# EVOLUTIONARY TRADE-OFFS: EMERGENT CONSTRAINTS AND THEIR ADAPTIVE CONSEQUENCES

by

Bret S. Weinstein

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Biology)
in The University of Michigan
2009

Doctoral Committee:

Professor Emeritus Richard D. Alexander, Co-Chair
Associate Professor Robyn J. Burnham, Co-Chair
Professor Emeritus Gerald R. Smith
Associate Professor Beverly I. Strassmann
Assistant Professor Elizabeth A. Tibbetts

---

# DEDICATION

When I was a little boy my grandparents, Harry and Sophie Rubin, took me on a hike in Griffith Park, in the hills above Los Angeles. As we sat eating our lunch of dry salami, crackers and gouda cheese my grandfather handed me a magnifying glass and pointed me to the center of a flower. As I looked at all the levels of intricate structure, the two of them told me about Darwin, and how all the complexity that I saw was a consequence of a simple thing called selection. When I looked up I was not the same. This dissertation is a status report on a project my grandparents set in motion. They dedicated themselves to my brother and me. And I dedicate the project, and this thesis, to them.

| Sophie Rubin | Harry Rubin |
|---|---|
| 1914-1978 | 1913-2008 |

# ACKNOWLEDGEMENTS

This thesis is the outgrowth of many people's efforts and influences. My maternal grandparents, Harry and Sophie Rubin, had a great deal to do with cultivating the traits on which I have leaned most heavily. My brother, Eric, has been a strongly positive force in my life since childhood. His intellectual influence is evident throughout the work presented here, and much that is correct within it was built and refined with his invaluable insight and guidance.

Larry Clark at the Monell Center in Philadelphia was the first teacher I had that understood the Darwinian landscape at a deep level. His willingness to take a young undergraduate seriously in his graduate Chemical Ecology seminar had lasting effects.

Landing in Bob Trivers' Social Evolution class at UCSC was an extraordinary bit of good fortune. Listening to hours of Bob's insights, delivered at the chalkboard was a special pleasure, and has strongly affected my approach to selection, as well as to teaching. Bob also took me to Jamaica as his field assistant. That experience opened my eyes to dozens of important things about the evolutionary nature of humans, islands and the subtle dynamics at the frontier of evolutionary knowledge.

Dick Alexander's influence as my graduate mentor is not easily summarized. As recognized as he is for his achievements as a biologist, I believe that Dick has yet to get his full due. It has been rare in my experience to come up with a really good idea that Dick has not gotten to first in one way or another—an unusual situation given the huge number of topics subsumed under our mutual interest, 'selection'. He has been everywhere, on the map and off, and his powers of observation and analytical capacity have made for invaluable surveys of that landscape. At the same time, he is singularly non-territorial about the puzzles he has encountered, and the work he has put into them—Dick likes to hash out ideas with other people who are interested in the same questions.

That is the single most valuable trait an advisee could find in an advisor, to be sure. Dick's insights are all the more important because he has understood, better than anyone else I know of, the unique way selection has shaped humans in the last 6 million years.

Dick's open-door and open-mind created a special zone of intellectual integrity and rigor around his office in the museum. It collected refugees from all over the Biology department, and was a great place to refine ideas and test out new heretical formulations. Among the many people in that sphere that helped shape my thinking about selection, a number stand out. Andrew Richards, David Marshall, John Cooley, Laura Howard, Deborah Ciszek and David Lahti have each been influential. David Lahti and Deborah Ciszek have also been excellent collaborators. Chapter 2 of this dissertation was co-authored with Deborah Ciszek, and Chapter 4 with David Lahti, and the only reason their names are not on those chapters is that putting their names there violates an apparently inviolable University policy.

Jerry Smith has played two crucial roles in my academic/intellectual life. First, he saved me from my detractors on several occasions. Second, he has encouraged me to think on much longer timescales than are typical in biology, even amongst evolutionists. That instinct—to look at selection on very long timescales—is, I believe, the key to many of the patterns that remain stubbornly paradoxical and enigmatic, and Jerry is the primary source.

Arnold Kluge also played two important roles. First, he taught me the correct way to draw phylogenetic inferences, which is, in many ways, the most important tool in the adaptationist's toolkit. The second thing he did was bring the philosophy of rigorous scientific thinking up to the explicit layer. All of the people I mention here are rigorous thinkers, but Arnold introduced me, and many others, to the inner workings of that machine.

Heather Heying, to whom I am now married, has helped shape every facet of my work since we joined forces in 1989. We have talked in detail about every Darwinian concept in this thesis, and well beyond. Having her excellent evolutionary mind to generate, shape and refine ideas with has been immensely useful. Without our many thousands of hours of discussion about selection, I can't say where any of this work would be. I can, however, say with confidence that the thesis never would have been

completed, and I would not have gotten a professorship, absent her many talents and extraordinary patience.

Elisabeth Kalko and the late Charles Handley both invested heavily in me and my work. The thesis contains none of the empirical product of that effort, but much of what is here was produced on BCI and was more than just catalyzed by the Smithsonian bat project which they created, and to which they welcomed me.

Peter Howard has been extremely helpful in the final construction of this thesis. Others that have played an important and positive role include Jennifer Ast, Antonia Gorog, Gary Paoli, Pia Malaney, Nat Kurnick, Jae Choe, Arnold Marglin, Beverly Rathcke, George Estabrook, John Vandermeer, Jerry Urquhart, Chris Picone, Tom DiBenedetto, Dyan Haspel, Allen Herre, Randy Nesse, Cecile Bockholdt, Mike Hemann, Alec Lindsay, Paul Fine, Lissy Coley, Tom Kursar, Egbert Leigh, Steve Hubbell, Scott Peacor, Bill Rice, Austin Burt, Paula Schofield, Jack Longino, Nalini Nadkarni and many insightful undergraduates at The Evergreen State College.

Finally, I wish to thank my committee, Robyn Burnham, Richard Alexander, Gerald Smith, Beverly Strassmann and Elizabeth Tibbetts for putting up with much that was non-standard. Their coordinated efforts and insight made this thesis considerably better than it otherwise have been.

# TABLE OF CONTENTS

# LIST OF FIGURES

**Figure**

# ABSTRACT

Trade-offs are widely recognized in biology, but the rules that govern them are not yet well understood. Increased predictive power can be gained by treating trade-offs as emergent phenomena governed by laws that are also emergent. The law-like nature of trade-offs becomes evident when we subdivide examples by type.

Trade-offs can be A) probabilistic (e.g. the darkest individual in a population will rarely be the biggest), B) based on the mutually exclusive allocation of resources (e.g. roots vs. shoots), or C) based on extrinsic design limitations (e.g. high efficiency vs. extreme robustness). Only design trade-offs are law-like, although the other two types may be transformed into design limitations given strong selective pressures.

Between every two fitness-enhancing characteristics of an organism or mechanism, a design trade-off must logically exist, preventing simultaneous optimization. Selection's tendency toward optimization reveals that fraction of design trade-offs we come to empirically recognize. A particular trade-off can be evident both within and between species. Trade-offs may be obscured by insufficient selective time, noisy or fluctuating selective environments, and weak selection pressures.

A natural parallel exists between trade-offs in space, and in time. The interrelation between these phenomena on the one hand, and niche-partitioning, competitive exclusion, character displacement and phenotypic plasticity on the other is also considered.

Chapter One describes the proposed theoretical landscape. Chapter Two describes a senescence-causing trade-off between cancer prevention and tissue-repair capacity in vertebrates. Chapter Three relates the latitudinal diversity gradient to a gradient of design constraints that is a consequence of environmental fluctuation positively correlated with latitude on all relevant time scales. A natural reconciliation between niche assembly and community drift is proposed, and the effects of mate choice on diversity patterns is

considered. Chapter Four argues that facultative human moral self-restraint is an evolutionary response to an inescapable trade-off between the component of fitness that results from success in competition within one's lineage, and the component that derives from the success of one's group in competition with other lineages.

# Chapter 1

# EVOLUTIONARY TRADE-OFFS AS A CENTRAL ORGANIZING PRINCIPLE IN BIOLOGY

Trade-offs are negative relationships between desirable characteristics, and are widely recognized across biology, but there is much about them that remains unclear. The potential importance of trade-offs as an organizing principle has been discussed, and previous attempts have been made to sketch the landscape (e.g. Stearns, 1992, pp. 72-90), but much of the explanatory power of trade-offs remains unmapped and untapped. What follows is an attempt to increase the utility of the trade-off concept by subdividing the concept into natural types.

One of the reasons that trade-offs are incompletely known within biology is that even the simplest biological organisms are unimaginably complex. Because we are always working with a crude and incomplete understanding of the adaptations that comprise a given organism, the natural tendency of evolution to modulate and balance competing concerns may be obscured.

Additionally, as Stearns (1992) observes, trade-offs may be hidden by the way in which we study. One might easily get the wrong idea, for example, about the well known trade-off between a plant's allocation of resources to above-ground and below-ground structures. Clearly an individual plant must allocate each unit of resource in one direction or the other, so there must logically be a trade-off. But you would not get that impression if you measured the masses of roots and shoots of individuals sampled as you moved from high altitude to low, because plants of a given species will tend to be larger in all regards at lower altitudes, so root mass and shoot mass will be *positively*, rather than negatively, correlated. In order to see the trade-off, we must control for variables that either encourage or tax the plant as a whole.

One solution to the problem of trade-offs obscured by complexity (including those masked by other variables) is to step out of the traditional boundaries of biology and into parallel realms in which trade-offs are better understood. Engineers, for example, work with trade-offs on a daily basis. And it is quite common to hear agreement among them regarding principles, even laws, that have equal reason to be expected in biology.

It is commonly asserted in engineering, for example—from computer science to aviation—that two properties of a machine, program or mechanism can not be simultaneously optimized. There is, however no such consensus in biology, and that is striking because the corollary of that idea, Macarthur's (1961) Jack-of-all-trades principle, is known and accepted, strongly implying the applicability of the stronger version from engineering within biology.

It is possible that there is some threshold of complexity above which these engineering principles cease being applicable, but that constitutes special pleading unless evidence of such thresholds emerges between the level of the most complex machines, and the simplest organisms. Until then, we are justified in cautiously peering into fields where the complexity is simple relative to the noisy biotic systems we primarily wish to untangle (Csete and Doyle, 2002).

The other advantage of learning from engineered machines about the constraints that shape the adaptive landscape (what engineers sometimes call 'design space') is that, unlike biological organisms, one can have a very complete understanding of exactly what each feature of a machine is intended to accomplish. That just isn't the case in biology—the dewlap of an *Anolis* lizard is a signal, but of what? That is unclear. And without complete knowledge of the advantage provided by a trait, it is difficult to do a cost/benefit analysis, which is at the core of understanding how any product of adaptation trades-off against any other.

Trade-offs can usefully be divided into three types, probabilistic trade-offs, allocation trade-offs, and those trade-offs arising from inherent design-constraints. And dividing trade-offs in this way allows us to see that each has unique properties that must influence how they interact with adaptive evolutionary forces.

Probabilistic trade-offs are the weakest of the three, and many such trade-offs will be dismantled if sufficient selective advantage arises from doing so. Suppose, for example that a female frog prefers males that are unusually large, and at the same time, males that are unusually blue. And suppose further that size and color are largely independent polygenic traits. From that female's perspective, preference for large size and blue color are likely to trade-off against each other. There will, for obvious reasons, be few individuals in the remote right tail of the distribution for size, and there will be few individuals in extreme blue tail of the color distribution. As a consequence, she is likely to have a very difficult time finding individuals in both tails and, as with any trade-off, she will have to prioritize the two considerations. That being said, if size and color are not at odds for some functional reason, then selection by females could produce, over time, large, intensely blue, males, thus eliminating the initial trade-off. If on the other hand, there is a significant fictional relationship between size and color such that being extreme in one regard has costs with respect to the other, then the probabilistic trade-off will be converted into either an allocation trade-off, or a design constraint trade-off depending on the functional nature of the relationship.

And the same can be seen in the world of machines. The likelihood that the camera with the best light-metering will be the camera with the sharpest lens will initially be low, unless there is demand for a camera with both an unusually sharp lens and an unusually powerful light-meter, in which case, there being no obstacle to the production of such a camera, one might well be built.

The second kind of trade-off, the type arising from the allocation of a limited resource, is illustrated by the root/shoot example above. This is the type on which Stearns' (1992) review of the topic is primarily concerned. Allocation trade-offs arise in any instance where a resource must be divided amongst competing concerns. The existence of such trade-offs does not therefore depend on the past action of selection—an allocation trade-off says nothing at all about the value of any particular division. A family could spend twenty percent of its budget on food, ten percent on housing, investing the remainder in lottery tickets and that would qualify as a three way trade-off, even though the budget makes no financial sense. Likewise, an individual songbird could spend all its time searching for mates out of season, thus failing to forage sufficiently to maintain

homeostasis and the trade-off would be just as real. Of course, in general, the division of resources exhibited by organisms will quickly honed by selection to reflect an adaptive division of the resource. The reproductively optimizing force of selection will tend to convert such trade-offs into something that behaves like a design-constraint trade-off (described below), with one important difference: allocation trade-offs can be eliminated, at least in the short term, with supplemental resources. This is why the root/shoot trade-off is obscured by changes in altitude—lower altitudes act like a supplement of resources above that available at high altitudes. One can see the same effect using fertilizers, supplemental light, water, etc. And the analogy holds into the realm of machines.

Consider the plight of a photographer trying to capture a picture of a moving object while maintaining a large depth of field, such that things at various distances from the camera are in sharp focus. The amount of light needed to get the right exposure is a simple sum affected by two parameters, the time the shutter is open, and the size of the aperture in the lens. For a given level of incident light, a particular exposure is required. If the variables were continuous rather than discrete, then there would be an infinite set of combinations of lens openings and shutter-speeds that would yield the right exposure. But the particular combinations would yield different costs and benefits. Those combinations with large apertures and fast shutter-speeds would freeze motion, at the cost of a narrow depth of field (only a narrow band of objects at a given distance away would be in focus). While, slow shutter speeds with small aperture openings will give a large depth of field, at the cost of moving objects being blurred. But the trade-off between depth of field and the freezing of motion evaporates if we supply large quantities of extra photons (as with a flash), allowing us to produce the same exposure with the lens opening small *and* the shutter set to a brief period.

The final trade-off type is the design-constraint trade-off. These trade-offs occur simply because the same form cannot be optimized for two different tasks simultaneously. Unlike the other two types, trade-offs that derive from design constraints are emergent phenomena, unobservable until revealed by selection of sufficient strength and duration to bump species up against them. The most important aspect of design constraint trade-offs is that they are insensitive to resource supplementation, and thus produce hard limits on what selection can and can not do. There are two important sub-

categories within design-constraint trade-offs: the degree of hardness is not universal. Some design-constraint trade-offs are local optima that can be exceeded once an 'innovation' arises that allows circumvention. Others are global limits that can not be exceeded for non-biotic reasons.

Design-constraints are likely to be the richest form of trade-off in terms of untapped explanatory power, because they exist inherently at the frontier where the biota meets its limitation. They are, in a sense, the primary geological force acting on the adaptive landscape—Dawkins' metaphor of Mount Improbable may explain how simple processes can produce instances of mind boggling complexity, but design-constraints are the reason Mount Improbable takes the shape it does, why creatures don't rise forever toward functional perfection.

But design-constraint trade-offs are also the most easily misunderstood because, unlike allocation trade-offs that exist irrespective of selection, and unlike probabilistic trade-offs that are destroyed by selection, design constraints are invisible until the particular quadrant of the adaptive landscape in which they exist is explored by selection. And an invisible trade-off may be mistaken for non-existent in the absence of a good theoretical basis for imagining what must be there. If we mistake the revelation of a trade-off for its invention, then we miss their most powerful feature: we do not have to ask if there *is* a trade-off between capacity A and capacity B of an organism, nor should we be surprised when we find that there is. We are in an important sense justified in *expecting* these limits to exist between every two functional adaptations. And though many of these of these hidden relationships may be insignificant in their effect, a large number are, in one way or another, central to the way species divide time and space.

Examples are the key to understanding design-constraint trade-offs, and there are many to choose from. Let's start with the evolution of flight in bats from a gliding ancestor. True powered fight is superior to gliding in many important regards. A flyer is both more agile and more efficient than a glider, but between agility and efficiency, there is a well established trade-off: agile bats (e.g. flower visiting Glossophagine bats, which could properly be called 'humming bats') have short paddle like wings, while open space bats that travel long distances, have long narrow wings (e.g. Molossid bats that roost in very large colonies and must therefore fly farther to escape the intense local competition

for food). And no matter how well fed an animal is, there is no apparent way to evade this trade-off, wings are either optimized in the direction of efficiency or in the direction of agility, or they are a compromise between the two.

Significantly, birds show the same pattern, hummingbirds and frigate birds representing their respective ends of the same trade-off. And the fact that organisms discover extrinsic limits that are the same has deep implications within evolutionary ecology. The fact of a high degree of lottery competition (Hubbell, 1997) existing in habitats that also apparently favor a high degree of specialization (Fine *et al*, 2006) is easily reconciled given the existence of extrinsic limits that constrain all taxa that reach them.

But the diversity we find along a given design-constraint trade-off (different organisms having evolved to accept different points on the spectrum of possibilities), coupled with the paleontological evidence of the evolutionary trajectory taken by the lineage on the way to the trade-off, suggests another important pattern. Birds and bats are both thought to have evolved from gliding ancestors. And it stands to reason that the emergence of powered flight, a complex trait, would initially have been crude—both inefficient and clumsy. And as selection refined the trait, the intermediates would have been increasingly agile and efficient, on average, each generation slightly more so than the last. Creatures from this phase of evolution would appear not to be choosing between mutually exclusive advantages, but rather improving generally.

And we have seen this same pattern in machines. The Wright Flyer (the first successful powered airplane) was both incredibly clumsy and inefficient, and also the first to take to the air because it was less so than any competing design. That initial innovation over successful gliding designs engendered almost unimaginably rapid refinement. At first there were not many types of planes, all planes carried one or two passengers exposed to the air, and little else. But, these designs allowed furious refinement, improving all the characteristics simultaneously. Ultimately, of course, trade-offs did emerge and designs diversified, carrying a large load necessitating a substantially different design than avoiding enemy fire, to take one example.

Another important pattern can be gleaned from the history of aviation. There have been numerous innovations that allow designs to evade some previously limiting barrier.

The ailerons supplanted the Wright brother's wing-warping technology, increasing the maneuverability of planes. Tricycle landing gear replaced tail-dragging designs, improving control and safety. And similar claims can be made for jets over propellers, swept wings over perpendicular ones--the full list being extremely long.

There is also something interesting in the fact of certain designs persisting within 'niches' despite the existence of designs that are, in one sense or another, superior. Many propeller-driven planes persist and 'outperform' jets for tight maneuvering and low overall cost. And other types of niches also exist amongst machines. Helicopters are useful where no runway is available for takeoff and landing, but this comes at a cost in terms of the upper limits of the design. The upper limits for helicopter speed are quite low compared to the limits on planes because, as the speed of the aircraft approaches the speed of the tips of the rotor through the air, the retreating blade (the one moving in the opposite direction) is effectively standing still relative to the air, thus generating no lift, causing the aircraft to flop over. So, we can recognize a second trade-off in this machine example, between the fast flight and the ability to land in a confined space, and this one is the basis for a type of niche partitioning, much like the inversions of competitive dominance we see between congeners in some habitat pairs (Fine, 2006).

Given the above arguments, design-constraint trade-offs become the most profitable focus for a biologist. What pattern should we expect of design trade-offs in complex entities? We have already touched on the fact that engineers expect the optimization of any particular parameter to have negative impacts on the upper limits possible for every other parameter, suggesting a universality to trade-offs—between every two desirable characteristics, there exists a trade-off that can not be exceeded. This claim strikes many biologists as a bridge too far, there being no intuitive currency to the idea that coat color should negatively impact speed, for example. But the reality is that there are feedbacks that force such a theoretical relationship to exist, even if it is not manifest in living examples. Coat color is, after all, related to the tendency of a creature to lose or retain heat. It also has a metabolic cost, and the apparatus that assembles pigments no doubt involves mass that must be carried. But the fact of pair wise trade-offs being ubiquitous is hidden by two facts within biology. First, organisms are never optimized for a single task, but instead must succeed in at least several ways to pass on

their genes at all. Second, some trade-offs involve inherent negative relationships, but with a bargain at one end such that evolution clumps all creatures at the same end.

Consider the claim that there is a trade-off between the safety and the efficiency of a car. It is true that you can remove the seatbelts and airbags from a car, and improve kilometridge by so doing. But the gains are so small and the cost in terms of safety are so large that no reasonable person would do it. On other hand, reasonable compromises can be made between these same two parameters. All else being equal, smaller cars are more efficient and less safe because A) in a collision the smaller the object, the more violently it is accelerated in a new direction, and B) because it necessarily has less deforming material with which to dissipate the force of impact. The fact of a significant trade-off and a trivial trade-off surrounding the same two parameters, safety and efficiency, is instructive. It implies that trade-offs are not always straight lines as we typically conceptualize them. Instead, they are more likely to abide by the law of diminishing returns, whereby increasingly extreme measures produce accelerating costs as one moves farther to either end of a given trade-off—much as you can make an efficient car that is very unsafe, you can make an unbelievably safe car that is too expensive to drive. We, of course, do neither and diversity is only seen in that middle ground where reasonable people can disagree, or where different tasks demand different priorities.

And that may be the most important thing to know about trade-offs. The reason we see so many implied by biotic diversity is likely related to the fact that a diminishing returns curve stretched between two desirable characteristics has a central section over which the tangent is close to 45 degrees (see chapter 3). We are likely to see diversity in and around that section, but not at the ends such that we don't even intuit the existence of the ends—the ends are purely theoretical, selection driving species toward the center where reasonable species can differ over particular strategies, divvying up space and dime in ways that demand explanation.

What follows are three chapters that tackle four significant questions with the same broad concept of trade-offs. Each involves a problem within evolutionary ecology (broadly defined) that has, at least in some regard, persisted for a long period without a consensus emerging about its evolutionary solution.

Chapter two involves a hypothesis that cancer, rather than being one of many effects of senescence, is actually the opposite of senescence. Building on George Williams' theory of antagonistic pleiotropy, the chapter proposes that the risk of cancer in early life is far greater than recognized, and that complex and highly effective mechanisms have evolved that constitute a tumor failsafe, reining in runaway cell lines so that they rarely interfere with our ability to reproduce. The specific mechanism proposed involves the shortening of telomeres with each cell division in most somatic tissues of vertebrates. Selection adjusts the upper limit on cellular reproduction, the Hayflick limit, on a tissue by tissue basis such that each tissue exhibits an independent balance along the design-constraint trade-off between tissue repair, and tumor resistance. When this hypothesis was first published, it was a radical idea. In the years since Weinstein and Ciszek (2002), many aspects have been tested and the idea now enjoys wide acceptance, though there is still disagreement about whether this is the central mechanism underlying vertebrate senescence, or one of several.

Chapter three seeks to explain the latitudinal diversity gradient using the logic of design-constraint trade-offs. It is based in two central ideas, the first being that wide climatic fluctuations on all relevant timescales in the temperate zones creates an temporal hazard which resist adaptive solutions. As species evolve toward competitive efficiency during mild periods (e.g. interglacial) they shed the robustness factors that got them through the last harsh bottleneck (e.g. glacial period) because those factors have a cost in competition and no present benefit. That renders such species vulnerable to extinction when the harsh conditions return. Creatures that resist this tendency become vulnerable to competitive exclusion during mild periods, and the oscillation between mild and harsh conditions sets up waves of extinction, there being no level of compromise between robustness and efficiency that resists both hazards indefinitely. The second set of arguments in this chapter surround the 'jack of all trades' principle and its consequences in relatively stable environments where competition exerts constant selective force in favor of specialization. It is argued that such selection is likely to break up widely distributed generalist species in tropical habitats into narrowly distributed specialist fragments through parapatric speciation (see Fine, 2006), likely facilitated by reinforcing

selection exerted by females choosing males that exhibit evidence for adaptation to the local environment.

Chapter 3 further addresses questions related to sexual selection. Females are thought by most models of sexual selection to favor males on the basis of honest indicators of good genes from which their offspring would benefit. But, if females disfavor males that carry 'bad' genes, generation after generation, then bad genes should become uncommon, rendering most of the efforts made by females in this direction a needless expense. That expense should favor females that are indifferent to indicators of quality because they get the benefits of past female choosiness without present cost. That should trigger a wave of female indifference, which should set the stage for bad genes to creep back in, drifting to increasing prevalence, at which point, female choosiness should become increasingly valuable and, therefore, common. Yet we don't see evidence of waves of female indifference in species with choosy females. And there is a second problem with such models. If good genes are indicated by costly displays in males, then whatever advantage they provide to a female's offspring must be diminished by the cost she inflicts on her sons by favoring costly male display. Unless the benefit to daughters outweighs the cost to sons, the displays and the preference for them should both be disfavored. In this chapter I propose two, compatible hypotheses to account for female vigilance in mate choice. The first involves the recognition that, especially in stable habitats (e.g. relatively aseasonal tropical habitats), 'good genes' are likely to be defined relative to local conditions that favor optimization in one direction in location A and, because of the trade-off principle, an divergent direction in location B (e.g. Fine, 2006; and see Hereford, 2009 for review of evidence of local-adaptation) Females in these locations are likely view a given male in opposite terms. If his genes are superior at A, they are likely to be inferior at B and visa versa. By imagining a local rather than global meaning to 'good genes' female vigilance becomes the expectation. Males from other locations are likely to be locally inferior, no matter how robust the may appear. And the fact that such males are always being created in adjacent habitats forces females to discriminate in every mating period.

In widely fluctuating habitats, a different explanation is proposed. The tendency of temperate lineages to evolve towards decreased tolerances during mild periods creates

the hazard discussed in chapter 3. But females are in a position to resist that tendency, thus retaining robustness factors in their lineage if, as a period of harshness recedes, they exhibit a preference for males with excess capacity demonstrated through costly display. In this way, females can function as a 'tensioner', compensating for the relaxation of selection that accompanies mild periods. A key prediction of this evolutionary tensioner idea is that such displays should be recoverable such that males can skimp when times are harsh (redirecting effort and materials toward survival), and display full strength when times are good.

The final chapter describes a trade-off proposed to underlie moral self-sacrifice in humans. It derives from the fact that ancestral humans had little ability to leave their group and thus virtually all human fitness depended in the long run on the well being of that lineage. That linkage implies that human fitness is a vector comprised of two components, the two existing in a trade-off relationship. One component of a human's fitness derives from the person's level of competitive success within that individual's group. The other derives from the success of one's lineage against other lineages. Jockeying for position within the lineage necessarily comes at a cost to the lineage. Thus David Lahti and I (Lahti and Weinstein, 2005) propose a model in which humans monitor threats to their lineage and exhibit a facultative tendency toward infighting when the lineage is well positioned, and an inverse tendency to pull together and act familially when the lineage is weak or jeopardized. Lineages are kin groups projected in time. 'Lineage selection' is a projection of kin selection deeply into the temporal dimension and, as chapter 4 demonstrates, has the ability to account for the evolution of extreme self-sacrifice, among other traits, without resorting to 'group selection'.

## Literature Cited


Csete, M. E. and Doyle, J. C. (2002) Reverse engineering of biological complexity. *Science* 295: 1664-1668.

Fine, P.V.A., Miller, Z.J., Mesones, I., Irazuzta, S., Appel, H.M., et al. (2006). The growth-defense trade-off and habitat specialization by plants in Amazonian forests. *Ecology, 87*, S150-S162.

Hubbell, S.P. (1997). A unified theory of biogeography and relative species abundance and its application to tropical rain forests and coral reefs. *Coral Reefs, 16*, S9-S21.

Lahti, D.C. and Weinstein, B.S. (2005). The better angels of our nature: group stability and the evolution of moral tension. *Evolution and Human Behavior, 26*, 47-63.

Macarthur, R.H. (1961). Population Effects of Natural Selection. *American Naturalist, 95*, 195-9.

Stearns, S.C. (1992). *The Evolution of Life Histories*. Oxford: Oxford University Press

Weinstein, B.S. and Ciszek, D. (2002). The reserve-capacity hypothesis: evolutionary origins and modern implications of the trade-off between tumor-suppression and tissue-repair. *Experimental Gerontology, 37*, 615-27.

# Chapter 2

# LIFE'S SLOW FUSE: TELOMERES, TUMOURS AND THE EVOLUTION OF VERTEBRATE SENESCENCE

**Abstract**

This paper joins the evolutionary theory of senescence (antagonistic pleiotropy) with recent findings in experimental gerontology. We argue that, in vertebrates, a telomeric fail-safe mechanism prevents tumour formation by limiting cellular proliferation. The same system unavoidably results in the progressive degradation of tissue function with age. This senescent decline is caused by the combined effects of uncompensated cellular attrition and increasing histological entropy, both of which begin at sexual maturity. Extrinsic causes of mortality produce selection that pushes species toward an optimal balance between tumour suppression and tissue repair. With that trade-off as a fundamental constraint, selection adjusts telomere lengths—longer telomeres increasing the capacity for repair, shorter telomeres increasing resistance to tumour formation. In risky environments, where extrinsically induced mortality is frequent, selection against senescence is comparatively weak, generally favouring better tumour suppression and thereby a reduction in telomere lengths. In less dangerous environments selection more strongly opposes senescence, tending to lengthen telomeres. In iteroparous organisms selection further tends to co-ordinate rates of decline between tissues, so that no particular organ generally limits life-span. The implications of these hypotheses for experimental methods are significant and lead us to question the generality of a number of widely cited results. In particular we are concerned that captive-rodent breeding protocols, which are designed to increase reproductive output, simultaneously exert strong selection against senescence and virtually eliminate selection that would otherwise

favour tumour suppression. This unnatural selective regime appears to have had dramatic effects on laboratory mice, elongating their telomeres and, consequently, making them both unusually prone to tumour formation and unusually resistant to senescence. Use of these animals may have confounded the interpretation of numerous experiments. A strain of telomerase-negative mice, for example, required several generations to show signs of accelerated ageing. This generational delay has led some workers to discount the significance of telomere erosion and Hayflick limits in the normal senescence of individuals. The inadvertent elongation of mouse telomeres may also have serious public health implications, as it is likely that safety tests employing these models tend to overestimate cancer risks and underestimate the risk of tissue damage and its natural consequence, accelerated senescence.

**Introduction**

*Why do we get tumours* and *Why do we grow old*? These questions have become an obsession in the biomedical community. But we view them as pieces of a much larger puzzle: How is it possible for a highly differentiated, self-repairing organism composed of millions, billions or trillions of cells to live long enough, in a mutagenic environment, to reproduce, without a single cell escaping the normal developmental program and becoming a deadly tumour? We regard mechanisms that allow for extensive tissue repair, while inhibiting the frequent production of tumours, as major evolutionary innovations—prerequisites to the evolution of life history strategies like those of most vertebrates. We describe one such mechanism by synthesising knowledge from two approaches to vertebrate senescence: evolutionary theory and experimental gerontology.

Historically, these approaches have been practised almost independently. Evolutionists have remained largely unconcerned with the proximate mechanisms of ageing and gerontologists have been lax about the ultimate explanations which underlie their discipline. But comprehension of genetic and cellular machinery has now progressed to the point that evolutionary theory and empirical findings have begun to mirror each other. Not only can these two approaches now be profitably unified, but the resulting synthesis can accelerate progress in both disciplines. For that to occur, evolutionists and gerontologists need a common body of theory and knowledge as well as

a language which allows for meaningful discourse. Below we develop a basic grammar and illustrate the breadth of potential insights that derive from its use.

## I. Synthesising two views of the ageing process

**Senescence: the evolutionary approach.** All else being equal, longer lives provide more reproductive opportunities than shorter lives, therefore natural selection opposes senescence. Compared to the immense challenge of building a self-assembling, ten trillion cell organism (such as a human), the maintenance of such an organism should be relatively simple[1]. Yet selection has failed to eliminate senescence from any vertebrate.

Elaborating on Medawar[2], Williams[1] explained the persistence of senescence as follows: Even in the absence of senescence, all lives would be finite because every organism would ultimately die from accident, starvation, predation or disease. Since an organism is always at risk of death, natural selection should favour early reproductive opportunities over the potential for later ones. Accordingly, the force of natural selection is never stronger than at the typical age of commencement of reproduction (when reproductive potential is greatest) and its strength must decline from that point forward. Therefore, traits that have beneficial effects early in life will tend to spread, even if they are inseparably coupled with deleterious effects that manifest later in life. Individuals are thus endowed with youthful vigour, at the cost of inevitable senescence. The power of "antagonistic pleiotropy" to account for the evolution of senescence was mathematically demonstrated by Hamilton[3].

According to the theory, selection modifies a species' rate of senescence (in response to the distribution of mortality risks across the reproductive life-span) by adjusting pleiotropic balances between longevity and youthful vigour: The greater the risk of death between reproductive opportunities, the stronger the selective bias in favour of youth, the faster the rate of senescence becomes. Giant tortoises, housed in protective shells and living on remote islands, face few hazards during their reproductive lives, producing a rate of senescence that is almost imperceptible. Conversely, spawning salmon would face extreme hazards returning to sea and later attempting a second journey upstream. The very low probability of future reproductive opportunities has

yielded a semelparous life-history strategy in which all resources that can be mobilised are invested in progeny rather than maintenance. These are the extreme cases. Most vertebrates fall on the continuum in-between, selection producing finer adjustments based on parameters that affect risk such as body size, defensive adaptations and the ability of some vertebrates to fly out of harm's way.

Extrinsic hazards can only produce such life-history refinements if selection retains substantial power *during* the process of senescence. Though the force of natural selection declines from the onset of reproduction, it remains strong throughout the reproductive life-span, even as the effects of senescence are becoming increasingly evident. Williams emphasised this point citing the example of the athletic decline that afflicts men in their thirties, observing that "Surely this part of the human life-cycle concerns natural selection."[1]

This point is persistently misunderstood outside of evolutionary biology. In gerontology it is commonly asserted that senescence results from "unselected" late effects of genes (e.g. refs. 4-6). Although the declining force of selection does *eventually* approach zero, that fact is insufficient to explain senescence early in the reproductive life-span, when selection is still very strong[1].

Even in the extreme cases of senescent failures that occur so late that they may actually be inaccessible to selection (such as Alzheimer's disease), the effects are only out of selective reach *because* senescence has already evolved. Extrapolating from the mortality rates of humans on the cusp of maturity, Ricklefs and Finch[7] point out that "…if not for aging, 95% of us would celebrate our centenaries and 50% would reach the seemingly astonishing age of 1200 years" (see also ref. 8, p. 29).

Selection continually minimises deleterious effects that manifest during the period of reproduction and offspring-rearing. If we mistakenly believe that senescence is the product of "unselected" effects, then we may harbour unwarranted hopes for therapeutic reduction of senescence. Conversely, if we view senescence as the unavoidable costs that remain after selection has acted to minimise harmful effects, then we will correctly view senescence as the same daunting challenge for medical science that it has apparently been for natural selection.

***A curious lack of antagonistic pleiotropies.*** In the four decades since antagonistic pleiotropy was first proposed, no clear case of a senescence-causing pleiotropy has come to light in vertebrates. This is odd in view of the fact that unique predictions of the theory have been clearly demonstrated[9-11]. We suspect that the failure to find relevant pleiotropies is the result of confusion over the term itself.

Williams' logic applies to traits with early benefits intrinsically tied to late costs. It is not necessary for the costs and benefits to derive from multiple traits (as a narrow definition of "pleiotropy" might seem to require), they may instead result from a single trait for which the cost/benefit ratio increases with age[12]. Furthermore, the trait need not result from an individual pleiotropic gene; the combined effects of multiple genes may produce emergent costs and benefits that cannot be separated by selection. To fall within the rubric of antagonistic pleiotropy, it is both necessary and sufficient that (1) the instantaneous cost/benefit ratio of a trait or system is initially less than one, becoming greater than one at some point after the onset of reproduction and (2) the cumulative cost/benefit ratio is less than one for the average individual.

**Telomeres and senescence: the experimental approach.** In 1961, Hayflick and Moorhead[13] made an important breakthrough in the experimental study of senescence. They disproved the notion that normal vertebrate cells could divide an indefinite number of times in vitro. They showed that normal somatic cell lines were limited in the number of population doublings they could undergo before growth slowed dramatically, then ceased. Later studies showed that the number of cell divisions occurring before the 'Hayflick limit' co-varies (between taxa) with life-span[14-16] and decreases in humans with cell-donor age[17]. For many years these findings lacked a mechanistic explanation, but a front-runner has now emerged[18] (see also ref. 19).

DNA polymerase is unable to duplicate the tips of chromosomes, so a small amount of DNA is lost with each successive cell division[20,21]. This progressive erosion would be catastrophic if important genes were located at the ends of chromosomes. But the ends of eukaryotic chromosomes consist of long, non-coding, repetitive sequences

known as telomeres[22,23]. Telomere loss may explain the mortality of somatic cell lines as the erosion of telomeres below a critical length appears to trigger the shutdown of replicative machinery[24].

There must also be some means of adding telomeric DNA to chromosome ends, otherwise germlines would be mortal as well. The reverse transcriptase *telomerase* elongates telomeres[22,25-27]. Telomerase is active in gametogenesis and undetectable in the vast majority of adult somatic tissues[28].

Several lines of evidence support the telomere erosion hypothesis for Hayflick limits:

- Telomere length diminishes with cell-line age in vitro and in vivo[29-31].
- A remarkably diverse array of *immortal* somatic cell lines (from tumours, which lack Hayflick limits ) express telomerase[28,32].
- Somatic tissues from patients with Hutchinson-Gilford (H-G) and Werner's syndromes (diseases of dramatically accelerated ageing) have reduced proliferative capacities in vitro. H-G patients have short telomeres at birth[17]. Werner's patients experience rapid erosion of initially normal telomeres[33].

The association of aberrant telomeres with apparently accelerated ageing suggests that Hayflick limits may explain more than just the mortality of individual cell lines. The limited proliferative capacity of somatic cells may underlie a general mechanism of body-wide senescence.

This possibility led to an experiment, which yielded equivocal[34] results. A strain of laboratory mice with two disabled copies of a gene necessary for telomerase activity was produced[35]. This telomerase-negative strain did exhibit accelerated ageing, but only after six generations. Even then, the effect was not uniform. Mice in the sixth generation seemed to senesce prematurely in some tissues and not others. These results strengthened the argument that telomeric erosion is involved in somatic senescence, but suggested that the role of telomeres in the phenomenon of senescence might be limited to those few somatic tissues with high endogenous rates of turnover[36]. The six generation delay was

taken to imply that normal senescence, of the type that occurs in a single generation, must involve important undiscovered factors[34]. We will reinterpret these results below.

***Telomeres and cancer.*** The connection among telomeres, Hayflick limits, and the phenomenon of senescence is important whether telomeres are the primary mechanism or just one of several. But, telomere regulation has significance beyond the issue of our gradual decline with age. Telomere regulation appears central to another great enemy of the old: cancer.

Activation of telomerase appears to be a necessary step in most transformations of normal tissue into tumors[26,28]. The connection of cancer and senescence to the same mechanism is not serendipity, it is a window into a fundamental trade-off, the balance of which we may find difficult to improve.

**The reserve capacity hypothesis: A synthetic search for the missing pleiotropy.** Juxtaposing an evolutionary perspective on senescence, with the gerontological and oncological view of telomeres, it appears that limits on the proliferative capacity of somatic cells (Hayflick limits) evolved as tumour suppressors that rein in runaway cellular proliferation, but that these same limits preclude indefinite somatic maintenance, causing gradual degradation of function. It seems the telomere/telomerase system is an antagonistic pleiotropy of the type Williams[1] predicted.

Kipling[37] briefly proposed a similar interpretation (without reference to Williams' theory), but there has been no apparent discussion of his idea or its implications. Others[6,38] have used the term "antagonistic pleiotropy" in this context, but have evidently failed to appreciate the advances made by Williams[1] over Medawar[2]: the declining force of natural selection with age is not sufficient to explain senescence during prime reproductive years. Only when senescence is recognised as an inherent consequence of design trade-offs can we fully understand the nature of ageing.

A new term facilitates discussion of vertebrate telomeres. We will use *reserve capacity* to refer to the remaining quantity of population doublings that a differentiated

cell can undergo (in vivo) before reaching its Hayflick limit. Reserve capacity decreases over time with cell division.

The relationship of Hayflick limits to tumours is relatively straightforward. When a cell is damaged such that it begins to over-proliferate, it ultimately reaches its Hayflick limit and proliferation ceases. The greater the reserve capacity of the progenitor cell, the larger the resultant mass of growth-arrested daughter cells will be. We regard this mass of cells as a *proto-tumour*, each cell possessing the first of several mutations necessary for tumorigenesis and cancer.

Because cells will tend to retain more proliferative potential early in an organism's life, younger individuals should tend to produce larger proto-tumours than older individuals. Since each cell in a proto-tumour presents an equivalent opportunity for the acquisition of telomerase activating mutations, we predict that proto-tumours produced early in life carry a proportionally higher risk of becoming mature tumours than proto-tumours generated late in life. This effect will be exacerbated by the fact that proto-tumours formed at an early age will tend to have more time in which to accumulate further genetic changes. The risk from any particular proto-tumour should diminish with time, as growth-arrested cells expire and are lost. Risk reduction may be accelerated if apoptosis is triggered by proto-tumour formation, but this would accelerate the exhaustion of the neighbouring lineages that ultimately replace the lost cells.

***Somatic senescence due to cellular attrition and increasing histological entropy.*** To our knowledge, no explicit mechanism linking Hayflick limits to the phenomenon of vertebrate ageing has been proposed. We offer the following first approximation.

Development continually increases histological differentiation and specialisation, which are maximal when an organism becomes a reproductively capable adult. Throughout life, damage and programmed cellular turnover result in cells being lost from the soma and replaced. When cellular lineages exhaust their reserve capacity and are lost, they must be replaced by neighbouring lineages, if they are replaced at all.

We propose that the uncompensated loss of some cellular lineages, coupled with the replacement of other lineages by neighbours (adapted to slightly different roles),

diminishes the optimal arrangement of cell types. By our model, body-wide senescence results from the combined effect of (a) uncompensated cellular attrition and (b) increases in what might be called *histological entropy*, both of which will diminish an organism's efficiency at accomplishing whatever tasks differentiation initially evolved to address. Senescence of this type should progress at a non-linear rate, accelerating with age as fewer cellular lineages maintain and repair an ever larger proportion of the body.

The ageing of human skin appears to progress as our model predicts. Skin thickness decreases approximately 25% between the fourth and eighth decade of life[39], and entropy increases:

> "The epidermis of older individuals exhibits a marked variation in thickness (often in the same histologic section) and a disparity in the size, shape and staining quality of the basal cell nuclei under light microscopy. There is also a loss of the orderly alignment of cells along the basement membrane and a disruption of the gradual upward uniform differentiation present in the epidermis of younger individuals… Electron microscopic studies show that the basal cells of the flattened epidermis of old individuals lack villi… Deletion and derangement of small blood vessels is found in aged skin, with sun-damaged skin being the most severely affected."[40]

Cardiovascular disease may provide an example of the negative *consequences* of uncompensated cellular attrition and increasing histological entropy. Cells in portions of the vascular system that sustain relatively high levels of wear and tear have short telomeres, implying a history of cellular replacement[41] and likely attrition of cellular lineages. These areas fail to produce a protective layer of cells characteristic of younger tissue, and consequently have an increased propensity to develop atherosclerotic plaques[41].

***One source, three sinks.*** Vertebrates use reserve capacity in growth, maintenance, and repair; each process erodes telomeres, reducing proliferative potential. Though *antagonistic pleiotropy* and *accumulated damage* hypotheses have traditionally been

viewed as alternative explanations for senescence, the finite reserve capacity approach integrates them. Damage, even if it is functionally repaired, will accelerate the ageing of tissue by limiting the capacity for future maintenance and repair. The liver of a heavy drinker, for instance, may function essentially as well at 40 as it did at 25, but should fail more rapidly than the liver of a non-drinker, even if alcohol consumption ends before damage is evident. Any factor that damages tissue, including mutagens, pathogens, mechanical wear or trauma, oxidative stress and free radicals, will promote a local increase in that tissue's rate of senescence.

Selection should tend to optimise reserve capacities based on a species' timing of reproduction and the typical rate of cellular repair and turnover as well as the extrinsic risk of mortality. Although telomere erosion begins at whatever point in ontogeny telomerase is inactivated in the soma, selection should adjust reserve capacities so the loss of cellular lineages does not begin before the usual age of first reproduction. In iteroparous species, selection should further act to co-ordinate reserve capacities among tissues so that senescence is synchronised throughout the body, thus minimising the fitness cost that would accompany early senescence in any particular organ (as per refs. 1,3,42).

But, because of the stochastic nature of environmental insults, past selection cannot predict the reserve capacity needs of *individuals* nor the organs on which they depend. An otherwise healthy individual may die from the premature senescence of a particular tissue (despite the synchronising force of selection) if the tissue has had an unusual history of damage. Because rates of damage differ between conspecific individuals, we should also expect dissynchrony of senescence rates between individual animals, even in populations that are genetically homogeneous.

Selection can adjust telomere lengths based on species' averages for parameters such as the number of cells in each tissue of the body and typical rates of damage and mortality. But selection based on averages will not produce ideal telomere lengths for individuals. The optimal telomere length on a chromosome passed from a 5'6" father to his 6'1" son will necessarily be a compromise (longer than optimal for the father and shorter than optimal for the son). This constraint may explain why the positive interspecific correlation between body size and longevity (addressed in ref. 1) is reversed

within species. For example, even when the effects of obesity are controlled for, larger humans[43,44] and dogs[45,46] tend to be comparatively short lived. The extra cell divisions required to become larger, by diminishing reserve capacity at maturity, may shorten life-span by reducing the capacity of larger individuals to maintain and repair their tissues. We expect smaller individuals to suffer a greater per cell risk of developing tumours due to longer-than-optimal telomeres at maturity. At the same time they should show an increased resistance to other senescent effects. Since smaller individuals are composed of fewer cells, we do not expect their increased per-cell tumour risk to fully counteract their decreased rate of senescence. Therefore, within a species, smaller individuals should tend to live longer.

**II. Reinterpreting experimental results:**

**Senescent cellular phenotypes: misregulation or adaptive response?** Upon reaching a Hayflick limit, many cell types begin expressing genes that were previously untranscribed, and cease expression of previously active genes[47]. Several workers have conjectured that somatic senescence of individuals results from the progressive accumulation of cells with "senescent phenotypes"[48-50]. To our knowledge no one has proposed a mechanistic connection between these phenotypes and organismal ageing. Instead, the phenotypic changes are asserted to result from "misregulation". The implicit assumption is that expired cellular lineages accumulate late enough in life that selection lacks the power to regulate their function to the benefit of the organism. We propose a contrary interpretation.

Williams[1] argued that late negative effects would spread if pleiotropically associated with early benefits. He went on to argue that selection would then produce modifiers that would minimise the harm caused by these late effects. We suggest that "senescent cellular phenotypes" are actually adaptations that *limit* the harm caused by the expiration of cellular lineages.

The hypothesis that *changes in gene expression associated with ageing are the result of misregulation* is apparently falsified by the very data used to support it. Ly et al. [50] compared gene expression amongst people from three age classes and children with H-

G progeria. They found that 50% of the genes whose expression is altered in ageing (both accelerated and normal) belonged to two classes, mitosis initiation and progression genes (e.g. spindle formation) and extracellular matrix (ECM) modification genes. If transcriptional changes were the result of misregulation then we should expect a random pattern of changes reflecting a lack of stabilising selection on gene regulation. The fact that regulatory changes were observed in groups of functionally related genes, suggests that the shift in gene expression results from selection rather than a haphazard process.

In addition to the functional relationship between affected genes, the seemingly co-ordinated direction of the regulatory changes was also suggestive of selection. Mitosis-related genes tended to be downregulated with age. This is unsurprising as "senescent" cells, which do not divide, are unlikely to need spindles or other mitotic machinery.

In contrast to the down-regulation of mitosis-related genes, some of the genes which modify the ECM were upregulated and others downregulated. Downregulated genes were primarily associated with construction of the ECM while upregulated genes tended to be associated with its disassembly. This is consistent with earlier findings which suggest that "senescent" cells decrease the production of collagen and increase production of collagenase, an enzyme which breaks down collagen and thereby facilitates the remodelling of the ECM[4].

We propose that selection has produced a system that locally breaks down the ECM as cells are nearing their Hayflick limits, thereby facilitating cellular replacement. Early in life, the ECM maintains the developmentally optimal placement of cells. But in some circumstances this system may act as an impediment to cell motility. As cellular lineages become unable to replace themselves, adjacent lineages may not be able to fill vacated spaces with the ECM intact. Selection may have programmed senescent cells to locally dismantle the ECM, paving the way for their eventual replacement by adjacent lineages.

**Lab mice and cloned sheep: life on strange islands.** If individuals disperse from a high risk environment to a low risk environment (e.g. a remote island) the resultant increase in

longevity will enhance the potency of selection on late-life effects, eventually slowing the rate of senescence[1]. The evidence that selection does this in the wild is strong[9-11]. We expect that, in vertebrates, selection adjusts telomere lengths to postpone senescence under such circumstances. This adjustment must come at some cost, such as increased risk of tumours and/or an increased burden from larger proto-tumours.

In the early part of this century, a small number of *Mus musculus* dispersed into a novel environment: the laboratory. In breeding colonies there is no predation, no resource limitation and the spread of pathogens and contaminants is controlled. Perhaps most importantly, breeders are retired at 8 months[51] so the mice that contribute most to future generations are those that begin reproduction early, and sustain a high rate of reproduction until the cut-off age. Such conditions are dramatically different than those in the environment mice originally evolved to exploit, likely favouring a different pattern of senescence.

The telomere systems of laboratory mice are hard to reconcile with the notion of Hayflick limits as tumour suppressors or as the cause of senescence. Compared to humans, lab mice have "ultra-long" telomeres, exceeding human telomeres by an order of magnitude[52]. Further, somatic tissues of lab mice produce telomerase, and can "spontaneously immortalise" in culture.

We predicted that the long telomeres observed in laboratory mice would be atypical for mice in general. Greider's lab tested this with a survey of telomere lengths in six species of laboratory mice with short histories of captivity. All six had telomere lengths approximately one tenth of those in common lab mice (C. Greider, pers. com.).

The unusual telomere system of lab mice may be an unintended consequence of captive breeding. Retirement of breeders after 8 months eliminates selection on late-life effects. Tumour-forming mutations take time to occur, and the likelihood of tumour initiation is presumably a function of the number of cells in the body, so in small bodied animals like mice, tumours may be rare in the first eight months of life, even without the telomeric fail-safe. Further, selection for sustained high reproductive output before 8 months should tend to eliminate any senescent effects occurring before that deadline. Selection acting to simultaneously increase early reproductive output and eliminate senescent effects may elongate telomeres. Because of the inextricable connection

between tumour suppression and somatic maintenance, telomere elongation should dramatically increase the risk of tumour formation, but any effects occurring after the breeding cut-off will be selectively irrelevant. Selection for early high rates of reproduction in the absence of selection for longevity should result in a strong propensity for these mice to eventually die from tumours. At all ages, lab mice should be more likely to die of tumours than wild mice raised in similar environments. Lab mice should also be unusually resilient to somatic damage and show few signs of ageing other than tumour formation. Alexander[53] presents evidence consistent with this pattern:

> "The most striking fact is that even very old [lab] mice (e.g., more than 2.5 years) when killed while still fit have remarkably few pathologies and are almost indistinguishable from young animals."

The hypothesis that an 8 month breeding cut-off should select for non-senescent, tumour prone mice seems, at first, paradoxical. Based on traditional evolutionary thinking, one might expect the elimination of selection on late life effects to *accelerate* senescence, not retard it. But in the case of mice, selection for high, sustained rates of breeding appears to be the dominant factor. This would likely not be the case in much larger vertebrates (which are necessarily composed of much larger populations of cells). Each cell that retains the ability to divide has the potential to become a tumour. Thus, *in the absence of a scaleable tumour suppressor,* the chance that a tumour will form and kill an organism is a function of the number of cells in that organism. In lab mice, the tumour suppressor has effectively been turned off, condemning them to form tumours but leaving an early-life window of reproduction within which there is minimal senescent decline. In a much larger organism, turning off the tumour suppressor would likely create such an onerous tumour-burden that reproductive age would never be attained. This is especially likely since larger animals tend to reach reproductive maturity later, increasing the chance that a tumour will arise before reproduction begins.

Unfortunately, it has been widely assumed and asserted that "ultra-long" telomeres are characteristic of "mice" or even "rodents" in general. The stark differences between lab mice and humans led de Lange [19] to argue:

> "…it seems very unlikely that mice use telomeres as a tumor suppressor system and perhaps with good reason. Since the telomere barrier to proliferation does not manifest itself until many cell divisions have passed, this mechanism may not be useful for a small animal in which a 2cm mass of misplaced cells could be life-threatening."

We agree that the telomere system of small animals would need to arrest very small growths to serve as a useful tumour suppressor, but the conjecture that "mice" do not use this system is premature. The tissues of wild mice might have very limited reserve capacities, thus protecting them from lethal growths and limiting their life-spans.

It is unfortunate that mouse strains with long telomeres were used to create the telomerase-negative mice. If the experiment were conducted using mice recently derived from the wild we predict that accelerated senescence would be observed in the first generation. But even in such an experiment we expect that the acceleration of gross senescent effects would be limited to high-turnover tissues because other tissues, which typically use reserve capacity to repair damage, will tend to senesce minimally in a protected environment.

The unique state of lab mice may lead to erroneous conclusions about tumorigenesis. For example, based on evidence from mice with ultra-long telomeres, Kipling[54] speculates that "...telomerase expression in mouse tumorigenesis is an innocent bystander rather than a necessary event." Clearly, telomerase activity, telomere length regulation and spontaneous immortalization must be investigated in newly domesticated mice to separate experimental artefacts from natural phenomena.

Care must also be taken in interpreting the pattern of ageing in animals produced through nuclear transfer cloning, such as the sheep Dolly. The nucleus that was used to produce Dolly was taken from an adult sheep[55], and thus had shorter telomeres than a normal sheep zygote, though as yet Dolly does not appear to be senescing abnormally[56]. Like lab mice, Dolly lives in a controlled environment, protected from the traumas, illnesses and impurities of a wild or even a typical farm habitat. We expect Dolly to senesce earliest in tissues with high endogenous turnover rates (because her need for

damage repair is likely to be minimal), and to display early senescence compared to sexually produced controls reared in the same protected environment. But compared to normal sheep, her senescence may not appear accelerated, as it is likely being slowed by her isolation from environmental insults.

**Retarding senescence with caloric restriction: natural phenomenon or laboratory artefact?** Caloric restriction (CR) is the only experimental treatment shown to dramatically increase longevity in vertebrates. Laboratory mice and rats placed on a restricted diet live significantly longer than controls[57]. This has been interpreted as evidence that resource limitation slows the process of senescence. But if, as we suggest, these animals have been selected to senesce minimally, then slowing senescence should have little effect on their longevity.

We have argued that laboratory mice should overwhelmingly die of tumours. CR may increase longevity in these animals by reducing the risk of tumour formation. CR animals are approximately 1/3 smaller[58] and exhibit slower cell replication[59] than controls. By stunting growth (reducing the number of cells in the organism), and by reducing the rate of cell division, CR may simply reduce the likelihood of tumorigenesis.

Such positive effects might also occur in CR vertebrates with wild-type telomeres. Further, by reducing body size, delaying maturation, inhibiting reproductive mechanisms and slowing cellular turnover, CR should postpone the exhaustion of reserve capacity. But CR, like famine, will likely interfere with normal homeostasis and repair, increasing vulnerability to environmental insults. The opposing nature of these effects will likely prevent CR from dramatically increasing longevity in vertebrates with normal telomeric tumour suppressors.

## III. Selective inactivation of the telomeric tumour suppressor

**The counterintuitive nature of early development.** If finite reserve capacity is an evolved fail-safe against runaway cellular lineages, we must give special consideration to those times and places where selection has disabled this mechanism. Telomerase is present in the somatic tissues of embryonic placental mammals, but activity ceases before

birth[38,60,61]. To illustrate why selection responds differently to telomere erosion in early versus late development, we will compare the distinct developmental profiles of two processes: cellular population doubling and resource investment.

In the absence of telomerase, telomere loss is a function of the number of cell population doublings, therefore division of a zygote into two cells would reduce the mature body's reserve capacity as much as the growth of a 5 trillion cell child into a 10 trillion cell adult (if all cells made an equivalent contribution to growth). The vast majority of cellular doublings occur in early development when the absolute number of cells is very small and the embryo is tiny compared to the adult it will become. In contrast, parental investment of resources is *lowest* in early development and grows with the embryo's size. Because of this asymmetry, the resources placed at risk by early foetal telomerase activity are minimal.

Spontaneous abortions are common in early foetal development, ending approximately 50% of human pregnancies[62]. Early spontaneous abortions are not without cost. In many species breeding periods are narrowly timed and an aborted pregnancy may eliminate a female's reproductive output for the year. The cost of early spontaneous abortions has apparently resulted in mechanisms that reduce such risks. For example, maternal aversion to complex foods during early stages of pregnancy is thought to protect the embryo/foetus from mutagens during an especially vulnerable period (reviewed in ref. 63). We propose that isolating the foetus from mutagens is particularly important while telomerase is active, when runaway cellular proliferation would necessarily result in abortion.

In humans the majority of prenatal cell divisions occur before the end of the fifth month of gestation, while telomerase is maintaining telomere lengths. The period of telomere maintenance ends, on a tissue-by-tissue basis, beginning in the fourth month and continuing through the fifth month[38,60,61]. After this point the foetus begins to accrue resources in the form of body fat. In contrast to the rate of cell addition, which peaks in the fifth month, then drops precipitously[64], the great majority of prenatal weight gain occurs in the later, telomerase-negative period. We interpret this developmental pattern as a mechanism by which selection has minimised the resources placed at risk by developmental telomerase activity.

Though early telomerase activity carries risks, a lack of telomerase during the period of rapid cellular doublings would result in a substantial erosion of the telomeres, accelerating the onset and rate of senescence later in life. It seems selection could solve this problem by lengthening germline telomeres, thus adding reserve capacity to the organism as a whole. Because selection has favoured telomerase activity (and its associated risks) over a simple lengthening of telomeres, we expect foetal telomerase activity also provides a significant benefit.

The nature of that benefit may relate to Williams' argument that selection should tend to synchronise senescence across the soma[1]. If finite proliferative capacities determine the senescence rates of different tissues, and if those rates are to be synchronised by selection, telomere lengths must be adjusted according to the typical rates of cellular turnover of different parts of the soma. Simply lengthening germline telomeres could not produce this synchronisation. If telomerase were never active in the soma, the reserve capacity of a particular tissue would simply be an inverse function of the total number of cell divisions that produced it from the zygote. In contrast, tissue-specific telomerase activity can establish inter-tissue synchronisation. This could be accomplished at any point in the lifecycle, but it is least costly in early development when (1) the investment placed at risk is minimal, (2) the foetus is insulated from most environmental mutagens, and (3) the number of potential runaway cells is relatively small.

The reserve capacity of mature tissues can be set by adjusting the number of cells in each tissue before telomere maintenance ceases. If it is demonstrated that organ senescence is prenatally synchronised in such a manner, it will firmly establish that patterns of senescence are products of natural selection, not incidental effects that occur in the absence of selection.

After somatic telomerase is shut down, growth via cell division will reduce tissue reserve capacity. Wistar rats that were growth-retarded prenatally (i.e. during telomere maintenance), but grew to normal size after birth, had shorter telomeres in their kidneys and shorter life-spans than control rats[65]. Among humans, women that were short at birth but grew to average or above average height had an increased risk of death from coronary

heart disease[66]. A similar pattern appears to exist in men[67], though it is confounded by mortality risks associated with obesity rather than "catch-up growth".

**Cellular over-proliferation in early and late life: tumours of two natures.** If the shortening of telomeres is part of an adaptive tumour suppressor mechanism, why are tumours most common late in life, when telomeres are shortest? Tumours may be divided into two classes: (1) tumours which arise when telomere lengths are exceedingly long or are being maintained by telomerase (these can occur at any point in the life-span, including childhood); and (2) tumours arising when telomeres have become critically short (primarily late in life). Reserve capacity limitation appears to counter early-life tumours so successfully that we may fail to realise that a serious threat would otherwise exist. The few systems in which telomere lengths are maintained provide a window into life without the telomeric fail-safe.

***Childhood leukaemia and lymphoma.*** Most of the tumours common in the elderly are essentially unknown in children and young adults. The most common childhood tumours, leukaemias and lymphomas, involve hyper-proliferative leukocytes (B- and T-cells) or their progenitors. Leukocytes are responsible for our specific immune response which "learns" to recognise pathogens. When a leukocyte is activated by a matching antigen, the cell proliferates, creating a population of cells with variations of the progenitor cell's receptor formula. Iterated clonal selection allows the system to hone in on unfamiliar pathogens and to track antigenic changes in an ongoing infection[68].

Although most leukocytes will never be stimulated, the subset that become activated must retain an extensive capacity to proliferate. Otherwise pathogens for which immune cells have initially weak affinities would remain elusive, and rapidly changing pathogens could exhaust the proliferative potential of the immune system. Instead, leukocytes produce telomerase upon antigenic activation, allowing for extensive proliferation[69,70]. We suspect that telomerase activity, which is necessary to the functioning of the immune response, but greatly diminishing the effectiveness of the

telomeric failsafe, results in a disproportionate childhood risk of developing leukaemias and lymphomas.

***Germline tumours.*** Testicular cancer is essentially absent in boys, but beginning at puberty (when gametogenesis begins) the incidence of testicular germ cell tumours jumps, peaking between the ages of 20 and 34[71]. Germline tissue does not senesce, so spermatogenic cells must maintain their telomeres throughout life, despite undergoing very high rates of cellular proliferation. Spermatogenic cells lack a telomeric failsafe, since they produce telomerase during gametogenesis[28], which likely explains the disproportionate occurrence of testicular cancer in young men. In female mammals gametogenesis occurs before birth, so there is no increase in risk of germ cell tumours with puberty. Indeed, minimisation of the fitness costs associated with germline tumours may account for the evolutionary shift of female gametogenesis to the prenatal period.

***Tumours late in life.*** Late-life tumours can arise by at least two pathways. A proto-tumour cell which earlier became developmentally insensitive to signals halting growth, may later gain a mutation that activates telomerase. This is statistically unlikely in any individual cell, but since the many cells in a proto-tumour will all carry the initial growth-stimulating mutation(s), the risk that one will gain an additional mutation increases with the proto-tumour's size.

The second pathway does not depend on telomerase or a population of cells at increased risk. Typically cells cease proliferation when telomeres become critically short. But a cell carrying a mutation that prevents such arrest may continue to divide and erode the telomere *below* the threshold necessary to stabilise the chromosome ends. When that occurs, chromosomes become unstable and fuse into closed structures[72]. Such chromosome instability has dramatic, unpredictable effects and may lead to excessive growth even in the absence of telomerase. For example, the erratic telomere shortening and resultant chromosomal aberrations characteristic of Werner's syndrome results in both tumorigenesis and accelerated senescence.

**A senescence 'rescue' mechanism: reactivation of telomerase in failing tissues.** Telomerase is believed to be inactive in nearly all healthy somatic tissues of adults, but we suspect this is a significant oversimplification. Selection should balance the risk posed by the early senescence of disproportionately damaged tissues against the risk of tumorigenesis. If relatively early senescence of a tissue threatens the survival of the individual, local activation of telomerase may be a worthy risk. If the exhaustion of cellular reserve capacities was not due to hyper-proliferation, then telomerase can extend the life of a failing tissue. We predict the existence of such a 'rescue' mechanism. However, if the rescued section includes a proto-tumour, telomerase activation will likely result in tumorigenesis. We expect localised activation of telomerase to increase with age (as the proportion of the body threatened by imminent senescence becomes increasingly large), and only a subset of such activation to be associated with tumours.

***Reinventing the veal: Novel effects in the cloning of calves.*** Lanza et al.[73] found that calves cloned from "senescent" fibroblasts were born with unusually long telomeres. This counterintuitive result suggests the possibility that the use of senescent cells may have inadvertently triggered the rescue mechanism during cloning or development. In any case, we predict that these cloned calves with long telomeres will have increased cancer rates compared to sexually produced calves raised in a similar environment, and will otherwise exhibit relatively delayed senescence.

***Explaining rapid decline in Hutchinson-Gilford progeria.*** The failure of telomerase reactivation may be relevant to Hutchinson-Gilford syndrome. H-G progeria is a homozygous recessive condition[74,75] which we predict results from two inactive copies of a gene necessary for telomerase functionality. Without telomerase, the erosion of telomeres during early development would be substantial, and could account for the abnormal ontogeny and early onset of senescence in H-G patients. The inability to rescue senescent tissues by selectively reactivating telomerase may account for the rapid decline of H-G patients compared to normal elderly people. Consistent with our theory, H-G patients are not known to get cancer[76].

**Telomerase activity in epithelial tissues.** Several types of basal epithelial cells (which must proliferate extensively for normal functioning) express telomerase (reviewed in ref. 77). Yet basal layers are not a common source of tumours in young people. There are at least two reasons: first, the basal layer is protected from superficial contact with environmental mutagens. Second, progeny of the basal cells are sloughed from the body regularly, likely purging hyper-proliferative cells from these tissues before they become a danger[78].

## IV. Conclusions

**An optimal window of reproductive opportunity.** Decreasing the rate of human senescence and the threat posed by tumours are desirable medical goals. Shay and Wright[79] (also see ref. 80) have outlined a research plan to accomplish both:

> "The key issue is to find out how to make our cancer cells mortal and our healthy cells immortal, or at least longer lasting. Inhibition of telomerase in cancer cells may be a viable target for anti-cancer therapeutics while expression of telomerase in normal cells may extend lifespan."

This illustrates the danger of isolating medical research from evolutionary biology. If one believes that senescence results from a lack of selection, then it may seem reasonable to pursue a technological solution to fill in where selection left off. But evolutionary theory indicates that senescence results primarily from trade-offs, not from incidental effects or a failure of selection. Once we recognise that longevity and tumour suppression are antagonistic goals, the first question we should seek to answer is: *How well has selection optimised the balance between these traits?*

It is not clear that selection has left much room for improvement. We suggest that a staggering majority of our proto-tumour cells are already mortal, allowing only a miniscule risk of tumorigenesis in the first four decades of life. And it is likely that selection has already extended our life-spans by modifying telomere lengths and co-ordinating the reserve capacities among our various tissues. It is a reasonable guess that

maximum longevity cannot be greatly extended without a dramatic increase in the rate of tumour formation, and that increasing the effectiveness of telomeric tumour suppression would accelerate the ageing process.

**Medical applications.** If a simple modification of telomere-system parameters would extend life without significant costs, selection would surely have made it. We are therefore skeptical of attempts to favourably modify telomere regulation in healthy people. But this does not imply that medical benefits cannot be derived from telomere regulation. In fact, such knowledge holds great medical promise. Telomerase treatment, in vitro, may rejuvenate tissues or organs before transplant, extending telomeres in accordance with the amount of cell division expected to occur in the recipient (but see ref. 81). This has been suggested for bone marrow transplants[82] and may be particularly useful for liver transplants in which fractions of a divided liver grow to normal size in multiple recipients.

Further, replacement tissues could be grown from a person's own cells, in the presence of telomerase, to provide a patient threatened by the premature senescence of a tissue with an MHC-matched replacement. This technique might be useful in treating early-stage HIV patients. HIV-reactive T-cells could be removed early in the course of infection, maintained in vitro, and treated with telomerase. When the in vivo T-cell count begins to crash, the invigorated cells could be reintroduced into the patient where they might greatly extend the latent phase of HIV.

Finally, given our increasing ability to detect and surgically or chemically eliminate tumours, we *might* one day be willing to accept an increase in our tumour risk in order to extend youth. The in vitro lengthening of zygote telomeres would likely produce that heritable effect.

Avenues of research likely to lead to viable therapies are those to which natural selection has not had access (e.g. surgery and in vitro methodologies). The idea that medical science will improve the cell-by-cell regulation of telomerase in healthy people, thereby extending youth while at the same time reducing cancer risks, is wishful thinking of the highest order.

**Antagonistic pleiotropy: a theory in retrospect.** The belief that senescence evolves because the harmful effects of genes are invisible to selection late in life, and thus accumulate by drift, is inadequate to account for the senescence of iteroparous organisms. Despite Williams'[1] elucidation of this point, a chronic confusion on this issue persists. This oversight has important implications for present and future work, implications which are brought into stark relief by errors in telomere research. A focus on drift as a causal agent has produced misinterpretations of empirical patterns (e.g. senescent cellular phenotypes) and may have obscured others (e.g. developmental co-ordination of reserve capacities between tissues).

Most importantly, a failure to understand the active way that environmental hazards selectively adjust patterns of senescence has resulted in a haphazard breeding strategy for model organisms such as mice. Inadvertent selection has altered our model systems in ways that obscure the very patterns we most wish to understand.

Not only are our model organisms unfit for studies of ageing, but because they have extraordinary reserve capacities, their use in the safety testing of drugs, pesticides and other chemical agents is likely to drastically underestimate somatic damage. Toxins which damage tissues, hastening the attrition of cellular lineages and thereby accelerating organ degeneration, may appear harmless when administered (even in high doses) to mice with telomeres long enough to last six generations. The same substance may produce irreversible effects in humans, which we may fail to recognise if they manifest after a delay of many years and appear similar to normal effects of ageing. We should therefore reconsider the use of substances deemed safe primarily because they proved harmless to "mice". At the same time, safety testing with lab mice may tend to overestimate cancer risks.

Based on the above analysis, we regard the theory of senescence developed by Medawar[2], Williams[1] and Hamilton[3] as remarkably foresighted and accurate. But as one might expect, there are a number of ways in which the nature of the reserve capacity mechanism is unexpected. Perhaps most significant, evolutionary theory predicted that senescence would be the result of a large collection of distinct pleiotropic effects acting across the soma. It was hypothesised that selection would come to synchronise these

many effects such that they would degrade somatic function across the body at an even rate, so no effect would itself be disproportionately costly.

We cannot rule out the possibility that there are a large number of senescence causing pleiotropies yet to be discovered. In fact, the unitary nature of development virtually ensures that there are. But it seems that the tissue by tissue adjustment of reserve capacity may have effects across the soma that match the expectations of Williams and Hamilton, without the need to invoke many independent genes. Are tumours such a potent selective force in vertebrates that tumour suppression alone could account for senescence by the logic of antagonistic pleiotropy? We suspect so.

We also suspect that selection's ability to regulate gene expression differentially with age allows the late-life silencing of many pleiotropic genes that may have originally contributed to senescence. It is therefore conceivable that only a few antagonistic pleiotropies continue to produce large enough senescent effects to be easily measured. Williams[1] foreshadowed this argument in his discussion of the evolution of senescence-reducing modifiers, but may have underestimated its significance.

In addition, we now question a potentially falsifying prediction offered Williams[1]. He hypothesised that no individual can have both an unusually vigorous youth and an unusually long life. We agree that no individual should be genetically predisposed to both, but an individual with long telomeres may exhibit slow senescence accompanied by an increased risk of tumour formation yet, by chance, not acquire mutations leading to cancer. We suggest that this particular prediction did not allow for the prominent role that environmental stochasticity plays in the senescence equation.

It is interesting that the hypothesised trade-off between youthful vigour and longevity does not map directly onto the reserve capacity hypothesis. Though tissue repair likely controls longevity in vertebrates, tumour suppression is not the same as "youthful vigour". It is possible that we have overlooked some feature of this system that does affect vigour. The accumulation of proto-tumours, for example, may have significant negative effects on organismal efficiency. If so, decreases in initial reserve capacities would result in smaller proto-tumours thus increasing vigour at a cost to longevity. More likely, "youthful vigour" is an unnecessarily restrictive concept and *any* fitness enhancing benefit in youth should be included on the vigour side of the trade-off.

Finally, we are intrigued by the implications of the reserve capacity hypothesis for the correlation between body size and longevity. We must stress that Williams[1] provides an elegant explanation for the widespread tendency of larger animals to live longer lives than smaller animals. Williams' argued that smaller organisms are subject to being eaten by a larger variety of predators, thus in small-bodied animals selection for longevity tends to be weaker, increasing rates of senescence. The exceptions to the general body-size trend clearly demonstrate the power of Williams' explanation. Birds and bats, for example, tend to be extremely long lived for their given sizes. This is presumably due to the anti-predator benefits associated with the ability to fly. But our analysis of reserve capacity suggests that, in vertebrates, a more fundamental selective effect may have acted in concert with predation to favour the evolution of the general body-size/longevity trend. As de Lange [19] implies, the smaller an organism is, the more threatened it will tend to be by a misplaced growth of some size. Because cell size does not increase with body size, effective tumour protection requires small vertebrates to arrest runaway growths at a smaller absolute number of cells. Therefore reserve capacity must be kept low, limiting the number of replacement cells available for each primary cell in the adult organism. This issue of allometry may underlie the tendency of body-size to correlate with longevity, and leads us to predict that birds, bats and other vertebrates disproportionately long-lived for their size will also be disproportionately burdened by larger proto-tumours and more frequent tumours.

**Perspective**. In the popular imagination, humanity is in a race to "reverse" human ageing and to "cure" cancer. Recent discoveries about the connection of telomeres to senescence and tumour suppression have fuelled speculation that we may be on the brink of accomplishing one or both tasks. We are not. Tumorigenesis is an ever-present threat to any large, highly differentiated, self-repairing organism. Proliferative limits provide a tumour failsafe, the inevitable cost of which is the gradual failure of our ability to repair damage. Though it may seem we are a species plagued by senescence and tumours, we are in fact the beneficiaries of selection's remarkable efficiency at simultaneously minimising the harm of these two opposing hazards.

**Acknowledgements**

This paper was submitted to *Nature* on August 1st, 2000, co-authored with Debbie Ciszek (submission number W08077). After it was declined review at *Nature*, a revised version was invited for publication at *Experimental Gerontology*. The resulting paper was published as follows:

> Weinstein, B. and Ciszek, D. (2002). The reserve-capacity hypothesis: evolutionary origins and modern implications of the trade-off between tumor-suppression and tissue-repair. *Experimental Gerontology, 37*, 615-27.

The acknowledgements that appeared in the original submission are as follows: The authors are grateful for the guidance of R. Alexander, S. Austad, P. Coley, R. Dawkins, C. Greider, M. Hemann, H. Heying, L. Howard, N. Kurnick, T. Kursar, D. Lahti, D. Marshall, D. Molbo, C. Murren, R. Nesse, W. Potts, A. Richards, N. Shefferly, G. Smith, F. Solomon, E. Weinstein, K. Weinstein, and G. Williams.

## Literature Cited


1. Williams, G. C. Pleiotropy, natural selection, and the evolution of senescence. *Evolution* **11**, 398-411 (1957).
2. Medawar, P. B. *An Unsolved Problem in Biology* (H. K. Lewis, London, 1952).
3. Hamilton, W. D. The moulding of senescence by natural selection. *Journal of Theoretical Biology* **12**, 12-45 (1966).
4. Campisi, J. The biology of replicative senescence. *European Journal of Cancer* **33**, 703-709 (1997).
5. Campisi, J. Aging and cancer: The double-edged sword of replicative senescence. *Journal of the American Geriatrics Society* **45**, 482-488 (1997).
6. Harley, C. B. in *Telomeres and Telomerase* (eds. Chadwick, D. J. & Cardew, G.) 129-144 (John Wiley & Sons, Chichester, 1997).
7. Ricklefs, R. E. & Finch, C. E. *Aging: A Natural History* (Scientific American Library, New York, 1995).
8. Finch, C. E. *Longevity, Senescence, and the Genome* (University of Chicago Press, Chicago, 1990).
9. Austad, S. N. Retarded senescence in an insular population of Virginia opossums (*Didelphis virginiana*). *Journal of Zoology* **229**, 695-708 (1993).
10. Reznick, D. N. Life history evolution in guppies (*Poecilia reticulata*): Guppies as a model for studying the evolutionary biology of aging. *Experimental Gerontology* **32**, 245-258 (1997).
11. Ricklefs, R. E. Evolutionary theories of aging: Confirmation of a fundamental prediction, with implications for the genetic basis and evolution of life span. *American Naturalist* **152**, 24-44 (1998).
12. Alexander, R. D. *The Biology of Moral Systems* (A. de Gruyter, Hawthorne, N. Y., 1987).
13. Hayflick, L. & Moorhead, P. S. The serial cultivation of human diploid cell strains. *Exp. Cell Res.* **25**, 585-621 (1961).
14. Hayflick, L. Biology of human aging. *American Journal of the Medical Sciences* **265**, 432-445 (1973).
15. Hayflick, L. Aging under glass. *Mutation Research* **256**, 69-80 (1991).
16. Rohme, D. Evidence for a relationship between longevity of mammalian species and life spans of normal fibroblasts in vitro and erythrocytes in vivo. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **78**, 5009-5013 (1981).

17. Allsopp, R. C. et al. Telomere length predicts replicative capacity of human fibroblasts. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 10114-10118 (1992).

18. Chiu, C.-P. & Harley Calvin, B. Replicative senescence and cell immortality: The role of telomeres and telomerase. *Proceedings of the Society for Experimental Biology & Medicine* **214**, 99-106 (1997).

19. de Lange, T. Telomeres and senescence: Ending the debate. *Science* **279**, 334-335 (1998).

20. Olovnikov, A. M. A theory of marginotomy: The incomplete copying of template margin in enzymic synthesis of polynucleotides and biological significance of the phenomenon. *Journal of Theoretical Biology* **41**, 181-190 (1973).

21. Watson, J. Origin of concatameric T7 DNA. *Nat. N. Biol.* **239**, 197-201 (1972).

22. Blackburn, E. H. Telomeres. *Trends in Biochemical Sciences* **16**, 378-381 (1991).

23. Moyzis, R. K. et al. A highly conserved repetitive DNA-sequence, (TTAGGG), present at the telomeres of human chromosomes. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 6622-6626 (1988).

24. Griffith, J. K. et al. Reduced telomere DNA content is correlated with genomic instability and metastasis in invasive human breast carcinoma. *Breast Cancer Research & Treatment* **54**, 59-64 (1999).

25. Blackburn, E. H. Telomerases. *Annual Review of Biochemistry* **61**, 113-129 (1992).

26. Feng, J. et al. The RNA component of human telomerase. *Science* **269**, 1236-1241 (1995).

27. Weinrich, S., L. et al. Reconstitution of human telomerase with the template RNA component hTR and the catalytic protein subunit hTRT. *Nature Genetics* **17**, 498-502 (1997).

28. Kim, N. W. et al. Specific association of human telomerase activity with immortal cells and cancer. *Science* **266**, 2011-2015 (1994).

29. Harley, C. B., Futcher, A. B. & Greider, C. W. Telomeres shorten during aging of human fibroblasts. *Nature* **345**, 458-460 (1990).

30. Hastie, N. D. et al. Telomere reduction in human colorectal-carcinoma and with aging. *Nature* **346**, 866-868 (1990).

31. Lindsey, J., McGill, N. I., Lindsey, L. A., Green, D. K. & Cooke, H. J. In-vivo loss of telomeric repeats with age in humans. *Mutation Research* **256**, 45-48 (1991).

32. Counter, C. M., Hirte, H. W., Bacchetti, S. & Harley, C. B. Telomerase activity in human ovarian carcinoma. *Proceedings of the National Academy of Sciences of the United States of America* **91**, 2900-2904 (1994).

33. Faragher, R. G. A. et al. The gene responsible for Werner syndrome may be a cell division "counting" gene. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 12030-12034 (1993).

34. Rudolph, K. L. et al. Longevity, stress response, and cancer in aging telomerase-deficient mice. *Cell* **96**, 701-712 (1999).

35. Blasco, M. A. et al. Telomere shortening and tumor formation by mouse cells lacking telomerase RNA. *Cell* **91**, 25-34 (1997).

36. Lee, H. W. et al. Essential role of mouse telomerase in highly proliferative organs. *Nature* **392**, 569-574 (1998).

37. Kipling, D. *The Telomere* (Oxford University Press, Oxford, 1995).

38. Wright, W. E., Piatyszek, M. A., Rainey, W. E., Byrd, W. & Shay, J. W. Telomerase activity in human germline and embryonic tissues and cells. *Developmental Genetics* **18**, 173-179 (1996).

39. Black, M. A modified radiographic method for measuring skin thickness. *Br. J. Dermatol.* **81**, 661 (1969).

40. Balin, A. K. in *Human Biologic Age Determination* (ed. Balin, A. K.) 343-373 (CRC Press, Boca Raton, 1994).

41. Chang, E. & Harley Calvin, B. Telomere length and replicative aging in human vascular tissues. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 11190-11194 (1995).

42. Maynard Smith, J. The causes of ageing. *Proceedings of the Royal Society of London, Series B* **157**, 115-127 (1963).

43. Samaras, T. T. & Elrick, H. Height, body size and longevity. *Acta Medica Okayama* **53**, 149-169 (1999).

44. Samaras, T. T. & Storms, L. H. Impact of height and weight on life-span. *Bulletin of the World Health Organization* **70**, 259-267 (1992).

45. Li, Y., Deeb, B., Pendergrass, W. & Wolf, N. Cellular proliferative capacity and life span in small and large dogs. *Journal of Gerontology* **51A**, B403-B408 (1996).

46. Michell, A. R. Longevity of British breeds of dog and its relationships with sex, size, cardiovascular variables and disease. *Veterinary Record* **145**, 625-629 (1999).

47. Smith, J. R. & Pereira-Smith, O. M. Replicative senescence: Implications for in vivo aging and tumor suppression. *Science* **273**, 63-67 (1996).

48. Campisi, J., Dimri, G. & Hara, E. in *Handbook of the Biology of Aging* (eds. Schneider, E. L. & Rowe, J. W.) 121-149 (Academic Press, San Diego, 1996).

49. Johnson, F. B., Sinclair, D. A. & Guarente, L. Molecular biology of aging. *Cell* **96**, 291-302 (1999).

50. Ly, D. H., Lockhart, D. J., Lerner, R. A. & Schultz, P. G. Mitotic misregulation and human aging. *Science* **287**, 2486-2492 (2000).

51. National Research Council. *Mammalian Models for Research on Aging* (National Academy Press, Washington, D.C., 1981).

52. Kipling, D. & Cooke, H. J. Hypervariable ultra-long telomeres in mice. *Nature* **347**, 400-402 (1990).

53. Alexander, P. in *Perspectives in experimental gerontology* (ed. Shock, N. W.) 266-279 (Charles C. Thomas, Springfield, Illinois, 1966).

54. Kipling, D. Mammalian telomerase: Catalytic subunit and knockout mice. *Human Molecular Genetics* **6**, 1999-2004 (1997).

55. Campbell, K. H. S., McWhir, J., Ritchie, W. A. & Wilmut, I. Sheep cloned by nuclear transfer from a cultured cell line. *Nature* **380**, 64-66 (1996).

56. Shiels, P. G. et al. Analysis of telomere lengths in cloned sheep. *Nature* **399**, 316-317 (1999).

57. Ross, M. H. in *Nutrition and aging* (ed. Winick, M.) (John Wiley & Sons, New York, 1976).

58. Sprott, R. L. & Austad, S. N. in *Handbook of the biology of aging* (eds. Schneider, E. L. & Rowe, J. W.) 3-23 (Academic Press, San Diego, 1996).

59. Wolf, N. S., Penn, P. E., Jiang, D. Z., Fei, R. G. & Pendergrass, W. R. Caloric restriction: Conservation of in vivo cellular replicative capacity accompanies life-span extension in mice. *Experimental Cell Research* **217**, 317-323 (1995).

60. Ulaner, G. A. & Giudice, L. C. Developmental regulation of telomerase activity in human fetal tissues during gestation. *Molecular Human Reproduction* **3**, 769-773 (1997).

61. Yashima, K. et al. Expression of the RNA component of telomerase during human development and differentiation. *Cell Growth & Differentiation* **9**, 805-813 (1998).

62. Stabile, I., Grudzinskas, G. & Chard, T. (eds.) *Spontaneous Abortion* (Springer-Verlag, London, 1992).

63. Flaxman, S. M. & Sherman, P. W. Morning sickness: A mechanism for protecting mother and embryo. *The Quarterly Review of Biology* **75**, 113-148 (2000).

64. Thompson, D. a. W. *On Growth and Form* (Cambridge University Press, Cambridge, 1943).

65. Jennings, B. J., Ozanne, S. E., Dorling, M. W. & Hales, C. N. Early growth determines longevity in male rats and may be related to telomere shortening in the kidney. *Febs Letters* **448**, 4-8 (1999).

66. Forsen, T., Eriksson, J. G., Tuomilehto, J., Osmond, C. & Barker, D. J. P. Growth in utero and during childhood among women who develop coronary heart disease: Longitudinal study. *British Medical Journal* **319**, 1403-1407 (1999).

67. Eriksson, J. G. et al. Catch-up growth in childhood and death from coronary heart disease: Longitudinal study. *British Medical Journal* **318**, 427-431 (1999).

68. Burnet, F. M. A modification of Jerne's theory of antibody production using the concept of clonal selection. *Austral. J. Sci.* **20**, 67 (1957).

69. Buchkovich, K. J. & Greider, C. W. Telomerase regulation during entry into the cell cycle in normal human T cells. *Molecular Biology of the Cell* **7**, 1443-1454 (1996).

70. Weng, N. P., Granger, L. & Hodes, R. J. Telomere lengthening and telomerase activation during human B cell differentiation. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 10827-10832 (1997).

71. Horwich, A. in *Testicular cancer: Investigation and management* (ed. Horwich, A.) 1-17 (Chapman & Hall Medical, London, 1996).

72. Greider, C. W. Telomeres do d-loop-t-loop. *Cell* **97**, 419-422 (1999).

73. Lanza, R. P. et al. Extension of cell life-span and telomere length in animals cloned from senescent cells. *Science* **288**, 665-668 (2000).

74. Khalifa, M. M. Hutchinson-Gilford Progeria Syndrome Report of a Libyan Family and Evidence of Autosomal Recessive Inheritance. *Clinical Genetics* **35**, 125-132 (1989).

75. Maciel, A. T. Evidence For Autosomal Recessive Inheritance of Progeria Hutchinson Gilford. *American Journal of Medical Genetics* **31**, 483-488 (1988).

76. Martin, G. M. in *Genetic effects on aging* (eds. Bergsma, D. & Harrison, D. E.) 5-39 (Alen R. Liss, New York, 1978).

77. Greider, C. W. Telomerase activity, cell proliferation, and cancer. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 90-92 (1998).

78. Cairns, J. Mutation selection and the natural history of cancer. *Nature* **255**, 197-200 (1975).

79. Shay, J. W. & Wright, W. E. in *Molecular Biology of Aging. Alfred Benzon symposium 44* (eds. Bohr, V. A., Clark, B. F. C. & Stevnsner, T.) 148-158 (Munksgaard, Copenhagen, 1999).

80. Shay, J. W. Toward identifying a cellular determinant of telomerase repression. *Journal of the National Cancer Institute* **91**, 4-6 (1999).

81. Wang, J., Hannon, G. J. & Beach, D. H. Risky immortalization with telomerase. *Nature* **405**, 754-755 (2000).

82. Holt, S. E. & Shay, J. W. Role of telomerase in cellular proliferation and cancer. *Journal of Cellular Physiology* **180**, 10-18 (1999).

# Chapter 3

## PLATEAUS, RIDGES AND DUNES: DIVERSITY GRADIENTS AND ADAPTIVE TOPOGRAPHY

The latitudinal gradient of species diversity has been recognized since 1799 (Humboldt, 1850) and is a strong contender for the most general ecological pattern on Earth, yet there remains no consensus on its root cause (Pianka, 1966; Willig *et al*, 2003; Hillebrand, 2004). The pattern exists in the great majority of extant taxa, on all the world's continents and, in every ocean (Hillerbrand, 2004), both at the surface and on the sea floor (Rex *et al*, 1993, 2000; but see Lambshead *et al*, 2002 for caveat regarding marine benthos).

The generality of the pattern suggests the action of a strong ecological force, acting pervasively over evolutionary time, a force that presumably continues into the present. For there to be no consensus on what strong, continuing ecological force is responsible indicates that something about our approach and/or our fundamental assumptions is incorrect and is blinding us to a process so general that it should, by all rights, be apparent.

Ecology, in the narrow sense, makes the implicit assumption that the causes of species distributions and the long-term interaction dynamics that explain them can be inferred from observations made on short 'ecological' timescales. Clearly some patterns, like forest succession (Connell and Slayter, 1977; Connell *et al*, 1987) can be. But the presumption that one can extrapolate from observations on short timescales is suspect for many other patterns of ecological interest, including the latitudinal diversity gradient. The latitudinal diversity gradient must, by its very nature, be explained in terms of species creation (speciation), species loss (extinction) and changes in species range (migration), all three being rare events that tend to elude observation. Searching for the causes of the

diversity gradient using the short time-scale tools of classical ecology may simply place the answer outside the scope of inquiry.

The rare event problem is exacerbated in ecology whenever the rare event in question is an inversion--a force that typically pulls in one direction, but occasionally pulls strongly in the other. In other words, we have reason to suspect that whenever selection favors distinct, antagonistic phenotypes on different timescales, that ecological investigations are likely to over-rate the significance of the common selective regime, and undervalue or miss the significance of the rare one.

That is particularly hazardous to our analytical framework given the power of rare, harsh conditions to bottleneck populations down to their most robust members. To the extent that something like drought tolerance is hereditary within a species, a thousand year drought will tend to produce a population in its wake that is highly resistant to dry conditions. When we study that population ecologically, we are therefore likely to observe a mismatch between the tolerances of the individuals in that population, and the range of conditions faced in an average lifespan, under average conditions. Further, when those tolerances are 'purchased' by selection at a cost relative to other desirable traits (e.g. competitive efficiency or dispersal ability) we may misunderstand the creature as simply inferior unless we see it in the rare circumstances that account for the phenotype (for a cautionary tale of this sort see Kenyon *et al*, 1993; then see Walker *et al*, 2000).

This might seem like a minor consideration were it not for the commonality with which we observe phenotypic trade-offs both within and between species (e.g. Walters and Reich, 1996; Weinstein and Ciszek, 2002). As important as trade-offs have become to modern ecological thinking, they are still vastly underrated with respect to their significance. We need a general theory of trade-offs with explanatory power. I will present a framework for incorporating trade-offs into our ecological and evolutionary theory in a manner that unlocks otherwise intractable puzzles like the latitudinal diversity gradient.

The trade-off framework presented here suggests a number of conceptual alterations to the way we address issues within ecology. First, it suggests several features in adaptive landscape space (the plateaus, ridges and dunes of this chapter's title) that are useful, logically necessary, but absent from the standard toolkit. Second, it suggests that

we should be very cautious about assuming the generality of ecological phenomena (e.g. the competitive exclusion principle; Hardin, 1960), that may correctly characterize dynamics in one biome or zone, and be absent from another. And third, for reasons already discussed, we should be extremely cautious about imagining that ecological timescales are sufficiently representative of long-term dynamics to allow ecological questions to be answered directly through observation and measurement under average conditions.

In the past when I have presented papers outlining the argument described in this chapter (first on Barro Colorado Island in 1996, at the University of Michigan in 2003, and again at the Perimeter Institute in 2009), I have argued that the discipline of ecology generally, and the question of latitudinal diversity gradients in particular, have suffered from the historical accident of ecology having been born in the temperate zone, leaving ecologists the impossible task of explaining the immense diversity of low latitudes according to the ecological rules that seemingly characterize the much lower diversity at high latitudes. I argued that Gause (1934) would not have thought to propose the competitive exclusion principle, and Hardin (1960) would not have thought to embrace it had they been working in highly diverse tropical forests or coral reefs.

The origins of ecological thinking are, it turns out, a bit more complex than I realized. A compelling argument can be made that ecology was in fact founded by a young Dane named Warming in the Brazilian tropics beginning in 1863 (Goodland, 1975). But the remains true that when the modern discipline of ecology emerged in the latter half of the 20$^{th}$ century, it was for some time almost entirely temperate. The result was that when interest in tropical ecology received serious ecological consideration beginning in the late 60s, the race was on to understand tropical diversity in the familiar terms of character displacement (Brown and Wilson, 1956), niche partitioning (Hutchinson, 1959) and competitive exclusion (Hardin, 1960) that had, to that point in the history of ecology, been so useful.

The result of attempting to shoehorn tropical systems into models built from temperate patterns has been less than impressive. In fact, the most accurate model we currently have to explain relative species abundances in tropical forests and coral reefs requires that we treat the species in question as if they are not interacting at all (see

Volkov *et al*, 2007 and the references therein) despite the fact that intense, ongoing competition for limiting resources is readily apparent in these systems (e.g. the trees are tall in lowland tropical forests).

The largest obstacle to credibly explaining the latitudinal diversity gradient in rigorous, predictive terms is knowing what features of the pattern are really in need of an explanation. Had the modern study of ecology begun where the majority of extant species are found, we would likely intuit that limiting similarity (MacArthur and Levins, 1967) can be zero (May and MacArthur, 1972; May, 1973), and that competitive exclusion is an important process, but not one we have an analytical basis to automatically expect for any given pair of ecologically similar species in sympatry. Starting from a tropical vantage point we would see that the central question demanding a mechanistic answer is not "how do so many similar species coexist in tropical habitats?" but rather "why are the temperate zones species-poor and dominated by generalists that differ widely with respect to competitive ability?"

**A framework for understanding trade-offs**

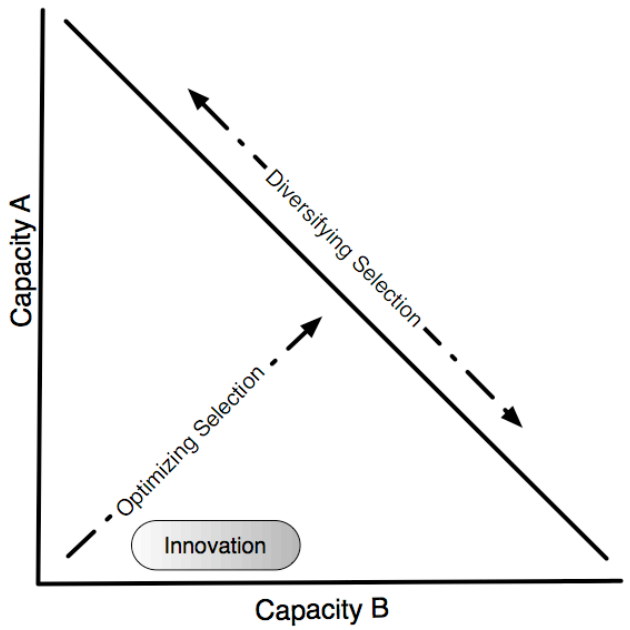
Trade-offs are inherently negative relationships between desirable characteristics. They are emergent phenomena in the sense that innovations (inherently crude structures) may be simultaneously refined in several respects at once, giving the impression that the characteristics in question are independent of each other. But optimization ultimately reaches a point—what economists call an 'efficient frontier'—beyond which simultaneous improvement is no longer possible. Along that trade-off function, further improvements to one trait negatively impact the other, and visa versa.

This emergent nature becomes apparent through a technological example. The Wright Flyer, the first airplane, was crude by any standard except comparison with its contemporaries. It was neither agile, nor efficient. But it ushered in an era of rapid refinement in which planes, including the Wright Flyer, were quickly improved in every regard simultaneously. During that era of refinement, the trade-offs that would eventually emerge would not have been apparent. Speed, safety, cargo and passenger carrying capacity, agility, efficiency all went up together. That pattern quickly reached an end,

though, and we now live in an era in which some planes are fantastically agile while other planes have ability to carry huge loads. But they are necessarily different planes, and there has been a powerful trend toward diversification. Many pair-wise trade-offs are simultaneously apparent, and functional diversity is found along many, but not all of these trade-offs functions.

This time series suggests three evolutionary phases, equally relevant to all realms with emergent design trade-offs, including biology, technology and economics. The three phases are innovation, optimizing selection (i.e. refinement) and diversifying selection (i.e. adaptive radiation). (See Figure 1.)

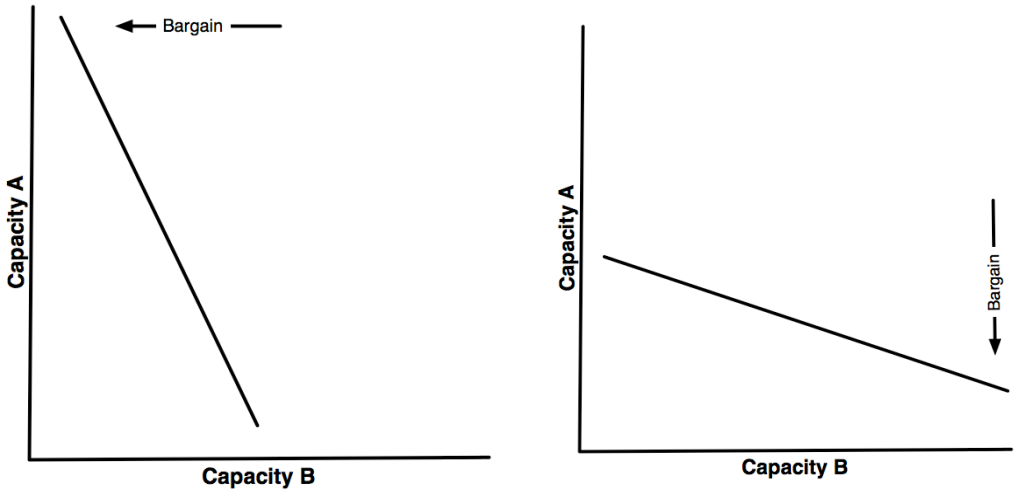
Figure 1:
**Three evolutionary phases: Innovation, optimizing selection, and diversifying selection**




This trade-off is sometimes referred to as a 'perfect' trade-off, meaning that along the frontier, the fitness benefit arising from increases in capacity A come at a cost in B that exactly balances.

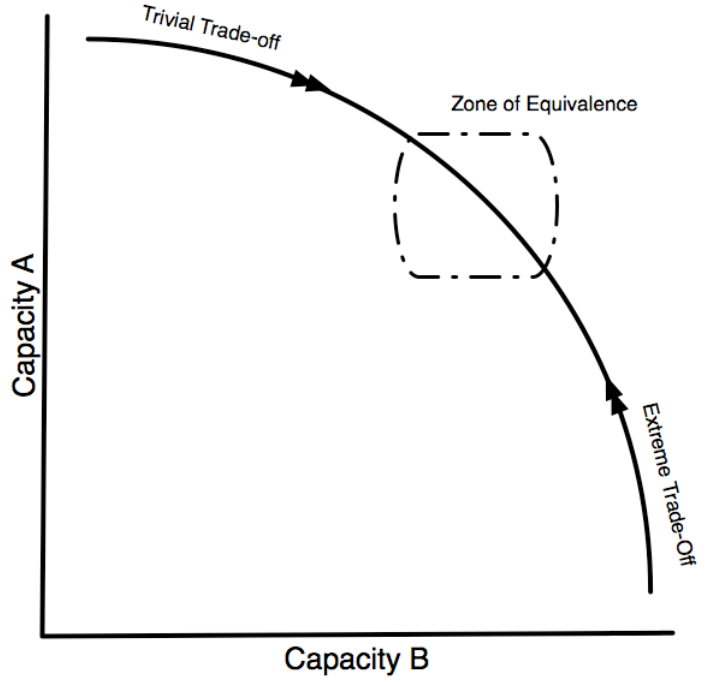
If you juggle real world trade-offs in your mind you will discover that there exist many **imperfect** trade-offs in which costs and benefits clearly do not balance (the slope is not 45 degrees) . No diversity is expected along such trade-offs because competitors that move in the direction that provides net benefits will out compete competitors that do not move. All competitors should therefore pile up at the bargain end of such a trade-off, and we observers may therefore not realize that a trade-off exists at all. Take for example the question of car safety vs. fuel efficiency. Eliminating seatbelts will improve the fuel efficiency of any car. But the fuel savings would be tiny, and the reduction in safety would be spectacular. So we don't expect to see diversity in strategies in spite of the fact that a genuine trade-off lurks there. The trade-off is too extreme (or trivial, depending on which characteristic you focus on) for rational competitors to disagree over it (Fig 2).

Figure 2:
**Imperfect trade-offs: Extreme and trivial**



But if we follow the example of safety and fuel efficiency through to its logical conclusion we find something important, that has a deep relevance to biological trade-offs. There exist foolish strategies and bargains on *both* sides of the safety vs. fuel efficiency trade-off. A fool could remove the seatbelts from his car to improve mileage, but his equally foolish neighbor could drive a military tank, thereby attaining a great deal of safety at an unacceptable cost in terms of fuel. We don't expect to see either strategy commonly employed. Instead we expect a range of vehicles with an inverse correlation between safety and efficiency (a pattern made noisy by the fact that there are many other trade-offs playing out simultaneously). The range within which we see actual, realized diversity should be bounded by the limits of where reasonable competitors could disagree labeled 'Zone of Equivalence' in Figure 3.

Figure 3:
**Zone of equivalence**




In this, as any other, diminishing returns curve, the greater the capacity in A, the larger the costs in B per unit of improvement in A. If we imagine selection acting on species with respect to characteristics that are functionally closely tied, we can begin to see why 'perfect' physiological, morphological, behavioral and ecological trade-offs are so commonly observed, and why they support the evolution and maintenance of diversity.

Species are pushed towards the diminishing returns frontier (the line that limits further improvement without compromise) by optimizing selection (Fig. 1). Species are pushed away from the ends of that curve by the relative bargains in the middle. Within the zone of equivalence, they are favored by resource competition to spread out. The zone of equivalence, if mapped in adaptive landscape space would look like a ridge, a set of contiguous peaks of equal altitude owing to their similarity with respect to fitness. Multiple trade-offs of this type interacting in adaptive landscape space (e.g. growth rate vs. resource defense vs. dispersal capacity) describe a plateau, a set of contiguous points over which strategic variation occurs without providing absolute competitive advantages.

If a species reaches a given frontier in design-space first, with no interspecies competition, it is likely to spread out (generalize) over the zone of equivalence. If there are other species in the zone of equivalence already, all species present will be competitively pushed toward any open spots, where competition is low. Once the line is saturated, there is no reason to expect limiting similarity, because within that zone, costs balance benefits as species move along the frontier. On the competitively saturated line, character displacement is no longer favored, and the number of species in a given niche can rise with time.

This model matches the initially bizarre observations that led Hubbell (2001) to propose the unified neutral theory of biodiversity, but also has the power to account for the relative species poverty of the temperate zones, as well as the tendency for competitors in the temperate zones to exclude each other. It also matches an important theorem in economics which states that 'efficient markets' fluctuate randomly (Samuelson,1965).

Tropical forests and coral reefs are fantastically diverse and fiercely competitive, yet the effects of competition can, to first approximation, be ignored in predicting relative species abundances and changes in population size (Volkov *et al*, 2007; Hubbell, 2001). The reason for that seeming paradox now comes into focus. Design trade-offs are not particular to a species. They are extrinsic constraints, arising directly, in a fundamental way, out of physics and chemistry, limiting what is possible within a given segment of design space. Hubbell (2001) has separately reached similar conclusions. Species that reach these limits will naturally tend to evolve toward the zone of equivalence, and once

there, they lack the ability to competitively exclude other species—They can't exceed the frontier without a new innovation. Falling below the line would render them competitively inferior, so selection holds them to the line. Leaving the zone of equivalence in either direction produces net losses. And within the zone, competitors are equal.

***What does this model predict about tropical ecosystems?***

1. Fierce, active competition with no net trend in population sizes of the species present—population sizes should drift as individuals fight to the death with a sea of equal rivals.
2. Every limiting resource falls to exceedingly low levels.
3. Recognizable niches, with no limit on species number within a given niche.
4. Specialization rates should be high.
5. Species ranges should be narrow.
6. Ecological tolerances should also be narrow.
7. It should be almost impossible to invade a mature mainland tropical habitat because invaders can, at best, come in as equal competitors (since they are bound by the same extrinsic design limits) and they inevitably come in at very low population numbers—an equal competitor, starting with a population size near zero, fluctuating randomly, is almost certain to go extinct before it becomes common enough to call it a successful invader. And if the competitor comes from slightly different habitat, it will tend to be an inferior competitor in the novel environment.

And what does this framework predict about the temperate zones? Why should they look any different?

The temperate zones fluctuate over a much larger range of climatic conditions on every time scale of biological relevance. And on timescales greater than an individual lifespan, this irregular fluctuation creates an inherent evolutionary hazard. The trade-off

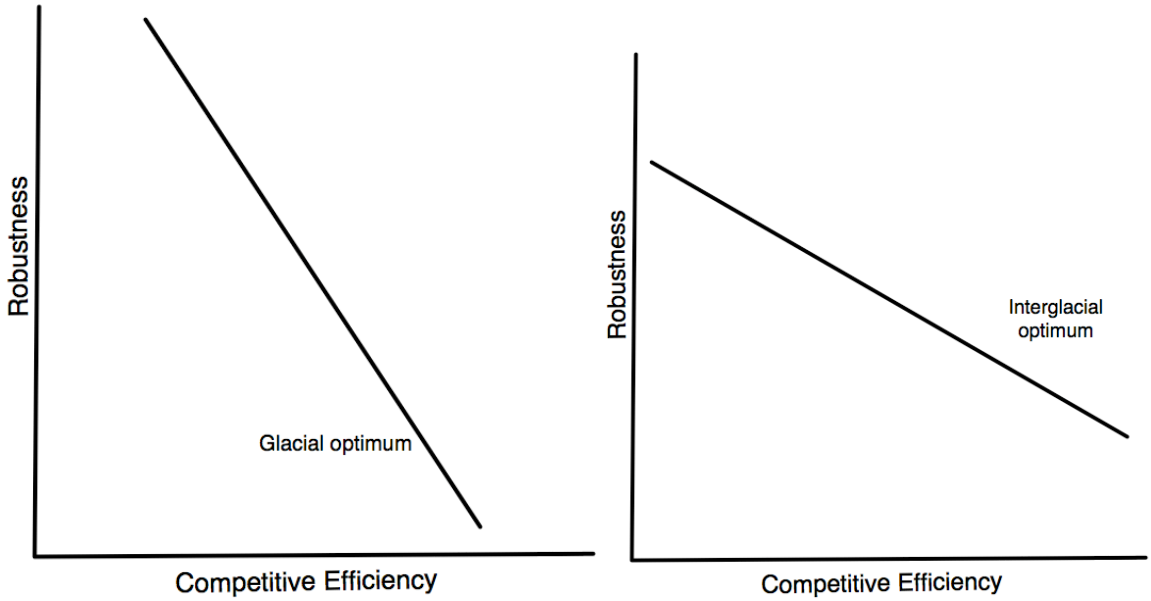
lines that describe the design limits will oscillate, preventing species from converging on a strategy that would render them equal competitors.

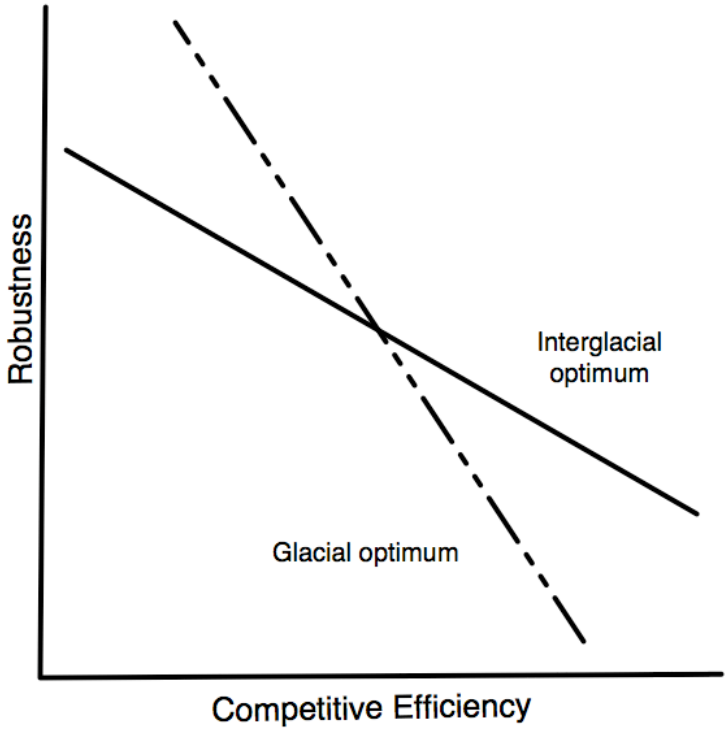
Consider, for example, glaciations, driven by long-periodicity Milankovitch cycles. Following a glaciation, all individuals persisting in temperate locations that exhibit severe climate alteration will tend to be robust, having descended from populations of individuals that weathered the severe event in one way or another. They will repopulate the landscape as the glaciers recede, creating broadly distributed populations of climatologically robust individuals.

But the tolerances that offer protection from harsh climate come at some cost, as argued above. Those costs would tend to be a bargain while the temperate zones are glaciated, but those costs would be a competitive disadvantage during the mild conditions that characterize the interglacial (Fig 4) . Competition in mild conditions will favor the elimination of the very robustness factors that allowed persistence through glaciation, as competitive efficiency becomes a bargain relative to climatological tolerance. And the populations that shed those robustness factors fastest will enjoy a comparative advantage over populations that retain them longer.

That flip-flop in dominance between biotic and abiotic hostile forces creates no safe refuge. Species that retain robustness and are slow to become competitively efficient will be vulnerable to competitive exclusion by species that are quick to adapt to the mild conditions. When the glaciers descend again, the dominant competitors in the mild period will be fragile with respect to the harsh climate. Each oscillation will tend to wipe species out on both ends: glaciation will drive many good competitors to extinction, and competition will eliminate many robust creatures. It is an adaptive dune, where peaks and valleys undulate, and species can’t help but bounce around in response.

Figure 4: **Robustness vs. efficiency in glacial and inter-glacial periods**



***What does this model predict about the temperate zones?***

The first thing to realize is that, unlike the tropics, it matters when, in a paleontological sense, you ask the question.

1. In all periods the temperate zones should be species-poor, climbing slowly due to migration and speciation after glaciers recede.
2. During an interglacial, species should exist over broad geographic ranges,
3. During an interglacial, species should exist with high levels of intraspecific variation.
4. Species in similar niches (but in different locations) should vary widely in competitive ability, being constantly in motion from a past optimum to a new one.
5. When similar species encounter each other, one should always be competitively superior and should exclude the other.
6. When humans transport creatures between temperate locations with similar average (interglacial) conditions, those organisms should tend to invade if their native habitat has lower climatic variance than their new location.
7. As a consequence, invasions *from* temperate habitat A *to* temperate habitat B should imply immunity *of* habitat A to invasion by creatures *from* B (rather than reciprocal invisibility).
8. Resources over which temperate species compete should be captured with much lower efficiency than in similar tropical habitats, leaving more in excess.
9. The level of within habitat excess resource should drop slightly as the interglacial period progresses and selection favors efficient competitors.

**How species come to divide up space, and resist the hazards in time**

The model above raises two interesting questions, the answers to which may resolve another longstanding puzzle: (1) In habitats where competition remains the dominant selective force (because fluctuations are narrow and/or highly predictable), how

do generalists diversify into specialists without physical barriers to interbreeding? And (2), is there any conceivable adaptive response that might diminish the tendency for competitors in widely fluctuating environments to drive each other to surrender their heritable robustness factors during mild times?

It is perhaps best to approach both questions from the other side by confronting a seemingly disjunct mystery. Costly sexual displays are traditionally explained in two ways. First, "sexy son models" (Weatherhead and Robertson, 1979) suggest that arbitrary sensory preferences (typically in females) favor corresponding elaborations in males. Females with such preferences therefore tend toward the production of sons that are attractive to other females with such preferences, leading to a positive feedback for the elaborations and the preferences. These models have a hollow ring, though, because the large cost typical of most displays ought to put the brakes on the positive feedback mechanism: the more elaborate the displays and the stronger the preferences, the greater the comparative ecological advantage drab males and display-indifferent females should enjoy. That same criticism can be leveled at the runaway portion of Fisherian models (Fisher, 1930).

The alternative to sexy-son and runaway models has traditionally been good-genes models in which male display functions as an honest advertisement of male genetic quality (for review see Zahavi and Zahavi, 1997). These models avoid the pitfall mentioned above, but they raise another question. If females chose genetically superior males by using displays as a proxy, inferior genes ought to be strongly disfavored. The result over time should be a drop in the value of the displays and the preferences for them as heritability is reduced by the fixation of good genes. When 'bad' genes become rare, indifferent females and drab males ought to have an ecological advantage due to their lesser investment in vigilance and display. That should produce an oscillating temporal pattern, where display and choosiness are amplified when 'bad' genes are common, and fall as 'bad' genes become rare, reduction in female vigilance allowing new 'bad' genes produced by mutation to creep up in frequency. But no such fluctuating pattern is known.

The Zahavi handicap model (Zahavi and Zahavi, 1997) provides a potential solution in that costliness itself is an integrative measure of male well-being. But it

requires the ecological benefits to a female's daughters to more than outweigh the display costs to her sons. And, although there may be situations in which that would be the case, it is unlikely to be stable because the costs to displaying sons remain high even when 'bad' genes are rare and the benefits to daughters are consequently small.

My contention is that sexual selection is misunderstood for two reasons. First, it is actually two distinct phenomena that have been studied as if they were one; Second, many sexual adaptations are responses to selection on lineages, rather than on individuals, which most models assume.

The first type of sexual selection is likely to be characteristic of habitats that are temporally stable over long periods. In such habitats, special variation in ecological conditions cants the generalism/specialization trade-off heavily in the direction of specialization. That is because in a widely distributed population, individuals that have broad capacities pay a cost for them that is likely not to be recovered in the narrower range of conditions under which individuals compete. Conspecifics that function more efficiently, under a narrower range of conditions, will tend to become competitively dominant under those conditions at the cost of being inferior under other conditions.

Therefore, as Darwin (see Stauffer, 1975), Fisher (1930), Endler (1977) and others have recognized, there should be a tension on broadly distributed species to fragment parapatricaly (for review see Coyne and Orr, 2004). The traditional problem for models of parapatric speciation is the need for sufficient 'reinforcement' to produce multiple species absent a physical barrier to interbreeding. Sexual selection has been proposed as a mechanism that can produce that effect. What has been less well understood is that local adaptation in stable environments provides a definition of 'good' genes and 'bad' genes that does not predict fluctuation. If two adjacent locations differ in their parameters such that competitively dominant individuals in one location are inferior at the other, then each environment is likely to be a perpetual source of 'bad' genes for the other, favoring female vigilance and therefore male display (see Hereford, 2009 for review of evidence for local adaptation and fitness trade-offs).

It is also true that a local adaptation model of sexual selection explains the otherwise enigmatic richness of sexual signals. Take for example the lekking behavior of the Satin Bowerbird, *Ptilonorhynchus violaceus* (Patricelli *et al*, 2003). Males build a

structure of purely esthetic value, adorn it with rare items that must be collected and defended against rivals, and they dance intricately around the structure. Now imagine a robust, healthy male adapted to ecological conditions in some other part of the species range—conditions dissimilar to the ones the female's offspring will face. That male might fly in, evict a resident male and attempt to mate. Were the female to simply assess his current condition, she might think him a good choice and in so doing select for her offspring genes that are well suited to some other environment (genes that are locally bad). On the other hand, if she assesses his ability to dance around the bower gracefully, she may well recognize him as an interloper, as it takes practice to dance smoothly over an idiosyncratic surface. If he dances well, he has been there long enough to become proficient dancing around the particular structure.

Now imagine the male flies in from afar, evicts a resident male from an attractive bower and practices the dance to perfection, all while defending and procuring rare (blue) objects and feeding himself. In such a case, the female would not be making a mistake mating with him because, wherever his genes may be from, they are well enough suited to **her** environment that he is able to forage, defend and practice the dance—all under local conditions—for a long enough period that he can dance with grace. If a male is healthy, well fed, has a nice bower and dances well around it, he necessarily well adapted to the local environment, whether he built the structure and collected the rare items or not. Similar arguments can be made for many sex displays and preferences in many comparatively stable environments, the key being that they provide an indicator of success coupled with an indicator that the wellbeing is being maintained locally.

But what about displays in fluctuating environments where the landscape favors broadly distributed generalists and local adaptation is not prevalent? Here a distinct possibility emerges. As discussed above, periodic harsh conditions will tend to bottleneck populations down to the most robust subset of individuals. As mild conditions return, intraspecific competition will tend to erode the heritable robustness factors that are required to get through harsh times, as instantaneous natural selection (i.e. selection to conditions in the immediate circumstance) favors competitive efficiency over robustness.

Lineages that respond to that instantaneous selection become vulnerable to fluctuation in the long term. But females are in a position to exert a countervailing force

that is capable of preventing their own lineages from going down that dangerous path. By favoring males that show an ability to acquire resources above and beyond ecological requirements, males are effectively advertising the presence of genes that are likely to be robust in harsh times.

It may at first seem like elaborate male displays would have the opposite effect, burdening creatures that invest in them such that they would be more vulnerable to harsh conditions. And that supposed tendency of costly displays to reduce the mean fitness of the population has been proposed and widely discussed elsewhere (Haldane, 1932; Lande, 1980; Kondrashov and Yampolsky 1996). But that discussion misses an important characteristic of many, if not all, such displays: costs can be reduced in response to environmental changes.

Peacocks (*Pavo cristatus)*, for example, re-grow their tails every year, and the size varies with the animal's current condition (Loyau, 2005). In other words, instead of making large ratty looking structures, males respond to harsh conditions by producing smaller tails. Presumably, therefore, during an extremely harsh event, males that would in favorable times have had the most excess resource, and consequently the most elaborate displays, would have much smaller displays. Males who would have had smaller displays under mild conditions would presumably be absent under harsh conditions, because their reserves would be inadequate for survival. Females would still favor the males with the relatively biggest displays, small as those displays might become.

According to this model, female choice in fluctuating environments functions as an evolutionary tensioner, counteracting the tendency of mild conditions to favor competitively efficient—and therefore climatically vulnerable—lineages. In order for this model to function, displays must be expensive, must track current condition of the creature in question, and be recoverable when conditions are not favorable. Displays being expensive and tracking current condition has been well documented (for Eastern Bluebirds [*Sialia sialis*] see Siefferman, 1998; for Brown Headed Cowbirds [*Molothrus ater*] see McGraw, 2002; for Blue Grosbeaks [*Passerina caerulea*] see Keyser and Hill, 2000; for Blue-Black Grassquit [*Volatinia jacarina*] see Doucet, 2002). And the recoverability of expense is evident in any system where calls or dances can be reduced in duration or intensity, or structures can be reduced under harsh conditions. It is also the

case that male-male competition (e.g. rutting behavior in bovids) can serve the same function relative to climatic fluctuation—robust males effectively demonstrating excess capacity that could be redirected in climatically unfavorable circumstances. The migratory behavior of anadromous salmonids may even be a manifestation of tensioning selection exerted by females, as it is clearly a demonstration of recoverable excess capacity on the part of males, demanded by females.

This model is closely related to the handicap model, but it has two important advantages. First, it does not require the benefits to daughters to exceed the cost to sons, because the cost to sons is reduced or recovered, potentially in its entirety, when natural selection is strongest in harsh times. It therefore predicts that sexual selection in such systems enhances lineage fitness rather than detracting from it. Second, it resolves the question of why bad genes persist in the face of strong female choice against them—the 'bad' genes are genes well suited to mild conditions and ill suited to recurrent harsh conditions.

## Literature Cited


Brown, W.L. and Wilson, E.O. (1956). Character displacement. *Systematic Zoology, 5(2)*, 49-64.

Connell, J.H., Noble, I.R., Slatyer, R.O. (1987). On the mechanisms producing successional change. *Oikos, 50(1)*, 136-7.

Connell, J.H. and Slatyer, R.O. (1977). Mechanisms of succession in natural communities and their role in community stability and organization. *American Naturalist, 111(982)*, 1119-44.

Coyne, J.A. and Orr, H.A. (2004). *Speciation*. Sunderland, Massachusetts: Sinauer Associates.

Doucet, S.M. (2002). Structural plumage coloration, male body size, and condition in the Blue-Black Grassquit. *Condor, 104(1)*, 30-8.

Endler, J.A. (1977). *Geographic variation, speciation, and clines*. Princeton: Princeton University Press.

Fisher, R.A. (1930). *The genetical theory of natural selection*. London: Oxford (original publisher: The Clarendon Press).

Gause, G.F. (1934). *The struggle for existence*. Baltimore: Williams and Wilkins

Goodland, R.J. (1975). The tropical origin of ecology: Eugen Warming's jubilee. *Oikos, 26(2)*, 240-5.

Haldane, J.B.S. (1932). *The causes of evolution*. London: Harper.

Hardin, G. (1960). Competitive exclusion principle. *Science, 131(3409)*, 1292-7.

Hereford, J. (2009). A quantitative survey of local adaptation and fitness trade-offs. *American Naturalist, 173(5)*, 579-88.

Hillebrand, H. (2004). On the generality of the latitudinal diversity gradient. *American Naturalist, 163(2)*, 192-211.

Hubbell, S.P. (2001). *The unified neutral theory of biodiversity and biogeography*. Princeton: Princeton University Press.

Humboldt, A.V. (1850). *Views of nature: or, Contemplations on the sublime phenomena of creation, with scientific illustrations*. London: Henry George Bohn.

Hutchinson, G.E. (1959). Homage to Santa-Rosalia or Why are there so many kinds of animals. *American Naturalist, 93(870)*, 145-59.

Keyser, A.J. and Hill, G.E. (2000). Structurally based plumage coloration is an honest signal of quality in male blue grosbeaks. *Behavioral Ecology, 11(2)*, 202-9.

Kenyon, C., Chang, J., Gensch, E., Rudner, A., Tabtiang, R. (1993). A C-Elegans Mutant That Lives Twice as Long as Wild-Type. *Nature, 366(6454)*, 461-4.

Kondrashov, A.S. and Yampolsky, L.Y. (1996). Evolution of amphimixis and recombination under fluctuating selection in one and many traits. *Genetical Research, 68(2)*, 165-73.

Lambshead, P.J.D., Brown, C.J., Ferrero, T.J., Mitchell, N.J., Smith, C.R., *et al*. (2002). Latitudinal diversity patterns of deep-sea marine nematodes and organic fluxes: a test from

the central equatorial Pacific. *Marine Ecology-Progress Series, 236(*129-35.

Lande, R. (1980). Sexual dimorphism, sexual selection, and adaptation in polygenic characters. *Evolution, 34(*2*)*, 292-305.

Loyau, A., Saint Jalme, M., Cagniant, C., Sorci, G. (2005). Multiple sexual advertisements honestly reflect health status in peacocks (*Pavo cristatus*). *Behavioral Ecology and Sociobiology, 58(*6*)*, 552-7.

Macarthur, R. and Levins, R. (1967). The limiting similarity, convergence, and divergence of coexisting species. *American Naturalist, 101(*921*)*, 377-86.

May, R.M. (1973). *Stability and complexity in model ecosystems*. Princeton, NJ: Princeton University Press.

May, R.M., and Macarthur, R. H. (1972). Niche overlap as a function of environmental variability. *Proceedings of the National Academy of Sciences of the United States of America, 69(*5*)*, 1109-13.

McGraw, K.J., Mackillop, E.A., Dale, J., Hauber, M.E. (2002). Different colors reveal different information: how nutritional stress affects the expression of melanin- and structurally based ornamental plumage. *Journal of Experimental Biology, 205(*23*)*, 3747-55.

Patricelli, G.L., Uy, J.A.C., Borgia, G. (2003). Multiple male traits interact: attractive bower decorations facilitate attractive behavioural displays in satin bowerbirds. *Proceedings of the Royal Society of London Series B-Biological Sciences, 270(*1531*)*, 2389-95.

Pianka, E.R. (1966). Latitudinal Gradients in Species Diversity - a Review of Concepts. *American Naturalist, 100(*910*)*, 33-46.

Rex, M.A., Stuart, C.T., Coyne, G. (2000). Latitudinal gradients of species richness in the deep-sea benthos of the North Atlantic. *Proceedings of the National Academy of Sciences of the United States of America, 97(*8*)*, 4082-5.

Rex, M.A., Stuart, C.T., Hessler, R.R., Allen, J.A., Sanders, H.L., Wilson, G.D.F. (1993). Global-Scale Latitudinal Patterns of Species-Diversity in the Deep-Sea Benthos. *Nature, 365(*6447*)*, 636-9.

Samuelson, P.A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review, 6(*2*)*, 41-9.

Siefferman, L., Hill, G.E. (2003). Structural and melanin coloration indicate parental effort and reproductive success in male eastern bluebirds. *Behavioral Ecology, 14(*6*)*, 855-61.

Stauffer, R.C., ed. (1975). *Charles Darwin's Natural Selection: Being The Second Part Of His Big Species Book Written From 1856 To 1858*. Cambridge: Cambridge University Press

Volkov, I., Banavar, J.R., Hubbell, S.P., Maritan, A. (2007). Patterns of relative species abundance in rainforests and coral reefs. *Nature, 450(*7166*)*, 45-9.

Walker, D.W., McColl, G., Jenkins, N.L., Harris, J., Lithgow, G.J. (2000). Natural selection - Evolution of lifespan in C-elegans. *Nature, 405(*6784*)*, 296-7.

Walters, M.B., Reich, P.B. (1996). Are shade tolerance, survival, and growth linked? Low light and nitrogen effects on hardwood seedlings. *Ecology, 77(*3*)*, 841-53.

Weatherhead, P.J., Robertson, R.J. (1979). Offspring quality and the polygyny threshold: the sexy son hypothesis. *American Naturalist, 113(2)*, 201-8.

Weinstein, B.S., Ciszek, D. (2002). The reserve-capacity hypothesis: evolutionary origins and modern implications of the trade-off between tumor-suppression and tissue-repair. *Experimental Gerontology, 37(5)*, 615-27.

Willig, M.R., Kaufman, D.M., Stevens, R.D. (2003). Latitudinal gradients of biodiversity: Pattern, process, scale, and synthesis. *Annual Review of Ecology Evolution and Systematics, 34,* 273-309.

Zahavi, A., Zahavi, A. (1997). *The handicap principle: a missing piece of Darwin's puzzle*. Oxford: Oxford University Press.

# Chapter 4

# THE BETTER ANGELS OF OUR NATURE: GROUP STABILITY AND THE EVOLUTION OF MORAL TENSION


## Abstract

Moral systems require individuals to act in service to their social groups. Despite the human tendency to view moral norms as invariant and constantly deserving of adherence, we vary not only in the moral norms we espouse but also in the degree to which our behavior reflects those norms. Nevertheless, moral systems exhibit patterns and complexity that suggest the action of natural selection. We propose that much observed variation in commitment to the group can be explained by a rule of *stability-dependent cooperation*, where the adaptive level of individual commitment varies inversely with the stability of the social group. This hypothesis is rooted in the understanding that humans are caught in an evolutionary trade-off between two methods of increasing reproductive success: competing with fellow group members, and increasing the stability of the group relative to other groups. If cooperation is stability-dependent, however, human groups in times of high stability and low cooperation may be susceptible to fast-acting extrinsic threats as well as self-destructive competitive races to the bottom. In light of this, we hypothesize that the absolutism and unchangeableness commonly attributed to moral norms serves a *group stability insurance* function, and present predictions from this hypothesis.



**Key words:** morality, evolution, inter-group competition, cooperation, altruism

> ”We are not enemies, but friends. We must not be enemies. Though passion may have strained it must not break our bonds of affection. The mystic chords of memory, stretching from every battlefield and patriot grave to every living heart and hearthstone all over this broad land, will yet swell the chorus of the Union, when again touched, as surely they will be, by the better angels of our nature.”
>
> -Abraham Lincoln, Second Inaugural Address (1861).

## 1. Introduction

The history of evolutionary approaches to morality has been characterized by debate between those who claim that “evolution has, as a matter of fact, constructed human beings to act for the community good” (Richards, 1987: 623-624; see also Ruse, 1986; Sober & Wilson, 1998), and a diverse opposition maintaining either that morality is selectively neutral, or that it requires us to combat the evolutionary process and “rebel against the tyranny of the selfish replicators” (Dawkins, 1976; see also Huxley, 1894; Williams, 1988). Despite the productivity of this discussion (Maienschein & Ruse, 1999), this paper argues that neither general perspective is sufficiently explanatory; moral action is not always adaptive, but neither is it neutral or maladaptive. We argue instead that *the propensity for moral deliberation* is the fitness enhancing characteristic, any given moral action shifting between adaptive and maladaptive depending on context. This paper draws attention to the significant variation humans exhibit in individual commitment to moral norms, and proposes that the variation represents generally adaptive responses to dynamic social environments.

Although psychologists are intensely aware of the importance of social influences on morality, they are less aware of whether and how individual commitment to moral responsibility varies with societal variables (Hartup & van Lieshout, 1995). This gap provides an opening for hypotheses that predict empirical trends based on an evolutionary consideration of the nature of morality. This paper presents two principles that predict variability in individual commitment to moral norms as a function of one’s perceived social context.

A background assumption for this discussion is that a moral rule tends to be manifest in consciousness as absolute, in two senses. First, when people promote one alternative as morally correct, they imply that it is superior to all others in some general way. Regardless of our actions or desires, we tend to treat moral rules as deserving of

absolute adherence. Second, humans tend to consider moral rules absolute in the sense that they carry an implication of permanence across time and space. Despite moral variation within and between individuals, humans tend to operate under the assumption of an underlying truth to moral rules that does not change (Mackie, 1981). This idea that moral absolutism is widespread and general is a refutable psychological and sociological hypothesis. If morals are deemed absolute in the first sense (deserving of absolute adherence), people should consistently endorse alternatives they deem morally right, even if their behavior or desires conflict with these moral precepts; and if morals are deemed absolute in the second sense (invariable in time and space), people should tend to evaluate the attitudes and behaviors of others according to their own moral belief systems, without regard for cultural or historical differences.

**2. Commitment to moral rules varies**

The central problem to be addressed in this paper is that despite moral absolutism, people's lives exhibit variation in commitment and adherence to the moral rules they recognize. Humans deliberate, calculate, and often struggle-- sometimes adhering to the rules, sometimes not. Perhaps the central paradox of morality is the fact that behavior does not always match the moral rules espoused by the agent. Moral rules are considered absolute, but adherence is facultative.

When social and natural scientists have asked why this apparent incongruity exists, their answers often fall into two broad categories. One general solution is to view moral rules as contrary to self-interest, such that the two are continually in opposition. The other, perhaps more common, solution is to see morality as always consistent with globally calculated self-interest, and our moral struggles and deliberation as internal conflict between short-term and long-term self-interest. Both these alternatives interpret variation in commitment to moral norms as a maladaptive byproduct of a weakness or inconsistency in human psyche. Either the difficulty of self-sacrifice or the difficulty of foregoing short-term benefits for long-term ones is proposed as the psychological constraint that limits compliance with moral rules.

An evolutionary perspective raises doubts about the explanatory power of both these solutions. If moral behavior were simply a hindrance to the competitive

ascendancy of the individual, one would expect it to dwindle, and if it were simply adaptive, then absolute compliance would evolve and presumably sweep to 'fixation'. Given this situation, Peters (2003) recently drew attention to the fact that evolutionary studies of morality have still failed to produce an effective explanation of the fact that humans appear to be disposed both toward and against prosocial or group-serving behavior.

What, therefore, needs to be explained? What is the hypothesized locus of adaptation? An attempt to explain the adaptive significance of perfectly moral behavior, although a common goal, is misguided since such behavior is never observed. We contend that a successful theory must address the adaptive significance of the *facultative* adherence to moral absolutes. From this perspective, the traits that appear to be adaptations are the capacity for moral behavior and the tendency toward moral deliberation, as distinct from the execution of any particular behaviors all the time. This is similar to Nesse's (2001) suggestion that natural selection may have shaped "commitment strategies" for effective use in society. If this is correct, the key to understanding morality from an evolutionary perspective lies in discovering the extrinsic factors that govern moral deliberation and moral commitment. The merit of an evolutionary explanation for moral behavior can be judged on its ability to predict what conditions will produce compliance versus defection. Can variability in human environments explain our plasticity in following the rules?

**3. Morality and intergroup competition**

Richard Alexander has shown that two related facts are key elements in an evolutionary understanding of morality. First, humans "evolved to live in groups, within which they both cooperate and compete and outside of which they presumably failed consistently"; secondly, "some acts of costly beneficence enable the survival of the entire group, when that outcome is essential for our own survival" (Alexander, 2004; see also 1992; 1987). Social grouping evolved in humans in an unprecedented way, with low within-group relatedness (relative to eusocial animals) and multiple breeding males within groups. Alexander built on earlier writers such as Darwin (1871) and Keith (1949) in explaining the evolution of this phenomenon. At first our grouping was

probably maintained by selection for predator avoidance, and later for cooperative group-hunting (Alexander, 1989), but eventually, as humans reduced their susceptibility to "hostile forces of nature", the main threat to an individual's reproductive success became other people, and competition among human groups became the primary function of group living. Human cooperation within groups, then, probably evolved as a way to compete between groups, as individual reproductive success was served increasingly by maintaining solidarity with one's fellows (Alexander, 1990). The most important mechanisms for this cooperation appear to have been (a) extensive and differential nepotism, and, arising in this context, (b) social reciprocity of two different sorts. The direct sort is the process of "indefinitely continuing interactions between intelligent beings in which each can benefit from cooperating with the other, and… defection… will in the long run represent net losses to the defector." (Alexander, 1992; see also Trivers, 1971). The indirect sort of reciprocity arises when multiple parties interact in the same way that two do in direct reciprocity. In a species with powers of observation, memory, and communication of individual reputations within a social group, rewards for cooperating (and punishments for defecting) can be administered by "society at large, or from other than the actual recipient of beneficence" (Alexander, 1979).

If humans have tended throughout their history to fail outside social groups, and if threats from other groups have rendered the suppression of competition within a group necessary for individual reproductive success, it is not surprising that individuals should often subjugate their own interests to those of their groups. Insofar as morality requires attention to group causes such as the welfare of others, morality functions as social cement and thus tends to contribute to long-term individual interests, i.e. fitness. Moreover, as indirect reciprocity became important for insuring service to the group, individuals perceived as morally upstanding would gain additional benefits through the approbation of others. These two interactive processes are broadly similar to the two kinds of games (“public goods” and “image scoring”) that recent experimental studies have employed to illustrate the dynamics of human cooperation (e.g., Milinski *et al.*, 2002).

Indirect reciprocity may be able to account for part of the facultative or inconstant nature of human commitment to moral norms. Indirect reciprocity is a system in which

one's reputation, built on others' observations of one's behavior in the past, affects one's prospects. It is not that a particular moral or immoral action inherently enhances or degrades the actor's fitness. The net effect is dependent on (a) the direct costs and benefits of the action, (b) the likelihood of being observed or of reports being believed, (c) the reputational shift that will result, and (d) the expected return on that shift in future interactions. Sensitivity to cues of these parameters and their net effect on fitness would account for a degree of nonrandom variation in adherence to moral norms.

Such, perhaps, is the adaptive value of the refined moral systems characteristic of extant human groups, where the proceeds from indirect reciprocity arguably have grown to be more relevant to decision-making than the proceeds from continued group persistence. For an average member of a modern group, the likelihood of suffering a significant fitness cost from the damage to reputation that an immoral choice can produce, is much higher than the likelihood of suffering a significant fitness cost from the loss of group unity that might arise from that choice. This is so even if Darwin and Alexander are correct in arguing that the need for solidarity against threats from other groups is precisely what drove the evolution of human cooperation, including those actions maintained by indirect reciprocity, in the first place. Thus, the system of moral reputation may have greater and more immediate fitness effects today than do the selective pressures that favored that system's origin.

To elaborate more fully the relationship between indirect reciprocity and human social structure, one can demonstrate that indirect reciprocity today depends on a concept of group service, but the existence of group service does not depend on indirect reciprocity. Indirect reciprocity is a means that needs an end; it requires a criterion for a 'good reputation'. If indirect reciprocity functioned independently of its evolutionary origin in intergroup competition, one's reputation would merely reflect shrewdness rather than group service, but service to others is a cardinal value fostered by indirect reciprocity. Humans expend significant effort debating the validity of claims of selflessness and bestowing praise for actions deemed selfless. Shrewdness does persist, because it can sometimes be an effective way to exploit the system, but it is discouraged by moral norms and thus suppressed by indirect reciprocity. The centrality of the concept of service or selflessness in moral norms (Roes & Raymond, 2002; Ridley, 1996)

suggests that within-group cooperation in the face of intergroup competition still underlies indirect reciprocity today. Otherwise, entertaining the idea that the actions of others can have group-serving motivations would be maladaptive. Moreover, in order for the earliest form of indirect reciprocity to produce fitness benefits, a belief in group service must already have existed, implying an independent evolutionary origin. The human social situation of within-group cooperation as a form of between-group competition explains why the concept of selflessness, rather than shrewdness, is the value encouraged by indirect reciprocity.

Following this reasoning, we propose that there was a period in the evolution of morality when group service was adaptive due to rising intergroup competition, but before indirect reciprocity became dominant. We are not aware of any previous proposal of such a period. Furthermore, the dynamics that drove cooperation at this intermediate stage in the evolution of morality may still be important today. Since moral norms are still pervaded by a strong group-service element, something about human social group dynamics are probably still today providing the values for indirect reciprocity, i.e., determining what behaviors will be productive of what kinds of reputation.

## 4. Group stability

Competition between groups implies that groups, like individuals, vary in how well they are doing. In the absence of a governing structure, such that groups are autonomous or nearly so, competition between groups for limited resources will function like competition between individuals, with variation in groups' likelihood of persistence analogous to the concept of fitness that evolutionary biologists use to compare individuals. (Since "fitness" in a biological sense generally refers to contribution to succeeding generations via reproduction, and since human groups do not reproduce as a whole or have discernable generations, we avoid the term "fitness" and use the term *group stability* to reflect this relative likelihood of group persistence.)

### *4.1. Principle of stability-dependent cooperation*

If humans are facultative in adherence to moral norms (section 2), and moral norms arose and are probably still maintained in the context of intergroup competition

(section 3), then variation in individual commitment to moral norms over time and space may reflect variation in the degree to which groups require service. Thus, the first of two principles we introduce to describe the dynamical function of morality in human history is *stability-dependent cooperation*. We propose that people vary in the relative importance they place on the individual versus the group in their working value systems or decision rules, because *individual sacrifice in service to the group at a given time and place is adaptive in inverse proportion to the stability of the group relative to its competitors*.

The general idea of adaptive variation in moral psychology has some precedent. Even proponents of a relatively strict developmental structure to morality have allowed for, and found indirect evidence of, apparently adaptive differences among cultures in the way morality is used to guide individual decisions (Nisan & Kohlberg, 1982; Edwards, 1975). Also, evolutionary psychologists and anthropologists have shown that rules of social exchange can vary in ways that are predictable from environmental conditions (review in Cohen & Vandello, 2001; Cosmides & Tooby, 1992). The principle of stability-dependent cooperation offered here predicts adaptive variation specifically in *commitment* to moral rules, as distinct from in the rules themselves or the development of their recognition. The foundation for this principle is the dynamic of natural selection in situations where individuals with divergent interests exist in collectives on which their persistence depends. A parallel dynamic best explains the overarching cooperation of genetic elements temporarily united in a genome. The genome works together, and subsets only rarely seek their own interests at the expense of other elements (Buss, 1987), because the persistence of a gene or chromosome depends on the survival and reproduction of the individual housing it. Cooperation to increase individual fitness is therefore usually the best strategy for a genomic element. When genetic elements behave competitively within a genome, as in T-haplotype mice (Lyon, 2003) this tends to produce negative fitness consequences for the individual, and thus for all other elements within it.

Likewise, individual humans depend on their social groups. Service to group causes fosters unity, and can decrease the effects of resource limitation (e.g., restrict hoarding and squandering, mitigate distribution of wealth effects, and otherwise increase

efficiency of resource use), thereby decreasing within-group competition and increasing the group's stability and competitive prowess (Frank, 2003; Alexander, 1979). Conversely, within-group competition arising from individual self-interest can be self-defeating: resource utilization becomes less efficient, and group unity erodes, increasing susceptibility to intergroup competition.

Some writers, often extrapolating from economic models, have hypothesized a general tendency of cooperation either to decay, or to fluctuate in regular boom and bust cycles (Nowak & Sigmund, 1998). Some have gone on to suggest that, given these proposed tendencies, in order for cooperation to be maintained and group persistence to be assured over time, a certain specialized trait must have evolved and persist at some threshold level in the population, such as "strong reciprocity" involving costly punishment of the selfish (Gintis, 2000) or "phenotypic defection" involving unintentional lack of service (Lotem *et al.*, 1999). Although punishment of various sorts of non-cooperators are certainly features of human culture (Axelrod, 1984), our hypothesis of stability-dependent cooperation is an alternative explanation for the persistence of groups. We propose that the reason why cooperation does not automatically collapse or cycle in the way suggested by economic models is because such models have not yet taken into consideration the general human tendency towards *facultative* adherence to moral norms, and the resulting negative feedback on booms and busts of cooperation.

One's reproductive success can be advanced either by serving one's group (thereby slightly increasing the reproductive success of all group members relative to others) or by more immediately serving oneself (thereby increasing one's reproductive success relative to that of other group members). In many situations, these two options imply a continuum of behavioral options, or even two mutually exclusive options. This is the point at which many evolutionary studies of human behavior apply the terms "altruism" and "selfishness", a misleading dichotomy that begs the question of which course of action is in fact adaptive for an individual in a given situation. A more precise set of terms would reflect the fact that certain behaviors are adaptive because they increase the between-group component of reproductive success (enhance group stability), whereas others are adaptive because they increase the within-group component (increase

individual fitness relative to other group members). Most behavioral options probably affect both components of fitness, which necessitates a continuum rather than a dichotomy. For convenience, we will use the terms "group-service" and "self-service" to refer to the two possible directions of movement on such a continuum.

The principle of stability-dependent cooperation implies that an individual's assessment of group stability is a major determinant of behavior in situations where group-service and self-service prescribe different courses of action. The more group stability is threatened, the more group-service is likely to be adaptive for an individual, while the more successful or stable the group, the more adaptive is a relaxation of individual commitment to the group and an increase in self-serving behavior, because the detriment to group stability of such relaxation is small. Security in the persistence of the group renders restraint from self-serving strategies less critical. If this hypothesis is correct, no single level of cooperation is adaptive all the time. Thus, we propose that the adaptation is not cooperation per se, but the propensity to evaluate the optimal level of cooperation in a given situation.

One prediction from this hypothesis is that studies of moral judgments in social context (e.g., Carpendale and Krebs, 1992; 1995) will find perception of group stability (a variable hitherto untested) to be a significant determinant of the outcome of moral deliberation. Also, the large proportion of unexplained variation in cross-cultural studies of moral intuitions (e.g., 39-44% in O'Neill and Petrinovich, 1998) may be reduced by taking a measure of perceived group stability into account. First, behavior should more closely approach the ideal of golden rules of general beneficence in times and places where group stability is threatened, whereas when group stability is more assured, "Do unto others..." may still be a mantra, but the evidence should indicate a more competitive edge to intra-group interactions, and a tolerance of such competition in the community. Second, patriotism is a value that should be most emphasized and displayed when the nation is threatened, and actions that undermine national unity should be better tolerated in a time of peace than in time of war. Third, generosity and magnanimity should be higher within less stable groups, and lower within more stable groups, since when one's group is doing poorly, group members are "all in it together", and should be more disposed to share and affiliate, and self-service should be viewed dimly. A sense of

communal struggle will lose impact in times and places of group success, however, and individuals may get away with, and tolerate in others, more "materialistic" and competitive behavior such as hoarding and extravagance. Indirect evidence for such trends already exists. For example, adversity tends to increase individual commitment to causes that are central to a person's values (Lydon, 1990; Brickman, 1987).

*4.2. Group stability and indirect reciprocity*

Indirect reciprocity provides the means for individuals to gain information about the dedication of others to the group, act to minimize parasitism by freeloaders or cheats, and allocate benefits to individuals in proportion to their level of commitment (Alexander, 1987). If, as proposed in section 3, the dynamic of within-group cooperation in the context of intergroup competition maintains indirect reciprocity, determining the bases upon which reputations are made and broken, then the workings of indirect reciprocity may be expected to covary with group stability by a two-part mechanism.

First, one's assessment of group stability will affect the attitude that one has towards the behaviors of others. When group stability is under threat, people will be especially assiduous in assuring that others serve the group, both because that service can aid group stability which is a high priority, and because one's own restraint from self-serving action raises one's vulnerability to competitive exploitation by any group members who fail to exhibit similar restraint. Thus, unilateral group-service is unlikely to be adaptive for individuals with the means to behave self-servingly; rather, group-service will tend to be adaptive only when other members are also practicing it, and each individual has a stake in ensuring the cooperation of others.

When group stability is relatively assured, the principle of stability-dependent cooperation implies that people will be more permissive of intra-group competition, the balance thus shifting away from group-service. Of course, one would prefer that everyone except oneself maintain a high level of group service regardless of group stability; but because apparent hypocrisy is disproportionately damaging to one's reputation, one cannot pursue self-interest alone except by deception (which is risky) or despotism (which requires rare power). Absent these options, the best way for one's competition to be tolerated by the group is to foster tolerance of competition in general.

On the group level this dynamic pattern of individual attitudes means that community enforcement of group-service will be variably strict depending on how members of the community perceive its stability.

The second reason to expect covariation between indirect reciprocity and group stability is that the assessment of the actions and attitudes of others involved in indirect reciprocity will be important in determining the optimal range of individual commitment to the group. People look to others not only to enforce their commitment, but also to determine what level of commitment is required. As group stability changes, the tendency of the community to enforce service to the group will also change, producing positive feedback, as each assessor is also a participant; individuals must track all such changes, behaving differently as the community shifts in moral emphasis. Erring on the side of self-service will incur a reputational cost; erring on the side of group-service will incur a sucker's cost.

In short, the dynamics of indirect reciprocity are expected to covary with dynamics of group stability because (1) individuals benefit by assessing others in different ways depending on group stability, and (2) as these assessments change, individuals will benefit by accommodating their own behavior to those changes.

*4.3. Dynamic moral tension: an illustrative model*

The dynamic moral tension hypothesized here can be illustrated as follows. Consider a boat race in which a number of multi-rower craft race against each other over a predetermined course. Prize money is divided such that the first boat receives 50% of the total, the second boat receives 50% of the remainder, and so on down the standings. Within each boat, the position of each rower dictates what fraction of the boat's total winnings are his, such that the person in the first seat gets 50% of the boat's total, the person in the second seat gets 50% of what remains, and so on down the boat. (Astute readers will recognize that the prize allotment scheme leaves a small sum un-awarded, which we would argue is best spent on imaginary beer for all participants). The rules permit individuals, alone or in collaboration, to dislodge boatmates from superior seats, but, as a practical matter, this cannot be accomplished while rowing.

In such a race, one can easily imagine that as boats fall behind they will be overcome by a sense of shared fate and the need to cooperate intensely so that they will have something substantial to divide. As any boat pulls ahead, individuals in the back of that boat (who stand to gain little from the win) will reasonably conclude that bettering their own standing within the boat is the best use of their efforts. Cooperation between competitors in the back of the boat is likely to arise, and breakdown of such alliances is increasingly probable as they move closer to the front of the boat and more prize money is at stake.

*4.4. When groups collapse*

One exception to the trend of stability-dependent cooperation follows from the fact that humans, although highly dependent on their groups, are not absolutely so. When stability is so low that the group might be doomed to dissolution, group members may consider the benefits of leaving the group to be greater than the benefits of serving it. Moreover, as individuals cease striving for the group when they believe the cause is lost, they will be accelerating the collapse of the group both by their withdrawal of aid, and by the effects of that withdrawal on the assessments of others. This consideration indicates a threshold effect, with a sharp decline in cooperation and thus self-fulfilling group dissolution, once hopelessness of group persistence begins to spread. The existence of this tendency, however, depends on a perceived probability of successfully integrating into new groups following past group failures. Where there is no such hope, individuals would be expected to go down with the ship, continuing to employ the only strategy with any apparent chance of success.

*4.5. Zero-sum vs. nonzero-sum dynamics*

We propose that this model generates behavior readily recognizable to observers of human behavior. However, it only represents that portion of human endeavor that is characterized by conditions that are zero-sum or nearly so. Zero-sum conditions are those in which resource utilization is complete, so one individual can only benefit at a cost to others. Because humans, like all organisms, tend toward carrying capacity, near zero-sum dynamics as the typical state may be a plausible hypothesis. Nevertheless, the

dynamics of moments in human history are worth considering when resource limitation temporarily decreases, to the point where a rapidly growing population can nevertheless become dominated by strategies that ramp up the extraction of resources from the environment. In such situations (e.g., dispersal onto a newly discovered landmass), intra-group competition, though still a functional strategy for individuals, may bear disproportionate opportunity cost because effort spent on infighting could be spent instead on resource extraction (thereby increasing the overall size of the pie), a strategy that in these situations is more profitable, and likely to increase the future size of the group, thereby reducing future susceptibility of the group to intergroup threats. Despite the potential importance of nonzero-sum dynamics to human morality (Wright, 2000), the remainder of this paper will continue to focus on the more conventional zero-sum situation.

## 5. Morality as group stability insurance

### *5.1. Variation in commitment can endanger group stability*

The principle of stability-dependent cooperation predicts an inverse correlation between the stability of groups and the tendency of their members to adhere to moral rules. However, if the dynamic of human cooperation were this simple, there are at least two potential problems which, when significant, could cause group stability to deteriorate too quickly for individual behavior in service to the group to increase to counteract it.

First, factors influencing group stability can be extrinsic, and therefore not under the control of the group. When factors like resource limitation and intergroup competition act quickly, group destabilization may be difficult to anticipate and prevent. If this is true, high group stability and its concomitant low levels of service to the group, will lead to an increased vulnerability of the group to fast-acting extrinsic sources of group instability.

Second, positive and negative effects on group stability are asymmetrical. As with many organized structures, an individual has more power to affect group stability negatively than positively. Under circumstances favoring group stability, each cooperator restrains self-service for the sake of the group, but generally contributes to group stability only in a small way. However, if individuals jockey for position within

the group, they can initiate a rapid decline in group stability, as the prospect of exploitation shifts everyone's adaptive strategy away from group-service towards self-service. If moral rules are less important to people in times of group stability, and the usual restraints on within-group competition are relaxed, the opportunity would be created for individuals to compete to slightly exceed their neighbors' moral decay. An individual would attempt to gain the greatest possible benefits from the group's moral relaxation. The result would be a "race to the bottom", where the bottom is the breakdown of group-serving cooperation and the outright neglect of group stability.

Of course, the dependence of individuals on their group means that when the disastrous race began to threaten group stability, the interests of everyone would be served by reversing the trend and maintaining the group. However, in cases where everyone's interests are served by community action that is costly to each individual if unilaterally pursued, a tragedy of the commons results (Hardin, 1968). Everyone may continue to pursue actions that are beneficial to no one in the end, resulting in group destabilization. Indirect reciprocity is unlikely to be able to rescue a community from this situation. (Milinski et al. (2002) concluded otherwise, but the situation being described here is different from their experimental milieux. In actual societies, reputational costs and benefits may return too slowly to counteract the immediate benefits accruing to competitors in a race to the bottom.)

Thus, in both classes of hazard—extrinsic threats as well as races to the bottom—the fast-acting nature of the changes is what is expected to jeopardize stability in human groups.

*5.2. Morality buffers variation in commitment*

Functional systems are often buffered against perturbations that threaten their integrity. For instance, in contrast to some proteins such as MHC that experience rapid evolution, histone proteins, which are critical to the stability of all eukaryotic genomes, have evolved an extremely low mutation rate (Kornberg & Lorch, 1999). If fast-acting or unpredictable forces can threaten the stability of human groups, and if susceptibility to these forces is affected by group members' adherence to moral rules, then a buffering system may be in place in moral systems that lowers the risk from such threats to group

stability. Therefore, the second principle we propose is the *group stability insurance function of morality*, whereby *certain features of human morality are adaptive primarily because they buffer group stability against fast-acting threats*. The two major features of morality that may be serving this buffering function are absolutism and viscosity.

*5.3. The function of moral absolutism*

Individuals in groups where moral rules appear to be constant and always deserving of adherence should be less likely to discard or neglect those rules than individuals in groups where rules have no such absolutist qualities. We hypothesize, then, that the air of absolutism surrounding moral rules has been maintained in human culture because it buffers the changes in attitude and behavior that would be engendered by stability-dependent cooperation. In particular, absolutism works against the natural slippage of adherence to moral rules that occurs during times of group stability, decreasing the susceptibility of the group to sudden extrinsic threats and heading off the tendency for a rapid competitive race to the bottom. Moral groups, on this hypothesis, insure their stability by "consecrating" their rules in the minds of their members, just as political theorist Edmund Burke suggested political groups do (Burke, 17XX: §159-164). Hence, the adaptive dynamics of social groups provide the basis for an explanation of how humans benefit by associating their moral rules with the most sacred and authoritative aspects of their culture, despite facultative adherence to these rules. We are not aware that any other hypothesis addresses this apparent paradox.

*5.4. The function of moral viscosity*

The second feature of the adaptive buffering system that we propose to be in place in human moral systems is a viscosity with regard to moral rules. Viscosity in this sense is suggested by the old notion of moral *character*, the quality of individuals that is significantly influenced by habit, and slow to change attitudes and behavior patterns once developed (Kohlberg, 1964). If humans are resistant to change and susceptible to entrainment or habit formation in morality, they will be less likely to engage in rapid changes of commitment level that can compromise the efficacy of indirect reciprocity and ultimately threaten group stability. They will also be less likely to track drastic

fluctuations in perceived group stability, decreasing susceptibility to sudden extrinsic threats.

Available psychological evidence does suggest that moral attitudes are viscous in this sense (Eisenberg *et al.*, 2002; Kagan, 1989; Eysenck & Eysenck, 1963). In fact, this appears to be a relatively old feature of the human psyche that functions in a variety of other contexts, and so is not unique to the moral sphere. Nevertheless, if the above considerations are correct, moral character or viscosity together with moral absolutism can be explained biologically as a system providing insurance of group stability. The system buffers the impact of threats to group stability at the level of individual adherence to moral norms.

**6. Status and power inequality as a modifier of within-group moral variation**

Discussion to this point, while not assuming egalitarianism within groups, has not dealt with the variation in moral commitment that results from inequalities of power and status. At an authoritarian extreme, the effect of power disparity will swamp the effects of group stability, for the ability of most individuals to modify their level of service to the group will be very limited. In general, predictions like those at the end of section 4.1 are more applicable the more freedom individuals have to make behavioral decisions, and are best tested on behaviors that are not legislated or coerced except by community expectations. Even when individuals do have such freedom, status inequalities still probably modify the expected dynamics.

One set of examples of such complexity relates to the adaptive strategies of people in positions of power. Like other people, the powerful benefit when group stability is high, but they also have a special stake in promoting group service, both because they get disproportionate shares of the profits of collaboration, and because competition is much more likely to move them down the intra-group hierarchy than it is to move them up. Two tactics that are therefore likely to be employed by powerful members of stable groups are misinforming the group by understating group stability (perhaps by manufacturing or exaggerating threats), and enforcing group service through penalties. In the terms of the boat race model presented in section 4.3, those seated in the fronts of boats, and perhaps especially in the front of leading boats, will tend to

exaggerate the threat from competing boats since they have little to gain and everything to lose from intra-boat competition. Moreover, in certain cases such as in democratic groups where the continued tenure of leaders depends on perceptions of their having done a good job, leaders may gain from preaching the high stability of the group, at least relative to when they were not leaders. These considerations illustrate the importance of distinguishing between actual and perceived group stability in predicting the optimal degree of group service. Such status-by-stability interactions also become important when stability is dangerously low. At some low threshold of group stability, group members may do better by leaving, but leaders will benefit by keeping others in the group, providing an incentive to misinform group members that the group is more stable than it is.

## 7. The multiplicity of social groups

This discussion has portrayed moral deliberation as unidimensional, from self-service to group-service. In fact the moral landscape is more complicated, since individuals belong simultaneously to different groups. Some may overlap, such as an ethnic group and a workers' union, and some are concentric, such as a neighborhood within a city within a state within a nation, but still others may be quite distinct in their domains of relevance. The rule of stability-dependent cooperation can certainly be dealt with on the simple continuum of self-service to group-service, but the actual decisions faced by individuals may often be a matter of how large a group to align with in a particular situation (the self being one end of that continuum), or which of two group memberships to prioritize (Mason, 1996). Analysis focused on one group identity may misinterpret prioritization of another as defection towards self-service.

## 8. Relation to two other perspectives on the evolution of morality

### *8.1. Cultural evolution of memes*

The hypotheses introduced above treat facultative adherence to morality as a trait that is adaptive. This contrasts with the view that the persistence of cultural elements, or "memes" (Dawkins, 1976), is unrelated to individual reproductive success. In the view of Dawkins and others, memes need not increase the bearer's inclusive fitness to persist;

rather, they evolve and adapt to each other in an autonomous system. This viewpoint surely has some utility in the short term, and individual cultural elements may subvert the interests of others and of the genes. However, the perspective that has led to the principles of stability-dependent cooperation and the group stability insurance function of morality is based on the assumption that radical separation of culture from individual fitness cannot commonly be the case in the long run.

Since the capacity for culture is genetic in evolutionary (and developmental) origin, in order to have become fixed in our species the system must have returned benefits to the genomes of its bearers throughout the period of its elaboration. The net average fitness effect of all memes, genetically speaking, must therefore have been positive. At whatever point the net effect of all memes on genes becomes negative, natural selection should disassemble the genetic capacity for memes. Recognizing this, most will acknowledge that memes must have been fitness enhancers early in their evolution, but some will claim that the system has more recently become autonomous, and memes need only be neutral in fitness impact (on average) in order to persist independent of genetic interests. But because all behaviors take time and effort and therefore have an opportunity cost, memes that are not beneficial are expected to be short-lived. When cultural elements are widespread and persistent, they are likely to have become so because they tended to benefit their bearers.

*8.2. Group selection*

The hypotheses developed here clearly depend on the interests of groups of individuals. Nevertheless, we have not relied on the group selection perspective (where natural selection is discussed at various levels including the social group, as if selection at multiple levels reflects multiple evolutionary mechanisms). In fact, even proponents of multi-level selection admit that selection at various levels can be reduced to a single mechanism (Sober & Wilson, 1998). An individual-level perspective in evolutionary discussions of human groups therefore has the advantage of discussing the complexity of natural selection without an apparent proliferation of evolutionary mechanisms. Moreover, as mentioned in section 7, humans are members of various social groups that are not always concentric as the notion of a "level" of selection would imply. Finally,

our individual-level perspective also refrains from making assumptions about the degree to which groups replicate or otherwise resemble genomes (Williams, 1966). Nevertheless, the hierarchical relationship between the competitive success of the group and that of the individual remains explicit in these hypotheses and the perspective underlying them. Individual fitness is considered overwhelmingly dependent on group persistence, and groups are assumed to vary in their capacity to persist. From this perspective, selection on the individual results in the individual's capacity to prioritize among several avenues of potential fitness return.

**9. Conclusion**

We have drawn implications from the understanding that humans, as social animals who are nevertheless genetically individualistic, must strike a balance between strategies for competition within a group, and strategies for increasing group stability. This assessment of the human situation follows from the evolutionary theory of human culture and morality developed by Alexander (1987, 1990, 1992, 2004).

In particular, we have proposed that much observed variation in commitment to moral norms is explained by a rule of stability-dependent cooperation, where the adaptive level of individual commitment is a function of the stability of the social group. If this is true (and the predictions from this hypothesis are numerous), variability in human moral commitment reflects our ability to track variation in the expected benefits of competition versus cooperation. However, groups in this situation would still be susceptible to fast-acting extrinsic threats as well as self-destructive competitive races to the bottom. In light of this, we have proposed that the absolutism and unchangeableness that people attribute to moral norms, features that have bewildered moral philosophers for centuries (Williams, 1985), might function as group stability insurance against fast-acting threats.

These hypotheses deserve testing for at least three reasons. First, they explain why morality has an air of absolutism despite the facultative nature of human commitment to moral rules. Second, they resolve the longstanding debate about the adaptive status of moral rules, by placing the locus of adaptation not in particular kinds of acts, but in the moral agent's ability to weigh options and choose a commitment strategy based on the current social environment. Third, they connect moral attitudes to

environmental variables, and thus have the potential to explain hitherto perplexing moral variation within and between individuals and cultures.

**Acknowledgements**

This paper was co-authored with David Lahti, and published in the proceedings following the festschrift for Dick Alexander, as follows: Lahti DC, Weinstein BS. 2005. The better angels of our nature: group stability and the evolution of moral tension. *Evolution and Human Behavior* 26:47-63.

**Literature Cited**


Alexander, R.D. (1979). *Darwinism and Human Affairs*. Seattle: University of Washington Press.

Alexander, R.D. (1987). *The Biology of Moral Systems*. Hawthorne, NY: Aldine de Gruyter.

Alexander, R.D. (1989). Evolution of the human psyche. In *The Human Revolution*. (Eds.) Mellars, P., and Stringer, C. Edinburgh: University of Edinburgh Press, pp. 455-513.

Alexander, R.D. (1990). How did humans evolve? Reflections on the uniquely unique species. *University of Michigan Museum of Zoology Special Publication*, *1*, 1-38.

Alexander, R.D. (1992). Biological considerations in the analysis of morality. In *Evolutionary Ethics*. (Eds.) Nitecki, M.H., and Nitecki, D.V. Albany, NY: SUNY Press, pp. 163-196.

Alexander, R.D. (2004). Evolutionary selection and the nature of humanity. In *Darwinism and Philosophy*. (Eds.) Hosle, V., and Illies, C. Notre Dame, IN: University of Notre Dame Press.

Axelrod, R. (1984). *The Evolution of Cooperation*. New York: Basic Books.

Brickman, P. (1987). *Commitment, Conflict, and Caring*. Englewood Cliffs, NJ: Prentice-Hall.

Burke, E. (1790). *Reflections on the Revolution in France*.

Buss, L.W. (1987). *The Evolution of Individuality*. Princeton, NJ: Princeton University Press.

Carpendale, J.I.M. and Krebs, D.L. (1992). Situational variation in moral judgment: in a stage or on a stage? *Journal of Youth and Adolescence*, *21*, 203-224.

Carpendale, J.I.M. and Krebs, D.L. (1995). Variations in level of moral judgment as a function of type of dilemma and moral choice. *Journal of Personality,63,* 289-313.

Cohen, D. and Vandello, J.A. (2001). Honor and 'faking' honorability. In *Evolution and the Capacity for Commitment*. (Ed.) Nesse, R. New York: Russell Sage Press, pp. 163-185.

Cosmides, L. and Tooby, J. (1992). Cognitive adaptations for social exchange. In *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. (Eds.) Barkow, J.H., Cosmides, L., and Tooby, J. New York: Oxford University Press, pp. 163-228.

Darwin, C. (1871). *The Descent of Man and Selection in Relation to Race*. London: John Murray.

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

Edwards, C.P. (1975). Societal complexity and moral development: a Kenyan study. *Ethos*, *3-4*, 505-527.

Eisenberg, N., Guthrie, I.K., Cumberland, A., Murphy, B.C., Shepard, S.A., Zhou, Q. and Carlo, G. (2002). Prosocial development in early adulthood: a longitudinal study. *Journal of Personality and Social Psychology*, *82*, 993-1006.

Eysenck, S.B.G. and Eysenck, H.J. (1963). The validity of questionnaire and rating assessments of extraversion and neuroticism, and their factorial stability. *British Journal of Psychology*, *54*, 51-62.

Frank, S.A. (2003). Perspective: repression of competition and the evolution of cooperation. *Evolution*, *57*, 693-705.

Gintis, H. (2000). Strong reciprocity and human sociality. *Journal of Theoretical Biology*, *206*, 169-179.

Hardin, G. (1968). The tragedy of the commons. *Science*, *162*, 1243-1248.

Hartup, W.W. and van Lieshout, F.M. (1995). Personality development in social context. *Annual Review of Psychology*, *46*, 655-688.

Huxley, T.H. (1894). Evolution and Ethics. In *Evolution and Ethics* London: Macmillan and Co., pp. 46-86.

Kagan, J. (1989). *Unstable Ideas: Temperament, Cognition, and Self*. Cambridge, MA: Harvard University Press.

Keith, A. (1949). *A New Theory of Human Evolution*. New York: Philosophy Library.

Kohlberg, L. (1964). Development of moral character and moral ideology. In *Review of Child Development Research, Vol. 1*. (Eds.) Hoffman, M.L., and Hoffman, L.W. New York: Russell Sage Foundation, pp. 383-432.

Kornberg, R.D. and Lorch, Y. (1999). Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell*, *98*, 285-294.

Lotem, A., Fishman, M.A. and Stone, L. (1999). Evolution of cooperation between individuals. *Nature*, *400*, 226-227.

Lydon, J.E. (1990). Commitment in the face of adversity: a value affirmation approach. *Journal of Personality and Social Psychology*, *58*, 1040-1047.

Lyon, M.F. (2003). Transmission ratio distortion in mice. *Annual Review of Genetics*, *37*, 393-408.

Mackie, J.L. (1981). *Ethics: Inventing Right and Wrong, ch.1*. Harmondsworth: Penguin.

Maienschein, J. and Ruse, M. (eds.) (1999) *Biology and the Foundations of Ethics*. New York: Cambridge University Press.

Mason, H.E. (1996). *Moral Dilemmas and Moral Theory*. New York: Oxford University Press.

Milinski, M., Semmann, D. and Krambeck, H.-J. (2002). Reputation helps solve the 'tragedy of the commons'. *Nature*, *415*, 424-426.

Nesse, R. (2001). Natural selection and the capacity for subjective commitment. In *Evolution and the Capacity for Commitment*. (Ed.) Nesse, R. New York: Russell Sage Press, pp. 1-47.

Nisan, M. and Kohlberg, L. (1982). Universality and variation in moral judgment: a longitudinal and cross-sectional study in Turkey. *Child Development*, *53*, 865-876.

Nowak, M.A. and Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. *Nature*, *393*, 573-577.

O'Neill, P. and Petrinovich, L. (1998). A preliminary cross-cultural study of moral intuitions. *Evolution and Human Behavior*, *19*, 349-367.

Peters, K.E. (2003). Pluralism and ambivalence in the evolution of morality. *Zygon*, *38*, 333-354.

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

Ridley, M. (1996). *The Origins of Virtue*. London: Penguin.

Roes, F.L. and Raymond, M. (2002). Belief in moralizing gods. In Amsterdam, The Netherlands.

Ruse, M. (1986). *Taking Darwin Seriously: A Naturalistic Approach to Philosophy*. Oxford: Blackwell.

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

Trivers, R. (1971). The evolution of reciprocal altruism. *Quarterly Review of Biology*, *46*, 35-57.

Williams, B. (1985). *Ethics and the Limits of Philosophy, ch.1-5*. London: Fontana.

Williams, G.C. (1966). *Adaptation and Natural Selection*. Princeton: Princeton University Press.

Williams, G.C. (1988). Huxley's *Evolution and Ethics* in sociobiological perspective. In *Issues in Evolutionary Ethics*. (Ed.) Thompson, P. Albany, NY: SUNY Press, pp. 317-349.

Wright, R. (2000). *Nonzero: The Logic of Human Destiny*. New York: Vintage Books.