheromones given off by the dominant female act on the hormonal systems of subordinate females to render them infertile (Faulks et al, 1991). Those pheromones are analogous to chemical cues from predacious fish in the case of *Daphnia*, and it is they rather than genotype that elicit the non-reproductive altruistic behaviour in the mole rat. In meerkat societies the reproductive efforts of subordinate females are deterred and disrupted by the dominant female (Young and Clutton-Brock, 2006). In the case of helper birds, individuals are forced to assume the non-reproductive (helper) position by external circumstances such as demography, rank and availability of nest cites (Rabenold, 1985) rather than the dictates of some "altruistic gene" in the helper. Yet these are all frequently cited examples of altruism, whose sustainability evolutionary biologists are confounded by as they attempt to explain it by assuming ASA.

In most social situations, it is where an individual ranks in the social structure, that determines whether it reproduces or not. In hyena and wolf packs for example, only the alpha male and female breed and the rest of the pack we must call altruists. However, upon the death of the alpha female, as observed in the naked mole rat (*Heterocephalus glaber*) by Lacey and Sherman (1991), any

of the non-reproductive (altruistic) females can undergo some hormonal changes and ascend to the role of the reproductive (selfish) female. A similar observation has been made with the termite, *Zootermopsis nevadensis*, in which a replacement is drawn from amongst the workers, upon the death of the king or queen (Johns et al, 2009). This means a phenotypic transformation of an altruistic worker into a selfish king or queen. Recall the sex change behaviour of the marine goby (*Coryphopterus personatus*) (Allsop and West, 2004) from the literature of phenotypic plasticity. In this case also, the same individual can be non-reproductive (altruistic) in one social circumstance and become reproductive (selfish) when the circumstances change. Such transitions between the altruistic and selfish phenotypes by individuals belie the assumption of an underlying genotypic dichotomy between the phenotypes in current models.

Even in the hymenoptera, amongst whom we find some of the most extreme cases of caste based phenotypic modifications, functional ovaries are maintained in the non-reproductive (altruistic) castes (Oster and Wilson, 1978) and some of them do indeed reproduce (become selfish) under given social circumstances (Ratnieks and Visscher, 1989). Gadagkar (1997, 28) agrees that a social organism would assume a subordinate role not because of any altruistic reasons but because it is the best of the available alternatives. In the social wasp *Ropalidia marginata*, Gadagkar (1997, 72) reports that individual wasps can act as queens or workers, in response to the opportunities available. He observed

further that often a worker would later drive its mother (the queen) out and become the queen. The change in status or phenotypic behaviour from worker to queen has also been reported in other social insects (Field et al, 2006). As Queller (2006, 42) observes in the eusocial insects, "workers are not leaping at every opportunity to be altruistic, they are coerced." Coercion as a trigger of altruism, in the absence of which an individual would rather remain selfish, is indicative of the plastic phenotypic deployment of a common genotype. Wensellers and Ratnieks (2006) also conclude from studies of ten social insect species that "it is mainly social sanctions" that keep individuals altruistic where they would otherwise have behaved selfishly. Emlen and Wrege (1992) report that in the white-fronted bee-eater (*Meropsis bullockoides*), young males are forced by older nest-owning males into helper status by harassment and disruptions of their attempts to set up their own nests. Such "altruistic" helpers can change their status to "selfish" reproductive nest owners whenever the opportunities arise in the future. In fact, in meerkat societies as Young and Clutton-Brock (2006) report, not only are our designated altruists (the subordinate females) able to express the selfish phenotype by reproducing when they get the opportunity, they are able to match the selfishness of the dominant female by murdering the infants of other mothers.

In all these examples, it is remarkably consistent that the altruistic and selfish phenotypes are determined by environmental circumstances rather than

74

genotype. One very crucial observation is that individuals are often able to make transitions between the two phenotypes in response to changes in their social environment. These facts are clearly, inconsistent with the notion of two separate genotypes for altruistic and selfish individuals as assumed by extant genetic models of altruism.

VIII. Inclusive Fitness and Reproductive Division of Labour

Hamilton's *inclusive fitness hypothesis* is credited anytime altruism is found to coincide with a high value of the *coefficient of relatedness* (r). It has in fact been suggested (Trivers and Hare, 1976) that the high value of (r) due to haplodiploidy is responsible for the evolution of reproductive division of labour in the social insects, and justifies a female honeybee for example helping to raise its sisters (r = 0.75) rather than producing its own offspring (r = 0.50). The question is: can *inclusive fitness* really explain reproductive division of labour on the assumption that the altruistic and selfish phenotypes differ with respect to an altruistic gene G? Inclusive fitness implies that the altruist benefits from the reproduction of the recipient of the altruism. For the altruistic gene to be sustainable, Hamilton (1963, 354) stresses that it "is not whether the behaviour benefits the behaver but whether it is to the benefit of the gene G." Curry (2006, 683) reiterates this point when he criticises those who base kin selection simply on "the overall proportion of genes that the individuals share." On how kin selection explains altruism, he writes (683): "Well, genes for altruism can spread if they help copies of

themselves that reside in other individuals." Unless we assume phenotypic plasticity, this would mean altruists helping only fellow altruists. If this is the basis of kin selection, then it fails to explain altruism in the eusocial organisms, because in those societies the altruistic workers mostly help the selfish queen and her offspring.

Thus, the theoretical requirement that the altruistic gene G be present in both altruist and recipient is logical, but it is a logic that exposes the flaw in the genetic assumption of the *inclusive fitness hypothesis*. The problem is that in reproductive division of labour it is almost exclusively the selfish individuals such as the queen bee or the alpha female (in mammals) that hold the purse strings of the evolutionary benefits of altruism; for it is they alone that leave progeny. How is the *inclusive fitness* dividend realized for the altruistic gene in those situations if we maintain that the altruistic gene is the underlying difference between the altruistic and selfish phenotypes (Haldane, 1932; Hamilton, 1963, 1964; Maynard-Smith 1964;Trivers 1971; Queller, 1985; Dawkins, 1989, 184; Okasha, 2009; Wilson, 2005)? If the altruistic gene is what sets the two phenotypes apart, it suggests that any relatedness between altruistic and selfish individuals is due to traits other than the altruistic trait. As Darwin (1859, 359) points out however, "peculiar habits confined to the workers and sterile females, … could not possibly affect the males and fertile females who alone leave descendants." Isn't that the assumption that generates the puzzle? That is, if the heredity factor that causes

altruism and the accompanying morphological modifications is confined to the workers, how does the queen "who alone leave descendants" reproduce them? Therefore, for altruism to be sustainable in societies with reproductive division of labour, the selfish individuals "who alone leave descendants," must necessarily carry the altruistic gene notwithstanding the fact that their behaviour is clearly selfish. If we suppose this however, the basis for proposing kin selection collapses.

Both Darwin (see Prete 1990) and Haldane (1932, 208) also came to this conundrum at some point in their analyses when they considered the circumstances under which workers and queens are derived in the honeybee, i.e. epigenetically from a common pool of eggs. They each concluded the queen also had to carry whatever hereditary factor it was that made the workers "altruistic." But for scholars of their respectful times who held the view of a one to one mapping of inherited trait to phenotypic expression, they found themselves with a conundrum; for how does the queen remain selfish in spite of carrying the altruistic trait? If both selfish and altruistic individuals carry the altruistic gene, then that gene cannot be seen to be the underlying cause of the phenotypic difference between the two individuals. The Mendelian may still pitch dominance and recessiveness to explain the latency. But that will not go far in light of the evidence of oscillations we have often seen between the altruistic and selfish

phenotypes within a single individual in a lifetime, which is not a feature consistent with the Mendelian dominance theory.

In fact, what Darwin (1859, 354) postulated, which would not be understood in genetic terms prior to our post-Mendelian era, was that what is transmitted from generation to generation is "a <u>tendency</u> to produce sterile members" among the fertile ones. "Tendency" indicates contingency in the expression of these traits; and it is some of those contingent factors that I am here enumerating. It is not one heritable factor (gene) for fertile individuals and another for sterile ones as table 1 suggests. It is a group trait, which tends to make the group produce some members that are sterile, according to Darwin. We have to bear in mind however, that these were blind chance hits that Darwin was making on modern genetic interpretations by his pronouncements. In his day he basically groped about in the dark on heredity. When Darwin formulated his theory, as Fisher (1929, vii) notes, the principle of inheritance was the vaguest element in its composition.

If we have to suppose therefore, that an individual that is behaving selfishly could actually be carrying the altruistic trait, then we will have to concur with the *Daphnia parallel* i.e. phenotypic plasticity as the underlying genetics of altruism. If a single genotype thus expresses both the altruistic and selfish phenotypes, it implies that there is no exclusive altruistic genotype of lower fitness. Consequently there is no need for an *inclusive fitness hypothesis*, which

purports to explain the paradox associated with that notional "altruistic gene." In almost every model of the evolution of altruism today there is a requirement of some form of conditionality as a necessary feature in the expression of altruism. However, there has never been any recognition or clear interpretation of this conditionality as a case of phenotypic dimorphism involving altruism and selfishness.

IX. Allelic Convergence

Contrary to the assumptions of the genotypic dichotomy supposed by the ASA models, both observation and theoretical requirements strongly suggest a convergence of the genetic traits of altruism and selfishness within each individual in social populations. Many including Darwin and Haldane have suggested, based on the way altruists and selfish individuals are determined in the honeybee, for example, that each individual in such populations has to carry both "traits." Hamilton's inclusive fitness hypothesis as we have seen above is valid only if the altruistic gene is present in both the altruist and the recipient, which in the case of the social hymenoptera means the selfish queen carrying the altruistic allele, since the queen is the recipient of most of the altruism. Hamilton (1964) suggests that inclusive fitness would be enhanced if it is accompanied by some capacity to recognize kin and subsequent discrimination in generosity of the altruistic behaviour. That would mean the altruist behaving selfishly when the solicitor is non-kin, and thus implying a capacity to express both traits.

Trivers' (1971) *reciprocity model* and its spin-off, the *cooperation* model of Axelrod and Hamilton (1981), suggest altruism may be sustainable if it is extended only to individuals who have helped the altruist in the past or are judged likely to help in the future. These suggestions entail a conditional expression of the altruistic phenotype. However, it is selfish behaviour when an individual withholds altruism under circumstances where some other individual would extend altruism. In modeling evolutionary explanations of altruism, we often consider two phenotypic behaviours – altruistic or selfish. In Queller's (1985) model, P_1 is the altruistic phenotype and P_0 is the selfish phenotype. For example, in a social population with an alarm calling altruism, an organism is altruistic if it calls the alarm on the approach of danger, and it is selfish if it does not. Thus, conditional altruism is tantamount to selfishness alternating with altruism in the same individual. That however, goes contrary to the assumption if genotypic dichotomy under the ASA models*, which is, that separate segregated alleles express the two separate phenotypes. Futuyma (1998, 586) points to many examples in nature that suggest an "Assessor strategy" in which the individual acts one way or the other based on the circumstances of each situation. In that case being selfish or an altruist is a temporary situational condition, which is subject to change in the next instance if the individual's social circumstances change. The capacity of each individual to behave with such

flexibility would require such individuals to have the underlying genetic capacity to express both behaviours.

Fehr and Schmidt (1999, 856) have shown through mathematical modeling that "in a public good game with punishment, even a small minority of selfish players can trigger the unraveling of cooperation." Similarly, as they demonstrate, "a minority of fair-minded players can force a big majority of selfish players to cooperate fully." Upon that, Fehr and Fischbacher (2003, 787) assert regarding human altruism that "it is not possible to infer the absence of altruistic individuals from a situation in which we observe little cooperation. If strong reciprocators believe that no one else will cooperate, they will also not cooperate." If the social environment can thus trigger the phenotypic transformation of an altruistic individual to a selfish one and vice versa, a Mendelian genotypic dichotomy could not be the cause of the phenotypic differences.

After discussing the selfish (i.e. "always defect") and the altruistic (i.e. "always cooperate") strategies of the hypothetical "Prisoner's Dilemma," Axelrod and Hamilton (1981) clarify that in nature the circumstances are quite different, in that organisms often have previous experience and memory, which do influence the choice between cooperating and defecting. Such a suggestion means in genetic terms that each individual organism has the option to express one of two phenotypes: 1. Cooperate (i.e. be altruistic) 2. Refuse to cooperate (i.e. be

81

selfish). In fact, that is one of the key conditions for the efficacy of Trivers' (1971, 36) *reciprocity model,* i.e. that "an altruist responds to [the] cheating by curtailing all future possible altruistic gestures to the individual." That simply means behaving selfishly towards such individuals in future. Why do we expect an individual that carries the altruistic allele rather than the selfish allele to nevertheless be able to behave selfishly some of the time? If that is indeed the case in nature as observations suggest, then that contradicts the genotypic dichotomy assumed by the ASA models. If extant models of altruism require an individual to express both the selfish and altruistic phenotypes depending on the social circumstance, it becomes difficult to say with confidence whether an individual behaves selfishly because it lacks the altruistic gene or because the circumstance do not allow the altruistic expression, even though the individual has the altruistic gene. That indeed makes the genetics of altruism so much like the genetics of helmets in *Daphnia*.

Beyond all these, direct evidence is beginning to trickle in from molecular genetic studies that clearly establish the alternate expressions of altruism and selfishness by a single plastic genetic factor. In the social amoeba (*Dictyostelium discoideum*) for example, a gene (*dimA*) has been identified (Foster et al., 2004; Thompson et al., 2004), whose differential expression triggers some cells to differentiate into non-reproductive (altruistic) stalk cells and others into reproductive (selfish) spore cells. A protein molecule, DIF-1 has been identified to

be the triggering factor. Even where differential sensitivity to DIF-1 has been suspected, the difference has been traced back to non-genetic factors such as "cell cycle position and growth history" (Thompson et al. 2004).

So the evidence for the claim here, of altruism and selfishness as alternate expressions of a single plastic genotype is incontrovertible. I am therefore not opposed to the invocation of conditionality by any other altruistic model. In fact I welcome that as a validation of my thesis, which rides on the empirical observations of conditionality as a feature of altruistic expression. The problem with the models I criticize is that the genotypic dichotomy assumption upon which they model altruism is inconsistent with the conditional alternate expression of the altruistic and selfish phenotypes. Such models start off with the Mendelian assumption that the altruistic and selfish phenotypes are genetic allelomorphs. Among the features of such models however, is often a conditionality clause, such that an individual is altruistic under one circumstance and selfish under another. Consider the following example. In his project to expand Hamilton's rule to explain reciprocal altruism, Queller (1985, 367) states the following evolutionarily stable strategy (ESS): "be altruistic to an untested partner, otherwise, behave towards the partner as he behaved towards you." So Queller here expects any individual in the population to be altruist if the partner he encounter's behaves altruistically, and to otherwise be selfish if the partner he meets behaves selfishly. Yet, this is Queller's stated genetic assumption for this

individual (367): "For simplicity, assume all members of a diploid population interact in pairs … and the genetic component of altruism is due to alleles at a single locus, and no overdominance." He goes on to define the variable "G = an individual's frequency of the altruism allele (0, ½, 1)." Of course in Mendelian terms 0 would mean the individual lacks the altruistic gene and 1 would mean he is homozygous. If any individual in the population must have one of these three possible values of G, which of them must he have in order to have the flexibility of being selfish or altruistic depending on the partner? Clearly therefore, this required flexibility in the expression of the altruistic behaviour undermines the idea of separate altruistic and selfish alleles in competition. On the other hand, if we respect the empirical data, and perceive altruism and selfishness as alternative phenotypes, which are both expressed in a single individual under different environmental circumstances, we avoid the conflict with our theoretical requirements.

X. Summary

This chapter has been a long conductive argument. It had to be so because I did not want to leave room for any thought that the evidence might be anecdotal and thus a reason for us to remain in our set ways. The empirical evidence presented here is not new to most evolutionary biologists. It is simply ignored because it does not fit well into the theoretical models we have been trained to use. This is reminiscent of the attitude of the early Darwinians towards Mendelism, which

rendered Darwinism sterile until they found a way to embrace the empirical reality of Mendelism. Similarly, today's empirical challenges to the status quo will continue to undermine our models of altruism until we address them with academic sincerity. So I will re-state the three bold claims in this article, which critics may reject by showing how the direct evidence I have supplied for them is flawed. 1. I have pointed out and supported by the relevant quotations and citations, that the best known theories of altruism today model altruism on the assumption of a genotypic dichotomy between the altruistic and selfish individuals in the population, which I say is wrong. 2. I then went on to suggest, upon a mountain of empirical evidence, that altruism and selfishness be viewed as plastic phenotypic expressions of a single genotype. I noted in this regard that models that assume genotypic dichotomy between the altruistic and selfish phenotypes often require altruists to withhold altruism i.e. behave selfishly under certain conditions; and that I say is inconsistent with their genetic assumption. In other instances, those models also require recipients of altruism who are clearly selfish to be able to transmit the altruistic trait, especially in the case of reproductive division of labour. Such requirements are more consistent with phenotypic plasticity than genotypic dichotomy. 3. Consequently, I conclude that if the empirical data thus point to phenotypic plasticity as the basis of the altruistic and selfish phenotypes in social populations, rather than separate altruistic and selfish genotypes, then models that seek to explain the sustainability of some

distinct altruistic genotype of lower evolutionary fitness (the puzzle of altruism),

may indeed be chasing a mirage.

Chapter Three

THE EVOLUTION OF PERSISTENT MALADAPTIVE TRAITS

ABSTRACT

It has been over a century and a half since the inception of the Darwinian evolutionary explanation, and the problem of how an altruistic population might evolve from a single incipient mutant in a selfish population remains unresolved. In this discussion, I identify the erroneous theoretical axioms at the source of the problem. They all stem from the genetic relationship extant models assume between the altruistic and selfish individuals in a population. I show how that erroneous assumption creates a barrier to the spread of an incipient altruistic genetype in a population and further demonstrate how the proper genetic assumption eliminates that evolutionary barrier.

I. Introduction

Traits that are maladaptive and nevertheless persistent in populations represent a troubling anomaly for Darwinism. Some, such as sickle cell anemia and sexual ornamentations in certain species, have been convincingly explained by the concept of trade-offs. However, others such as altruism, psychiatric disorders and other maladaptive behavioural traits have remained stubbornly insoluble. The problem of altruism especially has had a remarkable longevity, dating all the way

back to Darwin (1859 and 1871). The most intractable difficulty of all extant theories that deal with persistent maladaptive traits is explaining how they actually evolve from incipient single mutants rather than how an already existing frequency in a population may be maintained.

In this discussion, I draw heavily on work done in psychiatric genetics that provide insight into the true nature of behavioural genes, which is quite contrary to what evolutionary biologist suppose for altruism (which is also a behavioural trait). I go on to show how that erroneous genetic conception of altruism thwarts efforts to model its evolution. I then demonstrate how facile the evolutionary modeling of altruism becomes if the proper genetic description is made. Empirical evidence suggests that whether an individual expresses altruism or not depends on (social) environmental circumstances. Here we see how that conditionality (plasticity) in the expression of altruism and selfishness by a common plastic pleiotropic gene facilitates its propagation in populations. As a general model, the conditional expression of an incipient plastic genotype may enable it to increase in frequency in a population, if it has evolved mechanisms that trigger the alternative phenotypes in a way that enhances its fitness.

II. The Role of the Environment

By Darwinian principles, a mutation that has no net phenotypic advantage or disadvantage will not be affected by natural selection (Darwin, 1859, 108). This also occurs where a mutation does not express phenotypically (Kimura, 1968).

There are two conditions under which a maladaptive mutation with a plastic phenotypic expression may increase in frequency in a population. The first is, if the conditions that influence its phenotypic expression are sufficiently rare, thus keeping it latent most of the time. The second is, if the conditions that influence its phenotypic expression also result in a mechanism that in the long run over-compensates for the fitness depression of the maladaptive phenotypic expression: that is, if the benefit (in terms of fitness) to the gene is greater than the cost of the altruistic act. Hamilton (1964) suggested that this could happen in the case of altruism if the altruism is directed towards individuals who are kin. Trivers (1971) suggested reciprocity as another way it can happen without the necessity of kinship. Those two are different conditions under which a maladaptive phenotypic expression such as altruism may ultimately result in a net fitness benefit to the underlying gene. The problem with Hamilton and Trivers is that they set out their respective ideas as comprehensive explanations of altruism based on the assumption of altruism and selfishness as contrasting Mendelian alleles (allelomorphs). In actual fact both explanations work only because the gene that causes altruism has a plastic phenotypic expression such that carriers of it can also express selfishness under alternative conditions. That is anything but allelomorphic, and in fact those explanations, that assume separate genotypes, entail some logical difficulties, and are also strikingly contrary to the field data. For example in vervet monkeys, olive baboons and ground squirrels

the altruistic phenotype would not express unless the beneficiary fits a certain profile, in the absence of which the potential altruist would behave selfishly.

Due to the paucity of empirical data on the molecular genetics of altruism, I turn to related fields under the rubric of behavioural traits for some insight. There are a number of models supported by mines of empirical literature on behavioural genetics. Examples include a model for complex traits by Prichard (2001) and another for psychiatric disorders by Keller and Miller's (2006). In supporting *polygenic mutation-selection balance* as the best explanation for the persistence of harmful heritable mental disorders, Keller and Miller utilize the relationship

$$V_G = V_M/s$$

.where V_G is the equilibrium genetic variation of a polygenic trait, V_M is the increase in a trait's genetic variation due to new harmful mutations, and s is the average selection coefficient against the mutations. That equation explains variability by comparing the rate of new mutations to the rate at which natural selection eliminates those mutations. The most interesting element in this equation, especially for a behavioural trait, is the factor s, which incorporates the role of the environment. Keller and Miller attribute the large value of V_M for heritable mental disorders to the high mutation potential in the brain due to the large proportion of the body's total cell count that resides in the brain. For the small value of s, they link it to the idea that "mutations with milder effects are removed more slowly, so they tend to be more common." All these explanations

are quite plausible. But their effects are ultimately realized through the mediation of the environment. The way the expression of such mental disorder alleles vary with the environment is key to their persistent genetic variability, to which I turn in the next section.

III. Phenotypic Plasticity and Maladaptive Traits

Evolutionary biology has invested a great deal of theoretical capital on the putative concept of "allelic dichotomy" between the altruistic and selfish phenotypes in social populations. As demonstrated in chapter two of this research, the field observations are decisively contrary to that fundamental theoretical presumption. Further hints of the actual nature of the genetics of altruism may be gleaned from related research disciplines that deal with the relationship between genes and behaviour. In psychiatric genetics a great deal of effort has been invested into hunting down specific genes that predispose individuals to specific mental disorders. They have so far had little success in establishing any meaningful correlation between specific gene loci and specific psychiatric disorders (Keller and Miller 2006; Kendler 2005). The dismal progress in itself points to certain inferences about the genetics of such traits, which will be argued throughout this work.

A key feature of behavioural traits in general is the environmental contingency of their expression. Though the importance of the environment is very much bandied about in discussions of genes and behaviour, no one has

been able to fit together genes, the environment, plasticity and selection in the jigsaw puzzle that explains the apparent anomaly of persistent maladaptive traits. Instead, scholars often give due recognition to the role of the environment in the phenotypic expression of such traits and then move on to construct an exclusively genetic model that does not integrate the environment in any way

The environment is certainly very pre-eminent in the expression of altruistic and other behavioural traits, and we shall see why *phenotypic plasticity* presents the most robust explanation of the persistence of maladaptive traits in populations when we effectively incorporate the environment. As a matter of practical necessity, I render phenotypic plasticity in a very broad sense as the conditional phenotypic expression of an allele or genotype. In most situations the alternative phenotype is a default phenotype that results from the absence or inaction of the allele. Therefore, plasticity mostly involves situations in which an allele will either express phenotypically or remain latent, contingent upon factors other than its mere presence in an organism. The factors can be cast in two broad categories. The first category consists of other genetic factors such as: 1. Regulatory genes: These are genes that control the expression of other genes and could underlie some cases of plasticity (Windig et al 2004). In fact, for the purpose of this discussion, we can view regulatory genes broadly as genes whose presence facilitates the expression of other genes without regard to any mechanism. In that sense, this subsumes the genetic component of *multifactorial*

inheritance: Burghes et al (2001) have given a compendium of traits that are under multifactorial inheritance. 2. Additive gene thresholds: If the allele contributes additively to the trait, there may be a threshold requirement within the quantitative trait loci (QTL) (Falconer, 1989; Hazel et al, 1990), above which the allele would express phenotypically and below which it would remain latent, thus causing the alternative phenotype. 4. Recessivity: In Mendelian genetics, a recessive allele will express phenotypically only if an identical allele is present on the corresponding chromosome, otherwise it remains latent. In the typical Mendelian genotypic ratio of 1:2:1, the recessive allele has a latency rate of approximately 66.6%. 5. In another example of a genetic factor in the plastic phenotypic expression of another gene, Keller (2009) cites studies that reveal a peculiar caste determination system in some eusocial insects. In those systems there are two distinct lineages within a species. An egg that is fertilized by a male of a different lineage from the queen hatches an offspring that develops into a worker; whereas an egg fertilized by a male of the same lineage as the queen develops into a queen. So we can say here that the condition for phenotypic expression is whether the corresponding alleles are identical or contrasting at the locus involved.

The second category consists of environmental factors. Some of the environmental factors that influence the phenotypic expression of certain traits were cited in chapter 2. There is also a litany of them that influence the

expression of psychiatric disorders. Behrendt (2006) for example enumerated studies that show that schizophrenia can be caused by any of the following environmental circumstances: parental exposure to viral infections, obstetric complications, childhood brain disease, minor physical abnormalities, cognitive impairment and developmental delays. Roff (1994) proposes explaining phenotypic dimorphism under varying environmental conditions by assuming a continuum of additive genes with a threshold that is not fixed, but varies with the varying environment. In what he calls the Environmental Threshold (ET) model, the predisposing environment elicits the phenotypic expression of the trait by lowering the threshold additive genetic requirement. Thus, there are many traits that will either express phenotypically or remain latent contingent upon these factors. Often, very few members of a population are exposed to these predisposing co-factors and consequently only a proportion of individuals who carry such alleles in a population may actually express them phenotypically (Monroe and Simons 1991).

Plasticity/latency maintains maladaptive traits in populations by a very simple logic. According to Darwin (1859, 108) "Variations neither useful nor injurious would not be affected by natural selection." By the same token, alleles that fail to express phenotypically are not affected by natural selection (Dawkins, 1982). If circumstances that influence the phenotypic expression of an allele affect only a proportion of individuals who bear the allele in every generation, only

the affected proportion would express the trait phenotypically and consequently be affected by selection. The allele would be unaffected by selection in the rest of its carriers in the population in whom it is latent. Thus a maladaptive trait can potentially persist in a population if there exists any contingency that limits its phenotypic expression in such a way that only a segment of individuals who carry the trait actually express it phenotypically. That principle has been effective in maintaining even some harmful Mendelian traits such as sickle cell anemia and cystic fibrosis in populations. Since the heterozygous carriers of such alleles suffer no fitness impairment (and in fact may in some cases have enhanced fitness), they reproduce normally and therefore serve as a reservoir from which the harmful phenotypic trait is generated. In the discussion of altruism in chapter 2, we noted that almost any random member of such populations has the capacity to express the altruistic trait, suggesting that alleles that cause the trait are also present but latent in the non-altruists. That is indeed what guarantees the persistence of altruism in populations rather than the specious theories that hold sway today.

IV. The Polygenic Plasticity Effect

A case for *phenotypic plasticity* as the possible underlying explanation for the persistence of maladaptive complex traits hinges on two key features about such traits. The first is the heavy influence of the environment on their expression (Monroe and Simons 1991; Behrendt 2006; Allen and Badcock 2006)).

Environmental contingency is invariably indicative of plasticity in the phenotypic expression of an allele. The second is the strong consensus among psychiatric geneticists, for example, that psychiatric disorders - and in fact, behavioural traits in genera l- are generally under polygenic control (Keller and Miller, 2006; Kendler, 2005; Kendler and Greenspan, 2006). In addition to the environmental contingencies, the polygenic nature of such traits also introduces other contingencies such as additive genetic thresholds, epistatic constraints, and the requirement of certain regulatory genes. Thus both the environmental influence and the polygenic character of such complex traits contribute to depress their *penetrance* (i.e. the proportion of carriers of an allele who actually express the phenotype).

Let us here introduce the term *phenotypic penetrance* (PP) to mean the proportion of carriers of a particular genetic trait in a population who actually express that trait phenotypically. If natural selection only affects traits that are expressed phenotypically, then the lower the PP of a trait, the less would be the impact of selection on the trait. If one of the environmental conditions that would allow a mutant to express phenotypically requires the presence of another individual who has the mutation, but he is at this time the sole incipient mutant, then he would not express that maladaptive phenotype and will not be selected out. So such an individual gets to reproduce normally and thus increasing the number of copies of the mutation in the population. As we shall see later, extant

evolutionary models of altruism miss this point and assume the incipient altruistic mutant to be of lower fitness and thus unable to increase in frequency.

The *latency effect* is also generally overlooked in models of complex traits, and it has instead been suggested in those models that alleles for such traits persist because they have generally mild effects (Pritchard 2001; Keller and Miller, 2006). Is it really the case that severe mental disorders for example are not persistent and only mild ones are? Or is it being suggested that the alleles with mild deleterious effects simply contribute additively to produce the severely deleterious phenotypes? Under either scenario alleles with mild deleterious effects could not escape elimination by natural selection. As Dawkins (1976) argued, even a good gene can be eliminated from a population if it is in a team with a group of genes that perform poorly overall. Thus a group of alleles, each of which contributes in a small way to express an overall deleterious phenotype would all be eliminated jointly and severally by the action of natural selection on that phenotype. It appears rather more likely that alleles that predispose individuals to psychiatric disorders persists in populations not because they have mild effects which they express all the time, as Keller and Miller suggest, but rather because such alleles have no phenotypic effect most of the time due to constraints imposed by the environment and the other genetic factors discussed above.

The weakness of specific alleles as reliable predictors of specific mental disorders (Kendler, 2005; Pritchard, 2001; Risch, 1990) suggests rampant latency in psychiatric disorder alleles. As Monroe and Simons (1991) note regarding depression, there is always a proportion of the predisposed that never manifest the disorder. If the environmental trigger is rare while the genetic predisposition is common as Monroe and Simons (1991) further note for depression, latency could be particularly high.

In elaborating on some meta-analysis, Keller and Miller (2006) estimate, based on the odds ratios of some specific susceptibility alleles for schizophrenia, that given 1,000 people with the allele and 1,000 without it, 11 people in the first group and 10 in the second group would probably develop schizophrenia. That suggests a nearly 99% latency rate among such alleles. I think that however deleterious an allele is, it is going to be around for a very long time if it does not express the deleterious phenotype in 99% of the individuals who carry it. The statistic also suggests that there is almost equal probability of developing the disease between an individual with a particular susceptibility allele and one without it. We can infer from such data that there are many other susceptibility alleles, which are collectively extremely pervasive in the population. Consequently, for any two cases of schizophrenia, different alleles (or combinations of alleles) may be at play. This phenomenon is conveyed most effectively by Pritchard's (2001, 125) statement that "a defining feature of

98

complex phenotypes is that no single locus contains alleles that are necessary or sufficient for disease" (or I would say, for the expression of altruism).

That explains why it is difficult to draw a genotypic dichotomy between individuals who express such traits phenotypically and those who do not, in spite of the underlying genetic predispositions. As Hall (2010, 67) reports, "the very definition of a gene … is now vexed by multiple layers of complexity. What was once assumed to be a straight forward one-way, point-to-point relation between genes and traits has now become the "genotype-phenotype problem," where knowing the protein-coding sequence of DNA tells only part of how a trait comes to be." Thus, extant evolutionary models of altruism have continuing difficulties because they assume an underlying genotypic dichotomy between the altruistic and selfish phenotypes, whereas *the resolution of the paradox of altruism lies in the very fact that a selfish individual is not necessarily without the "altruistic gene."*

Given what we know now, Darwin's concern was ill-founded when he suggested that "he (the altruist) who was ready to sacrifice his life as many a savage has been, rather than betray his comrades, would often have no offspring to inherit his noble nature" (*The Descent of Man* 1871, 499). We now know that the "noble nature" of the altruist is inherited not only from the altruists, as extant evolutionary models suppose, but also from the non-altruists in the population and therefore continues to be present in any generation even when the altruists

leave no descendants. The priestly order of Roman Catholics has foregone procreation for over 60 generations dating back to the apostle Paul in the first century AD. If Darwin found sufficient reason not to be puzzled by their continued presence in human society (i.e. that the behaviour is not hereditary), those same reasons should be sufficient why worker castes persist among the *Hymenoptera* and ultimately why altruists persist among social animals (including humans).

It is estimated that thousands of different mutations could be contributing to psychiatric disorder susceptibility. Coupled with the high latency rates it is quite possible that almost every single individual of the human population will have some amount of these mutations (Keller and Miller 2006) - enough I would say to express the phenotype given certain environments. Monroe and Simons (1991) suggest for example, that an extreme environmental circumstance can cause depression in most humans. That is consistent with the observation earlier in this work that in the social animals, each individual organism has the genetic capacity to express the altruistic phenotype under the appropriate environmental cues. I think any human being would run into a burning building to save another if environmentally conditioned. I do not think it is a genetic predisposition that is limited to a few choice individuals in the human population.

Conceptually, genes direct every morphological and behavioural phenotype of the organism. As we have seen in this discussion however,

sometimes the environment plays a role in triggering or suppressing the action of certain genes. In such situations, the presence of the gene is no longer sufficient for the expression of the phenotype. An additional factor with complex traits such as behaviour is that there are often many different predisposing genetic factors involved. In such cases not only is any particular susceptibility allele not sufficient; none is particularly necessary for the expression of the phenotype. Thus it becomes impossible even with traits that exhibit some measure of heritability, to draw a straight genetic line of distinction between individuals who express the trait and those who do not. That is what imparts what I would call the *nature-nurture duality* on such complex traits.

The nature-nurture duality is a trademark effect of *Polygenic Plasticity*, which describes situations in which a trait is influenced by alleles at multiple loci, each of which has a plastic/conditional phenotypic expression. In such situations, there are often some indications of a genetic predisposition, but equally strong suggestions of environmental causes. Thus anybody can fish out the evidence necessary to support his/her biased position, leading to the never-ending debates. Beyond altruism and psychiatric disorders, which are showcased in detail here, there are other complex human behavioural conditions such as sociopathy, psychopathy, intelligence, homosexuality etc., for which equally strong arguments can be made for both genetic and environmental predispositions. Polygenic plasticity may be the ultimate explanation for their

nature-nurture duality. Most importantly however, it must be stated in fulfillment of the objective of this article, that it is also polygenic plasticity that ultimately facilitates the evolution and persistence of such traits in situations where they are maladaptive, such as in altruism.

V. Modelling the Evolution of Altruism

Having established that the altruistic and selfish traits are alternative phenotypic expressions of a single genotype (see chapter 2), the challenge now is to adapt our population genetic models to accommodate such deviations from the straight genotype to phenotype mapping of Mendelian genetics. This is the consideration we have to bear in mind all along as we model the probable path of the evolution of eusociality by natural selection. Unlike Hamilton (1964a and 1964b) who devotes two long articles and circuitous mathematics to explain how eusociality might evolve, for Haldane (1932) it is a straight-forward affair for which he needs just a single short phrase to explain. He writes (1932, 208):

> "In a beehive the workers and young queens are samples of the same genotypes, so any form of behaviour in the former (however suicidal it may be) which is of advantage to the hive will promote the survival of the latter, and thus tend to spread throughout the species."

We can see why it is so straightforward for Haldane and circuitous for Hamilton. Haldane builds upon the empirical fact that the two castes are of the same genotype, which means that a fitness gain for one is a direct fitness gain for the other, and we need not scale that by any pedigree factor. Hamilton on the other

hand approaches the problem from theoretical principles, and ignores this key empirical fact. Consequently, instead of viewing the workers and queens as genotypic clones with respect to altruism, he deploys the coefficient of relatedness (r) as the crucible for the disbursement of the fitness gains that accrue from an altruistic behaviour. This process culminates in his much heralded *inclusive fitness* theory, whose central claims have all now been refuted by empirical research (refer to chapter one for this discussion).

Between Haldane and Hamilton on eusociality, as shown above, we see the difference between explaining the empirical data on the one hand and theorizing from a collection of theoretical assumptions on the other. That kind of strict adherence to the empirical data was the basis for Darwin's great theoretical success. Even though Darwin was quite ignorant about the underlying genetics of the eusocial trait, he gave a rather decent account of how it might evolve through a gradual accumulation of profitable variations. I should also point out that the angle and concern from which he approached the problem was not that of behaviour (which we call altruism), but rather of morphology, given as he writes, that "we have an insect differing greatly from its parents, yet absolutely sterile, so that it could never have transmitted successively acquired modifications of structure or instinct to its progeny" (1859, 353). For all his limited knowledge of genetics, Darwin's explanations are never too far off the mark on any of these issues because on his theoretical journey, he sought an empirical foundation to

set every step. Every argument he made and every explanation he proposed was based on or supported by some empirical observation.

First Darwin explains why he is not puzzled by the distinct morphology of the neuter ants. But I had to fight to dismiss my own amazement that he saw it as a case of phenotypic plasticity, even though he did not have the benefit of knowledge of the underlying genetics. Again, being the empiricist that he was, he looked into nature for clues to the puzzle, and then made the following observation: "Let it be remembered that we have innumerable instances, both in our domestic productions and in those in the state of nature, of all sorts of differences of inherited structure which are correlated with certain ages and with either sex" (1959, 353). He goes on to cite examples of phenotypes that are expressed only under certain circumstances and environments, and then he concludes: "Hence, I can see no great difficulty in any character becoming correlated with the sterile condition of certain members of insect communities" (1859, 353). So for Darwin, just as certain phenotypes, even though available to all members of a species, are expressed only at a certain age, a given sex or some other circumstance, so is it that the heritable trait that underlies the distinct morphological features and instincts of the neuter ants are actually present in all members of the species, but are expressed only under the neuter condition. Of course, we now know that the neuter condition itself is a species-wide trait that is

expressed only by those individuals who experienced certain environmental circumstances such as diet, as in the honeybee for example.

Darwin points out that the real difficulty lies in explaining how such plastic phenotypes "could have been slowly accumulated by natural selection." Even that difficulty, Darwin thinks, disappears if we remember, as he writes, "that selection may be applied to the family as well as the individual" (1859, 354). He concludes that if it is advantageous that a community has the tendency to produce some members that are sterile with structural modifications, then the fertile males and females will flourish and transmit "to their fertile offspring a tendency to produce sterile members with the same modifications" (1859, 354). Once again Darwin is bang on here, and his solution is simple and unambiguous and in fact consistent with and actually stems from the empirical reasons for which he saw no puzzle in the peculiar morphology of the sterile caste. This is one of the brilliant insights of Darwin. As observes so astutely, the tendency or genetic capacity (trait) to produce neutered (altruistic) individuals is carried in the fertile individuals and transmitted amongst them from generation to generation. If a mutation occurs in a queen that makes her produce offspring, some of whom "are capable of work but incapable of reproduction," and if that is advantageous to the community, that community will spread. So Darwin provides a two-step solution to the evolutionary problem of neuter insects. First he recognizes that it involves a plastic trait common to all members, which expresses phenotypically only in

certain of the members who encounter certain circumstances. Secondly, if when the trait so expresses, it be advantageous to the community, the community will flourish and spread. Darwin supposes that colonies with more neutered workers will fare better than colonies with fewer workers, and consequently, the latter spread.

Hence, Darwin explains eusociality or reproductive division of labour as selection between communities rather than within communities. In general terms, it is the right characterization. For the task is to explain how the eusocial reproductive system may have evolved. At the fundamental level we have to suppose that the eusocial organism evolved from an ancestor that was not eusocial. For the eusocial trait to increase in the ancestral community, it has to have a fitness that is above the population average. Now, what I am going to do next is very critical. This is where the modern problem of altruism lies, which I hope to clarify once and for all. In the basic model of population genetics, the following representation is given for genotypes and their frequencies for a single locus trait (Halliburton, 2004, 133).

Table 1. Allele Frequencies Chart

Genotype	A_1A_1	A_1A_2	A_2A_2
Frequency	P^2	$2pq$	q^2
Fitness	w_{11}	w_{12}	w_{22}

The current practice in modelling the evolution of altruism is to assume the contrasting alleles (A_1 and A_2 as in the table above) to represent the altruistic and

selfish traits respectively, and that is a grave mistake, which has kept the problem of altruism intractable. For under such a scenario, a barrier to the evolution of altruism from a single mutation is erected. If altruism is to evolve from non-altruists, then we will have to imagine a mutation that causes altruism to occur in a single individual (FM, in figure 3) in a population (SP2) of selfish individuals. There is the concern under current models that because it carries the altruistic gene, this mutant will have a relative fitness that is lower than the population average in this population of selfish individuals. This concern is due to the thinking that the altruistic mutant will be behaving altruistically and incurring fitness costs, while the non-altruistic population would be reaping the rewards of that altruism. In such a scenario, our population genetic models predict that the altruistic gene should not increase in frequency, hence, a barrier to the evolution of altruism.

We should note that no current model of altruism demonstrates convincingly how this evolutionary barrier may be scaled. Inclusive fitness predicates upon the condition that the bulk of the benefits of altruism fall on close relatives who have a higher probability of carrying the altruistic gene. However, at this point (P2 in figure 3) in the evolution of the gene, only the mutant (FM) and no other individual in the selfish population carries the altruistic gene. Hamilton's rule (rb>c) is therefore useless at this point. Even if some individual is 99.9% related to the altruist (r = 0.999), he still does not have the altruistic mutation, and

107

the altruistic gene does not benefit. So Hamilton's inclusive fitness rule, that the altruistic trait would increase in frequency on condition that *the benefit of the altruistic act to copies of the altruistic gene in relatives is greater than the cost to the altruist* breaks down here because there is no altruistic gene in any other individual in the population, relative or not.

The group selection model touts the increased fitness of the group (that contain altruists) as a collective, without regard to which individuals benefit – relatives or not. So the group selectionists may contend that their model can proceed beyond this point, since the members of the group still benefit from the altruistic act, and they may thus gain some fitness advantage over neighbouring groups which did not have the benefit generated by the altruist. In fact, this happens to be the *modus operandi* of group selection, and it was what Darwin argued when he suggested that groups with more neuters or altruists would
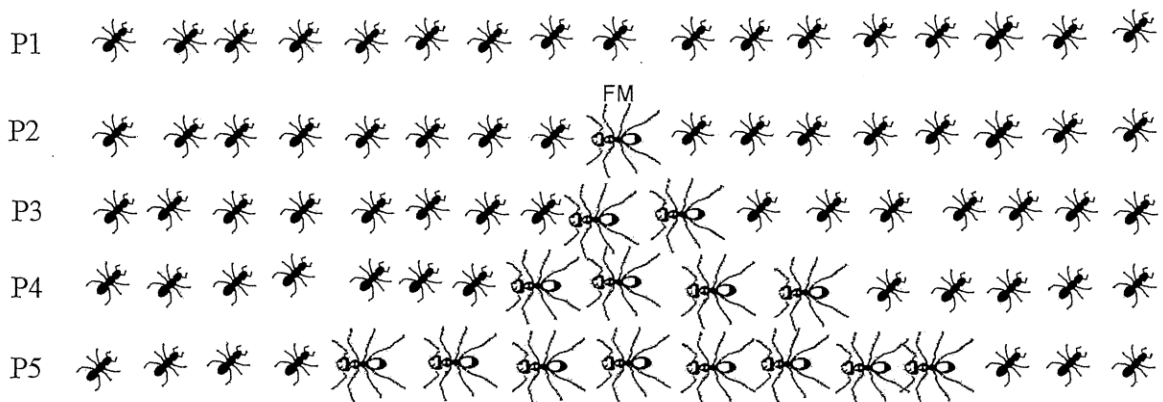

Figure 3. The current (ASA) model of a social community and how it evolves. The indigenous population (P1) is selfish. An altruistic mutant (FM) is born and is expected to spread in the population against the selection gradient. The final population (P5) is an admixture of altruistic and the original selfish individuals.

outcompete groups with fewer of them or none at all. In figure 3 for example, group strength would increase as the frequency of altruists increase from SP2 to SP5. Group selection however, does not have a credible account as to how the frequency of altruists in a group can increase in that manner. In fact, it does not even have a viable solution to the theoretically predicted attenuation of the altruistic allele in populations. What has been proposed (Haldane, 1932; Sober and Wilson 1998; Sober, 2000) is for such large populations in which the frequency of altruists is declining to periodically break up and reconstitute into smaller groups, some of which will have the threshold concentrations of altruists. However, other than a hypothetical scenario suggested by Maynard Smith (1964), there is a paucity of evidence for this in nature, with claims made only in contested anecdotal examples (such as *Myxoma* – Wilson, 2004).

If we are modeling the evolution of a eusocial population under current (ASA) assumptions, the first altruistic mutant (FM, in figure 3) would have to be a non-reproductive worker. She would be the only individual with the altruistic allele, but she would leave no descendants. So eusociality cannot proceed beyond the first mutant under any model that supposes altruism and selfishness to be contrasting genetic alleles (i.e. ASA). And if we deny ever supposing genetic allelomorphism of altruism and selfishness in the eusocial organisms, then we will have to accept Haldane's (1932) terse but correct account of

eusociality and reject Hamilton's (1964) convoluted inclusive fitness account as irrelevant to eusocial altruism.

The problems of extant altruism models are manifold. What I will do here is illustrate their underlying basis. Let me start by doling out some charity, and granting that against whatever odds, the first mutant altruistic individual survived and successfully reproduced, albeit fewer offspring than the rest of the selfish population. Now the inclusive fitness model might have the wings necessary to fly; that is, if we grant further, as the model requires, that the altruistic siblings and their mother cooperate or direct their altruistic behaviour towards each other. If these conditions are satisfied, then the altruistic individuals might flourish and grow in the population.

Now let us examine more closely these conditions under which an altruistic gene might increase in frequency through the principle of inclusive fitness and the implication of that on the genetics of altruism. The basic requirement under inclusive fitness or kin selection is that the altruists be able to discriminate in extending altruism. This condition by the way is also required under reciprocal altruism. The capacity to discriminate requires more than being simply altruistic. It requires the capacity to judge when to be altruistic and when not to be. That additional function suggests pleiotropy. A further implication is that there would be situations in which the altruist would withhold altruism or behave selfishly in order to discriminate against non-kin. This suggests plasticity, given

that the gene will express altruism or selfishness depending upon the circumstance. It is clear then that the kind of altruistic trait required under the inclusive fitness model, and for that matter any other model that requires conditional altruism, is a very complex trait that is both pleiotropic and plastic, and which may be controlled from several loci.

The point to note is that under the conditional altruism models, the so-called altruistic gene does allow, or make its bearers to behave both altruistically and selfishly. In a model that was to unify inclusive fitness and reciprocal altruism, Queller, (1985, 367) states the following evolutionarily stable strategy (ESS): "be altruistic to an untested partner, otherwise, behave towards the partner as he behaved towards you." This means that if individual x behaves selfishly towards individual y who is an altruist, y ought to be able to behave selfishly toward x in response, even though y carries the altruistic gene. In another Prisoner's Dilemma –ESS analysis, Fehr and Fischbacher (2003, 787) observe that "it is not possible to infer the absence of altruistic individuals from a situation in which we observe little cooperation. If strong reciprocators believe that no one else will cooperate, they will also not cooperate." If so, what then is the genetic difference between the individual we would call an altruist and the individual we would call selfish in a population, since the conditional altruism of kin selection, reciprocal altruism and ESS models requires that an individual be able to express both traits? So as you can see, by asserting conditionality in our

solution to the problem, we blur the genetic distinction we originally supposed between the two phenotypes, and which was the basis of the paradox; because one allele or genotype was supposed to cause a phenotype that lowered its fitness.

So the initial genetic distinction we supposed between the new altruistic mutant and the selfish population; and the general supposition in our models that the altruistic and selfish traits are allelomorphs become problematic. If for example we see a vampire bat refusing to give blood to a starving solicitor, would that be one carrying a selfish allele, or would it be one with the altruistic allele which is exercising conditional altruism? Are we to suppose then that there are two kinds of selfish individual in a social population- the one who is selfish on account of the selfish allele and the other who is selfish on account of conditional altruism, even though it carries the altruistic allele? This problem arises from supposing the altruistic and selfish traits to be allelomorphs and at the same time supposing that an individual that is carrying the altruistic allele can also conditionally express selfishness, the contrasting phenotype.

On the contrary, as I shall elaborate here, the allele we designate as altruistic is actually *a plastic pleiotropic sociality or cooperative genotype that can express both altruism and selfishness given the circumstances*. The other genotype, which is the original allele in the population, represents *the non-social, non-cooperative or the solitary lifestyle*. The italicized phrases describe the two

kinds of individual involved in the evolution of sociality. Extant models bungle fatally here in modeling it rather as an altruistic allele evolving from a selfish population, and making no distinction between the selfish individual in an evolved social community and the non-social individuals of the antecedent population.

Thus, if we are modeling the evolution of eusocial behaviour for example, we want to trace the steps that led to that kind of reproductive arrangement, as opposed to its antecedent solitary life-style. So before the emergence of the eusocial organism, there was a solitary organism that reproduced other solitary organisms, each of which went on to reproduce other solitary organisms. Current models call this pre-social organism "selfish," and take it to be one and the same as the selfish counterpart to the altruist in the evolved social population; and that I say is a blunder. The models suppose that the selfish baboon that would not help another in a social community is expressing the same genotype as the "selfish" pre-social baboon that was solitary. But that is simply not the case, especially if we suppose in our models that this selfish baboon could actually be carrying the altruistic trait and only being conditionally selfish.

If eusociality evolved in a Darwinian fashion, then we have to suppose that at some point a mutation occurred in a gem cell of one organism in a solitary population (SP1 in figure 4), which imparted onto the resulting offspring (FM) the tendency to stay in the nest and help the mother, given certain environmental circumstances. Most of the required environmental circumstances that would

trigger the altruistic or eusocial behaviour in this first mutant offspring (FM) would be absent at this point, because for the most part, such triggers are in themselves behaviours of other members of the eusocial group or altruistic society, which are coded by the pleiotropic action if the eusocial gene. Let us remember that when we talk of a eusocial organism we do not refer only to the altruistic members. The queen bee is also a eusocial organism just as the worker is. So is the alpha female in a naked mole rat colony. Such so-called "selfish members" of the social groups are also carriers of the eusocial gene, even though it expresses in them a different phenotype because they experience different life circumstances. The phenotypic role these "selfish members" play is often what induces the altruistic behaviour in the other members and we could not expect the individuals in the pre-social population from which the altruists evolved to be able to play those roles. A common example is coercion and preventive action which causes individuals to assume the altruistic, subordinate or non-reproductive role (Wensellers and Ratnieks, 2006; Queller, 2006; Young and Clutton-Brock, 2006; Emlen and Wrege, 1992). In the honeybee it is the diet that the workers feed the larvae that induces the altruistic phenotype in them. Then in reciprocal altruism the previous behaviour of a solicitor determines whether or not it can induce an altruistic response from a subject.
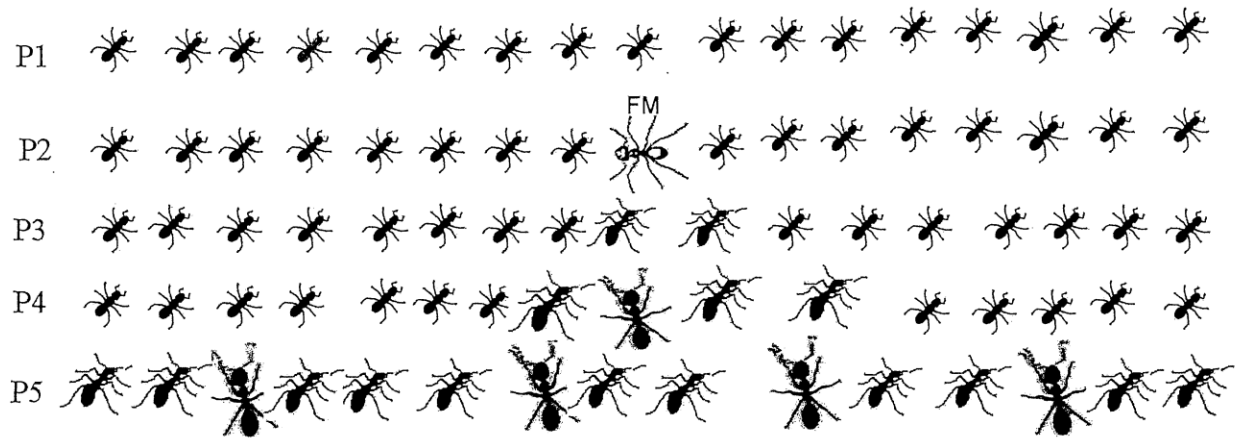
Figure 4. A schematic of the evolution of eusociality. P1 is the solitary population. P2 is the generation in which there arises the first eusocial mutant (FM). Morphologically differentiated reproductive individuals may appear in the third mutant generation (P3). Note that in the mature social population (P5). The sociality allele(s) is fixed and there are no more solitary or non- social individuals in the now separate species.

So rather than a simple altruistic trait arising in a selfish population, if we recognize instead, a pleiotropic and plastic sociality genotype arising from a non-social population, the barrier to its evolution disappears. For the conditions that elicit the altruistic behaviour are generated by other organisms with the sociality allele in a way that ensures a net fitness gain for the gene for social behaviour, and in the absence of the right social individuals to induce the altruistic behaviour, altruism is simply not deployed, and consequently, no fitness cost to the mutant altruist. The new altruistic genotype remains latent until the right circumstances that are favourable to its survival induce it to express phenotypically. The latency of the altruistic genotype is akin to seed dormancy. The seeds of many plants remain dormant until a specific set of circumstances obtain. Such circumstances are evolutionarily calibrated to coincide with the best

chances of survival of the ensuing plant when dormancy is thus broken. It is in the same way that circumstances that trigger the expression of the altruistic allele is evolutionarily calibrated to ensure its survival when expressed.

So coming back to our particular analysis, the first mutant social organism may not encounter the appropriate circumstance that would trigger the expression of the altruistic genotype it carries. Even if the mutant offspring were urged on by the eusocial instinct to remain in the parent's nest in spite of all the signals that it needs to strike out on its own, the parent, lacking the eusocial genotype, may not understand the gesture and may harass the offspring until it is forcefully weaned off. It is in the second generation of mutant offspring that we may see some eusocial phenotypes actually expressed. The parent to this generation is the first mutant, so with offspring of its own now, it will have the opportunity to phenotypically express eusocial parenting. In a naked mole rat family for example, the offspring will have the eusocial gene that urges them to stay and help the mother, and the mother will also have the same eusocial gene that will, in her circumstance and status, make her produce the concomitant pheromones and also engage in the behaviours that will prevent her offspring who remain in the nest from reproducing.

In the honeybee at this stage, the mother may raise her first few broods all as workers, since a workforce needs to be built. The first brood of workers may be numerous enough to relieve the mother of the foraging and nursing duties so

that she can dedicate herself to reproduction, i.e. become a queen after initially being a worker that raised the first brood. When the workers are in sufficient numbers, they may begin to nourish some of the larvae into queens by feeding them "royal jelly." Since this first eusocial mother never got the queen nutrition in her larval stage, she is probably not much different from the workers morphologically, and might also not exhibit the full retinue of queen behaviours. It is conceivable that one of the properly reared queens out of her daughters would replace her as queen.

In modelling the evolution of eusociality therefore, the eusocial behaviour as a whole (which comprises both altruistic and selfish behaviours) should be viewed against its alternative – the solitary lifestyle. Instead, in extant evolutionary models, we see the alternative phenotypes within the eusocial system being pitted against each other as competing evolutionary alternatives. However, the evolutionary alternative or allelic alternative to the non-reproductive worker is not the reproductive queen, because they are both products of the same genotype, and their phenotypic differences result from differences in environment rather than genotype. They are both phenotypes of one genetic constitution which is an evolutionary departure from the genetic constitution that codes for solitary living. So we need to consider the eusocial gene as a gene (more appropriately, a series of mutations) that adapts an individual to eusocial life. That means it is a genetic condition that can express either the worker or

queen phenotype in an individual depending on environmental factors (and or additional genetic factors); and in the preceding paragraph, I have given a simplified account of how that might evolve.

So, how might a cooperative gene evolve? Once again we have to start with a community of selfish individuals. Suppose there occurs a mutation in a gem cell of one of the selfish individuals. This mutation is a pleiotropic flexible gene, which I have argued the altruistic gene is. For that reason I will here on refer to what we commonly call the "altruistic gene" as the sociality gene; because it is a flexible gene that expresses not only altruism, but selfishness as well. By the same reasoning I will refer to the contrasting allele to that as the asocial gene, which is that possessed by the non-social, non-cooperating members of the population from which a mutation may cause a social allele to arise. Then as I suggested earlier on, sociality actually is a complex trait that may have resulted from a series of mutations at multiple loci. However, for the sake of simplicity, I have described the evolutionary process here using a single mutation, but the process can be iterated for a series of mutations, each of which adds one of the components of the complex trait.

I turn now to the evolution of sociality in the non-eusocial social organisms such as humans, baboons, vampire bats, lions etc. I pointed out at the beginning of this section how unlike Hamilton, Haldane had a simpler and more accurate account of eusocial altruism because he recognized the genetic commonality of

the selfish and altruistic members with respect to that trait. However, when Haldane moves on to model non-eusocial altruism, he commits the same error as most other scholars by modeling the altruistic and selfish phenotypes as allelomorphs. In chapter two I explained in detail why the altruistic and selfish traits are not allelomorphs, but rather plastic expressions of a single genotype. And like Hamilton on eusociality, this blunder takes Haldane's account of altruism through some circuitous mathematics and arguments, and in the end he leaves his account hanging on an unlikely set of ifs, none of which he seems to be confident about (Haldane, 1932, 208-210).

Now, in modeling the evolution of a sociality trait, we have to consider a gene (or a cumulative series of mutations) which endows an individual with the capacity to engage in social behaviour, such as reciprocal food sharing, cooperative hunting, mutual grooming and the other social interactions with other similarly endowed individuals on one hand as against individuals who lack that genetic capacity and therefore lead independent/solitary lifestyles on the other hand. So again, let us start with an asocial population such as a population of wildebeest in Tanzania's Serengeti, in which every individual acts in its own direct interest. Or we could imagine a solitary leopard hunting by itself and reproducing generation after generation. Let us suppose that in a non-social population such in the above examples a mutant is born that is capable of some kind of social

behaviour, such as food sharing as in vampire bats, grooming and aiding as in baboons, or nest helping as in the white-fronted bee-eater.

Again, as in the eusocial example earlier, this first mutant might not actually express any social behaviour because the environment that elicits that behaviour might not be present, since no other individual in the population at this stage has the genetic capacity to cooperate or reciprocate. Asocial animals largely have no interaction with each other except for mating. If the mutant has the grooming instinct for example, it cannot approach any of the non-mutants because its actions would be misunderstood or seen as bizarre. If it has a food sharing instinct and is willing to share its food, individuals in the population do not expect it and would not understand any gesture that would suggest the altruist is inviting them to share. Bigger individuals may approach, but only to seize, rather than share the food. If the altruistic mutant joins a fight to help an individual he considers to be an ally, he might get attacked by the supposed friend. So in a nutshell, the asocial population that surrounds this lone pioneer altruist will not provide the environment necessary for the altruist to dispense altruism. So from this and the earlier case of eusociality, we see that the phenotypic expression of the altruistic trait requires a social environment that is generated by other individuals who carry the sociality trait. In the absence of the appropriate social circumstances, the trait remains latent, and consequently not affected by selection. So the fitness barrier to the evolution of altruism does not occur here.

120

Extant models face this barrier because they suppose erroneously that the altruistic and selfish traits are allelomorphs. That supposition however, is not consistent with the empirical data (see chapter two). It is also logically and analytically unsound, as I showed earlier in this section.

Social interaction becomes possible in the second generation of mutants, who will be the offspring of the first lone mutant. This is because there would now be a few individuals who have the genetic capacity to engage in social interaction. The small group of mutant social organisms will therefore stick together and cooperate in many tasks of life. If the cooperation and mutual aid in which these social mutants engage raise their fitness above that of the asocial members of the population, they will flourish and their ranks will swell. So as you can see, the scenario envisioned in extant models, of a small vulnerable population of altruists being exploited by a selfish population does not really occur in the evolution of sociality. That wrong picture emanates from the wrong genotype/phenotype relations those models suppose.

VI. A Possible Objection

One of the remedies proposed to deal with some of the difficulties faced by the models that suppose altruism and selfishness to be allelomorphs (ASA) is that altruists might be able to recognize some feature in other altruists which could serve as the basis for extending to them the benefit of altruism. Dawkins (1976) used the analogy of a characteristic "green beard." Dawkins then went on to

suggest that it is conceivable that some selfish individuals could develop a mutation that would make them able to mimic the "green beard" even though they do not carry the altruistic gene, and this would allow them to exploit the altruists. Some prey animals for example have successfully used mimicry to fool predators to think they are poisonous.

This objection may indeed cripple the ASA models largely as a result of their own inherent contradictions. For the ASP model on the other hand, I will say this: In the account of the ASP model, I explained that the conditions that trigger the altruistic behaviour arise from the actions or behaviours that the sociality allele codes. It is not simply a morphological feature as a signal. Thus the conditions that trigger conditional altruism under the ASP models are behaviours that are intricately interwoven with the altruistic trait. For example, a vampire bat shares blood with only those who have given blood before or a baboon helping only those who have helped before. So in order to mimic this kind of "green beard," you will have to have the altruistic gene.

Suppose in the case of honey bees a mutant comes up, who is purely selfish and does not carry the plastic social (altruistic) allele. She can thus only be a queen. Her female offspring will also not carry the flexible social allele and would all be queen bound as well. What would follow is that the worker population of such a colony will rapidly decline and the community will collapse. For if the queen in such a colony reproduces only selfish queens then it would

soon be "all chiefs and no Indians" and that will not be sustainable. The fact that we still have eusociality today after tens of millions of years suggests that they must be immune to that kind of invasive strategy; or at least for now. Perhaps evolutionary history has just as many extinct social strategies as there are morphological strategies in the fossil record.

In the social amoeba, *Dictyostelium discoideum*, Foster *et al.* (2004) report that a common plastic genotype (*dimA*) alternatively expresses stalk cells, which do not reproduce, and spore cells, which produce and disseminate spores. The trigger for stalk formation is a protein (DIF-1) secreted by other cells in the aggregation. With such an arrangement, it is conceivable that a mutant could arise that does not respond to the DIF-1 molecule and would thus have a fitness advantage by committing all of its cells to spores and investing none in stalks. In a laboratory experiment, Foster *et al.* discovered that mutants with disrupted *dimA* gene did indeed ignore the DIF-1 trigger to form stalk cells, but they were also excluded from the spores in the *dimA* fruiting bodies. So here we have a case of plastic expression of an altruistic and a selfish phenotype by a single genotype, which has also evolved a very effective mechanism to exclude cheats.

On the same social amoeba, *Dictyostelium discoideum*, Strassmann *et al.* (2000) reported that fruiting bodies were prone to exploitation by cheats. However, they just identified any random DNA loci that differed between two separate clones that were then experimentally mixed. The proportions of cells

from the different clones were then evaluated in the different regions of the slug, and then subsequently, the stalk and spores of the fruiting body. However, as we see in the work of Foster *et al.* (2004), the criterion that matters is the possession of the *dimA* gene. It is the one that maintains the stalk forming altruism, and it does not matter what proportion of what clone is in the spores, as long as they all possess the altruistic *dimA* gene, that gene is guaranteed to be transmitted to subsequent generations, and the maintenance of the non-reproducing stalk forming altruism is not threatened.

So the evolutionary stability of a plastic gene that produces mixed populations of advantageous and disadvantageous phenotypes depends upon the efficacy of the mechanism it has pleiotropically evolved to regulate the alternative phenotype, which if well optimized, should also exclude cheats. If its exclusionary mechanism has loopholes that cheats might exploit, then its evolutionary stability is compromised. Social arrangements do not leave fossilized remains, but it is quite possible that there are many failed social arrangements in evolutionary history, just as there are many extinct anatomical and morphological phenotypes. In the case of *D. discoideum*, the condition for inclusion in the reproductive body is the possession of the *dimA* gene; and this effectively ensures that every single cell given the opportunity to reproduce will produce spores that contain the *dimA* gene. It is hard to imagine an easy way

around that. For that reason, that behaviour survives in that organism even if only for now, and for as long as it maintains a step ahead of the cheats.

VII. Summary

The objective of this final chapter was to give a coherent and satisfactory alternative account of the evolution of altruism that would be free of the difficulties faced by extant models. The new model hinges upon phenotypic plasticity as the genetic contrivance that fosters persistence of altruism and other maladaptive traits. It was established in chapter two that empirical evidence from the best known cases of altruism indicate unequivocally that certain specific environmental cues can elicit the altruistic phenotype in any random individual in such social populations. Further insight from psychiatric genetics which I presented in this chapter reveals that alleles that predispose individuals to such complex behavioural traits are numerous and ubiquitous in populations. The polygenic character of the traits introduces some constraints to their phenotypic expression which, in addition to environmental constraints result in very high rates of latency among the alleles of such complex traits. Deleterious alleles that are latent stay beneath the radar of natural selection and thus remain a guaranteed source of the maladaptive phenotypes. Thus environmental contingency of a trait is indicative of some measure of latency of the alleles that cause it and consequently a disruption of the normal course of natural selection. I showed these to be the circumstances that allow certain maladaptive phenotypes to evolve.

CONCLUSION

I have shown in this discussion that the bane of extant models is their failure to recognize the role the peculiar features of complex traits (discussed in chapter three) play in their evolution. The fundamental mistake of extant models is that they suppose altruism and selfishness to be genetic allelomorphs, and that error is the source of the paradox of altruism and the difficulty of modeling its evolution. The models for example pit food sharing against refusal to share food in vampire bats as competing genotypes, whereas they are behaviours that perhaps every individual in such communities can express depending on the social circumstance at the time. This thesis rejected the existence of any allele or genotype dedicated solely to the expression of the altruistic phenotype. Rather, what exists and whose evolution this thesis modeled is a plastic and pleiotropic sociality trait that expresses both altruism and selfishness under different circumstances.

The distinction of the phenotypic plasticity model, which is key to its explanatory success, is that it supposes the presence of the altruistic gene in both the individuals in the population who express altruism (the altruists) and those who do not (the selfish individuals). When it is understood this way, which in fact is the view supported by empirical observations, there should be no concern about altruists losing fitness to selfish individuals, and no need

consequently for all the safeguards and requirements of kinship and breaking up and reconstituting of groups. Another theoretical glitch in extant models is that they assert conditional altruism, in which the altruist can sometimes behave selfishly, and at the same time they assert a separate altruistic allele under the threat of elimination due to competition from a separate selfish allele. However, the first assertion binds the two phenotypes under a single genotype, and thus precludes the second. This logical inconsistency entailed in those models is yet another reason why they have languished for so long in the theoretical morass.

Bibliography

Agrawal, A. A. (2001) Phenotypic Plasticity in the Interactions and Evolution of Species. *Science* 294: 321-326.

Allen, N.B. and Badcock, P.B. (2006) Genes For Susceptibility to Mental Disorder are not Mental Disorder: Clarifying the Target of Evolutionary Analysis and the role of the Environment. *Behavioural and Brain Sciences.* 29 (3): 405-406.

Allsop, D. J & West, S.A. (2004) Sex Allocation in the Sex-changing Marine Goby (*Coryphopterus personatus*) on Atoll-fringing Reefs. *Evolutionary Ecology Research* 6: 843-855.

Axelrod, R. and Hamilton, W.D. (1981) Evolution of Cooperation. *Science* 211: 1390-1397.

Behrendt, Ralf-Peter (2006), The Social Environment Compresses the Diversity of Genetic Aberrations into the Uniformity of Schizophrenia Manifestations. *Behavioural and Brain Sciences.* 29 (4): 406-408.

Bernhard, H., Fischbacher., U and Fehr, E. (2006), Parochial altruism in humans. *Nature.* Vol 442|24 August 2006|doi:10.1038/nature04981.

Blackman, R. L. (1977), The existence of two species of *Euceraphis* (Homoptera: Aphididae) on birch in Western Europe, and a key to European and North American species of the genus. *Systematic Entomology* (I) 1-8.

Bowles, S. (2006) Group Selection, Reproductive Levelling, and the Evolution of Human Altruism. *Science* 314: 1569-1572.

Burghes, A.M.H., H. E.F. Vaessin, A. de la Chapelle (2001), The Land Between Mendelian and Multifactorial Inheritance. *Science*, September 2001.

Castro L., Serrano J., and Toro M. (1998), Conceptual Capacity to Categorize and the Evolution of Altruism. *Journal of Theoretical Biology.* 192(4) 561-565.

Chalmers, A. F. (1999) What Is This Thing Called Science? *Univeersity of Queensland Press* (3rd ed.).

Copernicus, Nicolaus, Three Copernican treatises: the Commentariolus of Copernicus, the Letter against Werner, the Narratio prima of Rheticus. *Edward Rosen (Transl.) Octagon Books (1971)*.

Copernicus, Nicolaus (1543), On the Revolutions if the Heavenly Spheres. *Running Press Book Publishers*, Philadelphia.

Curry, O. (2006) One Good Deed: Can a Simple Equation Explain the Development of Altruism? *Nature* 444: 683.

Darwin, C. (1871) The Origin of Species and the Descent of Man. *Modern Library, New York.*

Darwin, C. (1859) The Origin of Species. *Modern Library, New York* (1998).

Dawkins, R. (1982), Replicators and Vehicles. In: Currents Problems in Sociobiology. Eds. King's College Sociobiology Group. *Cambridge University Press.*

Dawkins, R. (1976), The Selfish Gene. *Oxford University Press* (1989).

Donnell, D. M.; Corley, L. S.; Chen, G.; and Strand, M. R. (2004) Caste determination in a polyembryonic wasp involves inheritance of germ cells. *Proc Natl Acad Sci U S A*. 2004 101(27): 10095–10100.

Deno, R. F. & Roderick, G. K. (1992) Density Related Dispersal in Planthoppers: Effects of Interspecific Crowding. *Ecology*. 73(4): 1323-1334.

DeWitt, T. J. & Scheiner, S. M. (2004) Phenotypic Variation from single Genotypes. In: Phenotypic Plasticity – Functional and Conceptual Approaches. ed. DeWitt, T.J. & Scheiner, S.M. *Oxford University Press.*

Emlen, S. T. & Wrege, P.H. (1992) Parent-offspring Conflict and the Recruitment of Helpers Among Bee-eaters. *Nature* 356: 331-333.

Evans, J. D., and D. E. Wheeler (1999), Differential Gene Expression Between Developing Queens and Workers in the Honey Bee, Apis mellifera. *Proceedings of the National Academy of Science*. USA Vol. 96, pp. 5575-5580.

Falconer, D. S. (1989) Introduction to Quantitative genetics. *Longmans, New York*.

Faulks, C. G., Abbott, D. H., Liddell, C.E.; George, L. M. & Jarvis, J. U. M. (1991), Hormonal and Behavioural Aspects of Reproductive suppression in Female Naked Mole-Rats. In: The Biology of the Naked Mole-Rat. *Princeton University Press (1991).*

Fehr, Ernst & Urs Fischbacher (2003) The nature of human altruism. *. Nature* **425**, 785-791 (23 October 2003).

Fehr, E. & Schmidt, K.M. (1999) A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics.* 114, 817–868.

Field, J. Cronin, A. and C. Bridge (2006), Future Fitness and Helping in Social Queues. *Nature* 441: 214-217.

Fischer, R. A. (1929) The Genetical Theory of Natural Selection. *Dover Publications (1958).*

Foster, K. R., Shaulsky, G., Strassmann, J. E., Queller, D., C., Thompson, C. R. L., (2004) Pleiotropy as a Mechanism to Stabilize Cooperation. *Nature.* Vol. 431, 693-696. Oct. 2004.

Foster, K. R., T. Wenseleers & F. L. W. Ratnieks (2006) Kin Selection is the Key to Altruism. *Trends in Ecology and Evolution.* 21(2): 57-60.

Futuyma, D. (1998) Evolutionary Biology. *Sinauer Associates Inc. Sunderland, Massachusetts.* (3rd ed.)

Gadagkar, R. (1997) Survival strategies – Cooperation and Conflict in Animal Societies. *Harvard University Press.*

Gildenhuys, P. (2003) The Evolution of Altruism: The Sober/Wilson Model. *Philosophy of Science* 70: 27-48.

Greene, E. (1989) A Diet-Induced Developmental Polymorphism in a Caterpillar. *Science* 243: 643-646.

Godfrey-Smith, Peter (2009) Darwinian Populations and Natural Selection (*Oxford University Press*).

Gould, S. J. (2002) The Structure of Evolutionary Theory. *Belknap – Harvard* (2002).

Haldane, J. B. S. (1932) The Causes of Evolution. *Longmans, Green and Co. London. New York. Toronto.*

Haldane, J. B. S. (1955) Population Genetics. *New Biology* 18: 34–51.

Haldane, *J.* B. S. (1957) The cost of natural *selection. J. Genet.* 55:511-24

Hall, Stephen (2010) Revolution Postponed. *Scientific American.* October 2010, 60-67.

Hamilton, W. D. (1963) Evolution of Altruistic Behaviour. The American Naturalist. 97: 354-356. Published in: Narrow Roads Of Gene Land*. W. H. Freeman Spektrum.* Oxford. New York. Heidelberg.

Hamilton, W. D. (1964) Genetical Evolution of Social behaviour, I and II. *Journal of Theoretical Biology* 7: 1-52.

Harvey, J. A., Corley L. S. & Strand, M. R. (2000) Competition Induces Adaptive Shifts in Caste Ratios of Polyembryonic Wasps. *Nature* 406: 183-186

Hazel, W.N.; Smock, R.; & Johnson, M.D. (1990) A polygenic model for the evolution and maintenance of conditional strategies. *Proceedings of the Royal Society of London (B)* **242**: 181–187.

Huxley, Julian (1942) Evolution: The Modern Synthesis. *George Allen & Unwin Ltd* (1974).

Johns, P. M., K. J. Howard, N. L. Breisch, A. Rivera, and B. L. Thorne (2009) Nonrelatives inherit colony resources in a primitive termite. *Proceedings of the National Academy of Science* vol. 106, no. 41.

Keller, M. C. & Miller, G. (2006) Resolving the Paradox of Common, Harmful, Heritable Mental Disorders: Which Evolutionary Genetic Models Work Best? *Behavioural and Brain Sciences* 29, 385-452.

Keller, L. (2009) Genetics of behaviour. Philosophical. *Transactions of the Royal Society- Biological Sciences* 364: 3209–3216

Keller, L. & Ross, K. G. (1993) Phenotypic Basis of Reproductive Success in the Social Insect: Genetic and Social Determinants. *Science* 260: 1107-1110.

Kendler, K. S. (2005) "A Gene for …": The Nature of Gene Action in Psychiatric Disorders. *American Journal of Psychiatry* 162: 1245-1252.

Kepler, Johannes (1609) New Astronomy. (Transl.) William Donahue. *Cambridge University Press* (1992).

Kitcher, Philip (1982) Abusing Science: The Case Against Creationism. *The MIT Press, Cambridge.*
Koestler, Arthur (1963) The Sleepwalkers. *The Universal Library, New York.*

Lacey, A. H. & Sherman P. (1991) Social Organization of Mole-Rat Colonies: Evidence of for Division of Labour. In: *The Biology of the Naked Mole-Rat. Princeton University Press* (1991).

Lakatos, Imre (1978) The Methodology of Scientific Research Programmes. *Cambridge University Press* (1978).

Lively, C. M. (1986) Predator-Induced Shell Dimorphism in the Acorn Barnacle (*Chthamalus anisopoma*). *Evolution* 40 (2): 232-242.

Marshall, Michael (2010) Sparks Fly Over Origin of Altruism. *New Scientist. Issue 2780,* September 29, 2010.

Mattila, H. R. and Seeley, T. D. (2007) Genetic diversity in honey bee colonies enhances productivity and fitness. *Science* 317: 362-364.

Maynard-Smith, J. (1964) Group Selection and Kin Selection. *Nature* 201: 1145-1147.

Maynard-Smith, J. (1972) On Evolution. *Edinburgh University Press.*

Maynard-Smith, J. (1998) The Origin of Altruism. *Nature* 393: 639-640.

Monroe S. M & Simons A. D. (1991) Diathesis-stress theories in the context of life tress research: Implications for the depressive disorders. *Psychological Bulletin* 110(3): 406-425.

More, J. (1995) Behaviour of Parasitized Animals. *Bioscience* 45(2): 89-96.

Morreale, S. J., Ruiz, G. J., Spotila J. R. & Sandora E. A. (1982) Temperature-dependent sex determination: current practices threaten conservation of sea turtles. *Science*. 11: 216(4551): 1245-7.

Nowak, M. A., Tarnita, C. E., and Wilson, E. O. (2010) The Evolution of Eusociality. *Nature*. Vol. 466: 1057–1062 (August 2010)

O'Donnell, S. (1998) Reproductive Caste Determination in Eusocial Wasps (Hymenoptera: Vespidae). *Annual Review of Entomology* Vol. 43(1): 323-346.

Okasha, S. (2002) Genetic Relatedness and the Evolution of Altruism. *Philosophy of Science* 69: 138-149.

Okasha, Samir, (2008), "Population Genetics", *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*, Edward N. Zalta (ed.). URL = <http://plato.stanford.edu/archives/fall2008/entries/population-genetics/>.

Okasha, S. (2009), Biological Altruism. *The Stanford Encyclopedia of Philosophy (Winter 2009 Edition)*, Edward N. Zalta (ed.), URL= http://plato.stanford.edu/archives/win2009/entries/altruism-biological/>.

Oster, G. F. & Wilson E. O. (1978) Caste and Ecology in the Social Insects. *Princeton University press*.

Packer, C. (1977) Reciprocal Altruism in Olive Baboons. *Nature* 265: 441-443.

Parsons, Keith (2006) Copernican Questions: *A Concise Invitation to the Philosophy of Science. McGraw-Hill, New York*.

Patel, A., Fondrk, M. K., Kaftanoglu O., Emore C., Hunt G., Frederick. K. & Amdam G. V. (2007) The Making of a Queen: TOR Pathway is a Key Player in Diphenic Caste Development. *PloS ONE* 2(6): e509.

Peirce, C. S., The Collected Papers of Charles Sanders Peirce. Vols. I-VI ed. Charles Hartshorne and Paul Weiss (Cambridge, MA: Harvard University Press, 1931-1935), Vols. VII-VIII ed. Arthur W. Burks (same publisher, 1958).

Pfennig, D. W. & Murphy, P. J. (2000) Character Displacement in Polyphenic Tadpoles *Evolution* 54(5): 1738-1749.

Pigliucci, M. (2001) Phenotypic Plasticity – Beyond Nature and Nurture. *The John Hopkins University Press.*

Pigliucci, M. (2009) An Extended Synthesis for Evolutionary Biology. *Annals of the New York Academy of Science* 1168: 218–228.

Prete, F. R. (1990) The Conundrum of the Honey Bees: One Impediment to the Publication of Darwin's Theory. *Journal of the History Biology* 23(2): 271-290.

Pritchard, J. K (2001) Are Rare Variants Responsible for Susceptibility to Complex Diseases? *American Journal of Human Genetics 69:124–137*

Popper, K. R., (1963) Conjectures and Refutations: The Growth of Scientific Knowledge. *Harper and Row (1968).*

Queller, D. C. 1985. Kinship, reciprocity and synergism in the evolution of social behavior. *Nature* 318:366–367.

Queller, D. C. (2006) To work or not to Work. *Nature* 444 (2): 42-43

Rabenold, K. N (1985) Cooperation in Breeding by Non –Reproductive Wrens: Kinship, Reciprocity, and Demography. *Behavioural Ecology and Sociobiolology* 17:1-17.

Ratnieks, L. W. & Visscher P. K. (1989) Worker Policing In The Honey Bee. *Nature* 342: 796-797.

Roff, D. A. (1994), The Evolution of Dimorphic Traits: Predicting The Genetic Correlation Between Environments. *Genetics* 136: 395-401.

Roff, D. A. (1986) The Evolution of Wing Dimorphism in Insects. *Evolution* 40: 1009-1040.

Rousset, Francois and Denis Roze (2007), Constraints on the Origin and Maintenance of Genetic Kin Recognition. *Evolution* 61-10: 2320–2330

Seyfarth, R. M. & Cheney D. L. (1984) Grooming, Alliances and Reciprocal Altruism in Velvet Monkeys. *Nature* (308) 541-543.

Shanahan, T. (2004) The Evolution of Darwinism – Selection, Adaptation, and Progress in Evolutionary Biology. *Cambridge University Press.*

Sherman, P. W. (1977) Nepotism and the Evolution of Alarm Calls. *Science* 197: 1246-1253.

Sober, Elliot (2000) Philosophy of Biology. *Westview Press. Boulder, Colorado.*

Sober, E. (1984) The Nature of Selection. *Dover Books, Princeton University Press.*

Sober, E. & Wilson, D. S. (1998) Unto Others: The Evolution and Psychology of Unselfish Behaviour. *Harvard University Press.* Cambridge.

Strassmann, J. E.; Zhu, Y.; and Queller, D. C. (2000), Altruism and social cheating in the social amoeba *Dictyostelium discoideum. Nature.* Vol. 408 December 2000. 695-697

Thompson CR, Fu Q, Buhay C, Kay RR, Shaulsky G. (2004), A bZIP/bRLZ transcription factor required for DIF signaling in *Dictyostelium. Development.* 2004 Feb;131(3):513-23.
Trivers, R. L. (1971) The Evolution of Reciprocal Altruism. *Quarterly Review of Biology* 46: 35-57.

Trivers, R. L, & Hare, H. (1976) Haplodiploidy and the Evolution of the Social Insects. *Science.* 191: 249-263.

Wade, M. J. (1977) An experimental study of group selection. *Evolution* 31:134–153.

Wenseleers, T. & Ratnieks, F. L. W. (2006) Enforced Altruism in Insect Societies. *Nature* 444(2): 50.

West-Eberhard, M. J. (1975) The evolution of social behavior by kin selection. The *Quarterly Review of Biology.* 50: l-33

West-Eberhard, M. J. (1986), Alternative Adaptations, Speciation, and Phylogeny. *Proceedings of the National Academy of Science.* Vol. 83:1388-1392.

Whewell, William (1847) The Philosophy of the inductive Sciences. *Frank Cass & Co, (1967).*
Wilkinson, G. S. (1984) Reciprocal Food Sharing in the Vampire Bat. *Nature.* 308: 181-184.

Wilkinson, G. S. (1984) Reciprocal Food Sharing in the Vampire Bat. *Nature.* 308: 181-184.

William, B. J. (1981), A Critical Review of Models In Sociobiology. *Annual Review of Anthropology.* 10:163-192.

Wilson, D. S. (1989) Levels of Selection: An Alternative to Individualism in Biology and the Human Sciences*. Social Networks* 11: 257-272.

Wilson, E. O. (2008) One Giant Leap: How Insects Achieved Altruism and Colonial Life. *BioScience .* January 2008 / Vol. 58 No. 1 (17-25).

Wilson, E. O. (2005) Kin selection as the Key to Altruism: Its Rise and Fall*. Social Research.* Vol. 72 (1): 159-166.

Wilson, R. A. (2004) Test Cases, Resolvability, and Group Selection: A Critical Examination of the Myxoma Case. *Philosophy of Science.* 71, 380-401.

Windig, J. J.; Kovel, C. G. F. D.; and Jong, G. D. (2004) Genetic and mechanics of plasticity. In: T. J. DeWitt and S. M. Scheiner, eds. *Phenotypic plasticity: functional and conceptual approaches. Oxford University Press*, Oxford.

Winston, Mark L. (1987) The Biology of the Honeybee. *Harvard University Press, Cambridge, Massachusetts*. London, England.

Wynne-Edwards, V.C. (1962) Animal Dispersion in Relation to Social Behavior. *Oliver & Boyd, London*.

Wynne-Edwards, V. C. (1986) Evolution Through Group Selection *Blackwell.* ISBN 0-632-01541-1.

Young, J. A. and Tim Clutton-Brock (2006) Infanticide by Subordinates Influences Reproductive Sharing in Cooperatively Breeding Meerkats. *Biology Letters.* 2(3): 385-387.

Zimbardo, P. (1971) The pathology of imprisonment. *Society, 9*, 4-8. p. 4.