ROBUSTNESS IN BIOLOGICAL AND SOCIAL SYSTEMS

J. C. FLACK, P. HAMMERSTEIN, AND D.C. KRAKAUER

1. Abstract

This is a draft. References will be added to the final version. In the meantime a brief list of recommended papers is provided at the end. In the final version of this MS, we will also add a brief discussion of the relation between robustness and the levels of selection problem, and we will expand on the discussion of robust decision-making rules.

2. Introduction

The large number of parts and interactions characteristic of biological and adaptive systems in general ensures that there are many ways in which a system might fail. Yet the number of defective individuals or mutants observed in nature remains relatively low. One trivial reason for this is that mutants are rarely viable and so eliminated before we observe them. A second reason is that biological systems have evolved means for correcting malfunctions and errors and are consequently *robust* to mutation and other kinds of perturbations. A system can be said to be robust if one or more of its structural or functional properties remains invariant despite exposure to nontrivial perturbations.

Robustness is well studied in biology, ecology, and engineering. In each of these disciplines the importance of maintaining performance in spite of a changing or unpredictable environment, component error and failure, and component conflict is well recognized. In addition, within biology and ecology, a fundamental goal has been to account for the origins of complex structures. The layering of robustness mechanisms in "anticipation" of fault and failure has been offered as an explanation for hierarchical complexity (sometimes called robust-over-design), generating many studies on the relation between robustness, complexity, and evolvability.

Studies on social evolution, and social systems more generally, have lagged behind biology in addressing robustness issues. Reasons for this include: (i) in social evolution the focus has been on trait economics –the cost and benefit scenarios that support the evolution of alternative behavioral strategies, often at the expense of studies that address *how* social systems and their constituent strategies persist in time and related mechanistic questions, and (ii) the recognition that the rigorous study of robustness requires carefully controlled systematic experiments that are hard to conduct on social systems.

In this chapter we discuss how ideas of robustness can inform our understanding of social evolution and social decision making and how robustness can be investigated. We provide a basic introduction to robustness concepts and introduce several generic robustness principles. We conclude with a discussion of how these principles might shed light on trait economics and the evolution of individual strategies and decision-making rules. To ground the discussion, we begin with a brief history of the robustness concept as it developed in the study of living systems.

3. Brief History of Robustness

The study of robustness finds it origins in the study of the phenotype. There have been two major trajectories in evolutionary theory with regard to the evolution of phenotypes. Although an oversimplification, we might term one the functional perspective and the other a developmental or constraints perspective. The functional perspective, consolidated during the Modern Synthesis, has focused on selective or mutational processes that neglect many details of the organism. The emphasis has been the origins and maintenance of genetic and phenotypic diversity, and secondarily on adaptation, either to an environment that changes slowly or to other strategies in a population that change quickly (frequency dependence). It is from this perspective that sociobiology, evolutionary game theory, behavioral ecology, and eventually evolutionary psychology grew.

The developmental or constraints perspective has focused on "self-organization"– the interaction between the dynamics of metabolic and physiological processes with the developmental process. Within this perspective are broadly three traditions. One associated with D'Arcy Thompson, another with C.H. Waddington, and a third with Theodore Boveri and Richard Goldschmidt. Intellectual descendants of D'Arcy Thompson focus on the constraints on phenotypic space imposed by physical laws –the scaling work of West, Brown, and Enquist is a modern example. The "Waddington camp" is interested in what might be called evolved constraints or frozen accidents. This includes the study of canalization–a developmental strategy for dealing with uncertainty in which an organism essentially makes itself impervious to fluctuations in the environment. Goldschmidt-Boveri decendants pursue the idea that organismal transformation, and hence changes to phenotypes, results from the reorganization of genes and is essentially a combinatoric and regulatory problem [**?**]. The questions of evolved constraints, canalization and regulation all bear on issues of robustness.

In general, the constraints perspective has been in the domain of the mechanistic and experimental biologist, and the functional perspective in the domain of the evolutionary biologist. With respect to robustness, the two threads began in the 1990's and early 2000's to merge as students of evolutionary theory recognized that robustness mechanisms are one of the

bridges connecting the dynamics of ontogeny with the dynamics of phylogeny. This is because robustness mechanisms limit phenotypic variation and thereby the strength of selection, and simultaneously provide a means of exploring alternative genotypes without compromising the phenotype, that is they promote neutrality [**?**].

4. Robustness: Basic Issues

When considering robustness it is worth bearing in mind the plethora of definitions the concept attracts. For an extensive list see Jen[**?**]. These definitions are often domain-specific. In ecology, stability or robustness is a measure of the preservation of species diversity upon species removal[**?**] or, the permanence of a trophic networks when perturbing variables of ecological interest. In medicine, robustness is associated with healing and compensation, neither of which imply invariance of phenotype but rather a restoration of a "wildtype" function [**?**] . In linguistics robustness relates to competence and comprehensibility despite incomplete information and ambiguity [**?**]. In paleontology robustness relates to the continuity of lineages across geological eras [**?**], and the persistence or rebound of lineages during mass extinction events. In metabolism, robustness relates to limited phenotypic variation with large changes in kinetic parameters [**?**], [**?**]. In cell biology robustness can describe how cell fate decisions remain constant when transcription regulation is stochastic [**?**]), or how RNA secondary structures are preserved when experiencing point mutations [**?**].

In each of these cases robustness relates to either: (1) non-detectable or minor modification in a "phenotype" following a large perturbation to the genotype, (2) non-detectable or minor modification in phenotype following a large perturbation to the phenotype from the environment (3) non-detectable or minor modification in function following a large perturbation to the genotype or phenotype with or without a correlated change in the phenotype. The important distinction between genotypic and phenotypic or environmental robustness is that in the first case perturbations are inherited, whereas in the second case, they are not. An additional complexity is that functional robustness can be achieved through phenotypic invariance or phenotypic plasticity. In the first case the phenotype resists change, and in the second case, the phenotype tracks perturbations. Genotypic and environmental robustness are often measured through the environmental (V_e) or mutational variance (V_m) of a trait, whereas functional robustness can be measured as the variance in geometric mean fitness. It is often the case that a single mechanism leads to all three forms of robustness in which case we observe *congruence* [**?**] between the genotype and phenotype-mechanisms that promote a correlated response to genotypic and environmental perturbations.

We began this paper by stating that a system can be said to be robust if one or more of its structural and functional properties remains invariant despite exposure to nontrivial perturbations. This can be stated more formally as, *system property S is robust if it remains invariant under a perturbation P that removes, disables, or otherwise impairs the performance of adaptive component C* (for further discussion, see Ay references in Recommended Reading). This statement, which captures the essence of what most researchers mean when they employ the term "robustness," is applicable to most of the examples given in the first paragraph of this section. Below we break each aspect of this statement down, giving examples to make the definition more concrete. In doing so, we provide a basic introduction to a few of the issues that arise in the systematic study of robustness.

4.1. **Perturbation.** Perturbations are exogenous or endogenous factors that increase disorder within a specified system. The range of perturbations can include environmental factors like UV radiation and internal factors like the propensity of a cell to mutate, generate stochastic outputs or fail. Chronic conflict between components arising from imperfectly aligned interests and competing objectives is an important source of functional errors. To be considered nontrivial, a perturbation must be capable, given its frequency and magnitude, of having an effect on that system. For example, if one person throws from four feet a paper airplane at another person, the perturbation is unlikely to injure or even be perceived. This might be considered a trivial perturbation. On the other hand, a large rock thrown from a comparable distance, is likely to be noticed. Ideally, the relative magnitude and frequency of the perturbation should be quantified and defined with in relation to the system component it is perturbing.

When a perturbation has no effect on a system despite being of a causally significant scale, there are two possible explanations. One is that there is a robustness mechanism protecting the target component, or the function it contributes to. The other is that the target of the perturbation is not causally important, as discussed below.

Knockout and knockdown are techniques used primarily in genetics and cell biology for disabling parts of systems – perturbations administered by the experimentalist. Often these are second order perturbations of putative robustness mechanisms and need to be administered conjointly with a perturbation that the mechanism is thought to buffer against in "the wild." Hence most robustness experiments involve administering a minimum of two perturbations. The number of perturbations that need to be administered experimentally will also vary depending on whether the goal is to determine whether the system is robust to a specified perturbation or whether a posited robustness mechanism is operational.

Natural perturbations (perturbations that occur in the "wild") in the context of system vulnerabilities can be interpreted in terms of the selection pressure driving the evolution of robustness mechanisms.

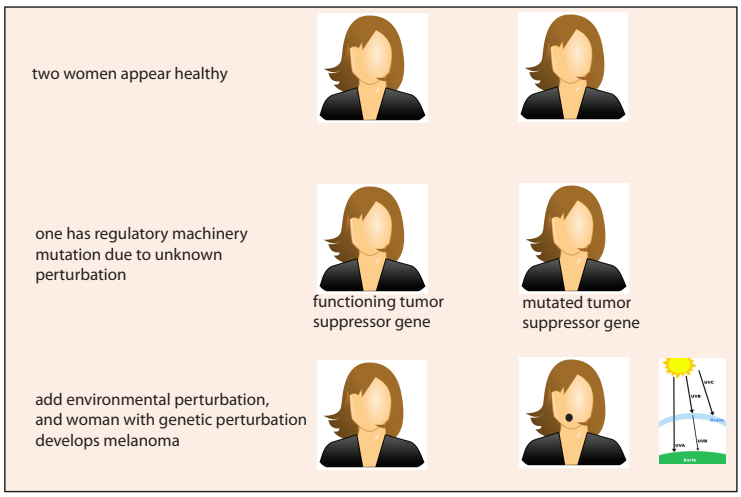

FIGURE 1. If the system has evolved robustness mechanisms, coincident dual perturbations will be required to reveal a lack of robustness. In the case shown here, one women possesses a robustness mechanism –a functioning tumor suppressor gene (TSG) that prevents rogue cells from proliferating and causing cancer. The second women's TSG is damaged. When exposed to UV radiation, the second woman develops a melanoma, indicating a lack of robustness against this perturbation, in this case because the robustness mechanism has failed *and* the UV radiation caused a cell to begin proliferating *and* this proliferation detrimentally affected some critical physiological function. An experiment with this two perturbation design can also be used to show that a posited TSG is in fact a robustness mechanism as long as it has been demonstrated independently that the the second perturbation –in this case, the UV radiation –can cause cells to proliferate uncontrollably.

4.2. **Functional Contribution and Exclusion Dependence.** An important consideration in the study of robustness is the magnitude of the *functional contribution* made by system component that is the target of a perturbation. Colloquially the functional contribution of a component is its relevance or adaptive value when it is unperturbed and performing "normally." To illustrate why this an important concept, consider a social network in which some nodes have a degree of 0, meaning they are disconnected from the remaining nodes, and in which the only way nodes can influence either the behavior of other nodes or global network properties is through their connections. We expect a perturbation to a node with degree

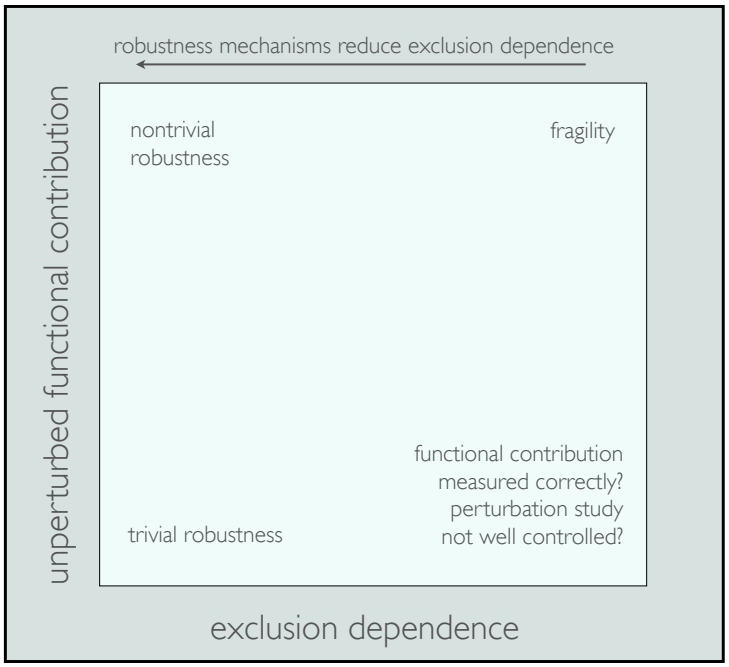

FIGURE 2. Functional Contribution and Exclusion Dependence.

0 to have no effect because it makes no functional contribution. At best the system can be said to be trivially robust to perturbations on such components. For further information see Ay references in Recommended Reading list at end of paper.

A second critical concept is *exclusion dependence* – that is, how much some specified structural or functional property of the system changes in response to a perturbation that impairs functionally relevant system components. As shown in Figure 2, the extent to which a system property is robust can be measured by taking into account these two quantities. Robustness mechanisms minimize exclusion dependence (see Ay references in Recommended Reading).

4.3. **Types of questions.** There are three questions that biologists typically ask when studying robustness.

(1) Is X a robustness mechanism (does it reduce the exclusion dependence measured on some system property S given a perturbation P that effects a functionally relevant component C?).
(2) Is system property S robust to a perturbation P affecting component C?
(3) Is system property S robust to a perturbation P that disables a known robustness mechanism C_r?

Each of these questions requires a different experimental design, in terms of the number of perturbations that must be experimentally administered and the timescales over which the consequences of the perturbations must be assessed. For example, to determine the robustness value of a posited robustness mechanism, the posited mechanism must be disabled so that its robustness contribution can be inferred from the effects of its removal. The knockout must be temporary to prevent the system from adapting – from building a new robustness mechanism or activating a latent, backup mechanism, as this would confound the results. In contrast, to determine whether the system is robust to a specific perturbation, a "long timescale" knockout is required in order that time-series data can be collected to determine whether the system destabilizes, restores function redundantly, or reconfigures establishing a new robustness mechanism.

4.4. **A Word of Caution.** No biological or social system is robust to all perturbations. In some cases, a perturbation that has negative consequences on one component can have neutral or even positive consequences on another. Robustness at one level of a system –the robustness of some global property– might even require sensitivity of components –a lack of robustness– at lower levels [**?**]. Hence in studying robustness, it is critical to always be specific about the perturbation, the component subject to the perturbation, the system property thought to be jeopardized by the perturbation, and the robustness mechanism that buffers the system against the perturbation.

5. Robustness Principles

Krakauer and Plotkin [**?**] [**?**] discuss three *principles* that have arisen in an effort to understand the evolutionary response to mutations. The principle of canalization, the principle of neutrality and the principle of redundancy. They contrast these with the *parameters* of robustness –those mechanisms by which these principles are realized. The principles and parameters metaphor is derived from linguistics [**?**] where the principles are the invariant properties of universal grammar and the parameters the local rules and practices of language. A focus on principles allows that several mechanisms varying parametrically but employing some form of redundancy can be classed together, and the same would be true for modularity and so on.

Here we extend the Krakauer-Plotkin list to give a brief introduction to robustness principles commonly invoked in biology. We then briefly discuss the relation of robustness and canalization to niche construction and the extended phenotype with the goal of clarifying a common misapprehension about what makes for a robustness mechanism rather than a common garden adaptive trait.

5.1. **Redundancy.** A common means of identifying the function of a gene is to perform a knockout experiment, removing or silencing a gene early in development. By assaying the resultant phenotype, the putative function

of the absent gene can be inferred. In many such experiments, there is no scoreable phenotype: the knockout leaves the phenotype in the wildtype condition. Biologists refer to a gene x on a background y as functionally redundant [**?**]. This is taken to mean that the target gene is one of at least two or more genes contributing to the phenotype epistatically [**?**].

Removal of a redundant gene x leads to compensation by other genes, the compensatory part of the genome. Let $f(\mathbf{y})$ be the fitness of a part $\mathbf{y}$ of a genome and $f(x, \mathbf{y})$ the fitness of this part with an additional gene x. If $f(x, \mathbf{y}) = f(\mathbf{y})$, this means that gene x is redundant with respect to part y. In the simplest case, y consists of a single gene being a redundant copy of x.

True redundancy might be rarer than "artifactual" redundancy, or experimental neutrality, in which the effect of perturbation remains below an experimental detection limit [**?**]. Assuming that we are able to detect small changes, the degree of redundancy describes the degree of correlation among genes contributing to a single function. Models of redundancy in biology tend to focus on the evolutionary preservation of redundant components - and hence employ population genetics approaches. More recently differential equation based models for the dynamics of regulatory systems following structural perturbation have also been explored [**?**].

5.2. **Purging – antiredundancy.** Whereas redundancy buffers the effect of perturbation, purging acts in the opposite fashion – amplifying the effects of perturbation – so as to ensure the purity of a population by making perturbed genomes more vulnerable to selective forces. Consider a genetic background $\mathbf{y}$, a particular gene z^+ and a deleterious mutation z^- of this gene. Let the fitness of an organism with this genetic background and the deleterious mutation z^- be denoted by $f(z^-, \mathbf{y})$, so that the fitness of the wildtype is greater than that of the mutation, $f(z^+, \mathbf{y}) > f(z^-, \mathbf{y})$. An additional gene x is called functionally antiredundant if its presence reduces the fitness of the mutant even further, $f(z^-, \mathbf{y}) > f(x, z^-, \mathbf{y})$. In the more general case, z can also consist of more than one gene.

Purging is only effective when individual replication rates are sufficiently large to tolerate the effects of removal of defective components. Thus apoptosis – programed cell death – is a common strategy for eliminating cells upon damage to their genomes or upon infection, provided these cell types are capable of regeneration. Nerve cells and germ cells produce factors that strongly inhibit apoptosis [**?**], as removal in these cases, has deleterious consequences. In case of severe infection it can make sense ot purge nerve cells [**?**].

Recent models dealing with purging-type phenomena have involved stochastic models assuming finite populations. The key insight from the study of antiredundancy is the ability of particulate, hierarchical systems to exploit cellular turnover to eliminate and replace deleterious components from populations.

5.3. **Multiple Pathways.** Organisms are surprisingly robust against the loss of single genes. Additionally, knockout studies investigating the benefits of duplicate genes find duplication accounts for substantially less phenotypic invariance than originally assumed. These observations suggest that phenotypic invariance is not the consequence of redundancy but of some other feature of the organism. One finding from the study of metabolic networks suggests an explanation. Metabolic networks that have evolved in fluctuating environments show high mutational robustness and are characterized by a greater numbers of multifunctional enzymes and independent paths [**?**, **?**].These findings suggest that robustness is a positive byproduct that is achieved as a consequence of building multiple pathways in response to a selection pressure to be able to process different kinds of metabolites due to a fluctuating environment.

5.4. **Feedback Control.** Elementary feedback control systems have three components: *a plant* (the system under control), *a sensor* (measuring the output of the plant) and *a controller* (generating the plants input) [**?**]. A measure of performance is often the degree to which the output of a plant approximates some function of the input to the controller. In biology a plant could be, RNA or protein concentration, protein kinase activation, immune effector cell abundance, or species abundance. Inputs in each of these cases would be transcription factors, protease concentrations, chemical agonists bound to receptors, antigen concentrations and death rates. The controllers are more often than not aggregates of several mechanisms. Feedback is a mechanism of robustness as it enables plants to operate efficiently over a range of exogenous input values. The question remains as to whether the controller is robust to variations in the plant - does it provide *robust stability*? For example, in biology, can a single feedback controller regulate the concentrations of several different proteins?

The theoretical literature in linear feedback control is very well developed in engineering. Biology has borrowed extensively from this literature. Nonlinear feedback control is another issue, and there are few canonical models [**?**].

5.5. **Modularity.** Repeated representations of functionally distinct character complexes capable of recombination or shuffling is an example of modularity. Modularity aims to capture structures that balance autonomy and integration. Within a complex there is strong integration, whereas populations of complexes are only weakly coupled [**?**]. This has also been called near decomposability [**?**] . In genetics, modularity involves a minimum of pleiotropy, in which sets of genes contributing to one complex or trait (for example organ system), make little contribution to other complexes or traits [**?**], [**?**] [**?**]. These modular genetic systems are found in different genomic contexts performing a similar function. Of course modularity can be defined at levels of organization above that of the gene [**?**] - the extent to

which organs operate independently during homeostasis. The dissociability of modules provides one means of damage limitation through encapsulation.

There are no collectively agreed upon models for analysing modularity in biosystems. To date quantitative genetics models have been used to explore the limits to the evolution of modularity and neural network models have been used to explore how modularity can lead to more efficient task management [**?**]. For a comprehensive review of the recent literature in the area of modularity see [**?**].

5.6. **Spatial Compartmentalization.** Compartmental systems are those made up from a finite number of macroscopic subsystems called compartments, each of which is well mixed. Compartments interact through the exchange of material [**?**]. The spatial compartmentalization of reactions leads to robustness by minimizing covariance among reaction components participating in functionally unrelated processes. Thus spatial de-correlation through compartmentalization substitutes for temporal correlation in biological functions. Robustness is achieved in at least two ways: (1) minimizing interference - chemical, epistatic or physiological, and (2) minimizing mutual dependencies and thereby attenuating the propagation of error through a system. The study of spatial compartmentalization is particularly rich in theoretical ecology and epidemiology [**?**] where it has been used to explore the maintenance of antigenic diversity, restrictions on pathogen virulence, and seasonal forcing. And more recently in molecular biology where proteins have found to be compartmentalized [**?**].

From a modeling perspective, compartmentalization is often approached from the perspective of metapopulation dynamics or coupled oscillators, in which space is assumed to be discrete (implicit space) and non-local [**?**]. An alternative approach is based on continuous space (explicit space) with local interactions and employs partial differential equations to study diffusion and advection of components [**?**]. A third approach assumes discrete space with local interactions employing coupled map lattices and cellular automata. A fourth approach analyzes the statistical connectivity properties of undirected graphs and their response to node or edge elimination [**?**].

5.7. **Distributed Processing.** Distributed processing describes those cases in which an integrated set of functions are carried out by multiple, semi-autonomous units [**?**], [**?**]. The most obvious example is that of nerve cells comprising the nervous system. Distributed processing, or connectionism, might be assumed to be a combination of modularity and spatial compartmentalization, but differs in that a single function is emergent from the collective activities of units, and correlated activity is thereby a desired outcome.

The robustness properties of connectionist models are: (1) the ability to identify incomplete patterns, (2) generalize from a subset of learnt patterns, and (3) degrade gracefully upon remove of individual nodes.

Connectionist models range from a simple application of linear algebra, dynamical systems and Hamiltonian representations of steady states, through to the use of statistical mechanics models of frustrated systems such as spin glasses.

5.8. **Conflict Management.** In most biological systems, component or individual interests are only partially aligned, leading to chronic conflict. This conflict can jeopardize performance and the formation of stable interaction networks critical to resource extraction and production[**?**]. Mechanisms have evolved in many biological systems to manage conflict, keeping it at levels thought to be useful for state-space exploration that might support adaptation when the environment changes[**?**]. Many of these conflict management mechanisms serve as robustness mechanisms that buffer resource extraction and production against the damaging effects of chronic conflict. Examples include third-party policing in animal societies, whereby individuals uninvolved in a dispute impartially intervene into the dispute, usually at low cost to themselves and low immediate cost to the conflict participants, and causing the conflict participants to disperse [**?**].

5.9. **Error-correction and Repair.** Another strategy for buffering against perturbations is to directly repair the damage perturbations induce in components. For example, in the cells of our bodies environmental factors and metabolic activity can cause damage to DNA. This damage can compormise gene transcription or the survival of cells after cell division. Organisms have also evolved means for repairing damaged DNA. These mechanisms, called DNA repair mechanisms, range from chemically reversing damage to removing damaged sections of one strand of DNA and re-synthesizing the missing piece using the second undamaged strand as a template. In the case of double strand damage, this requires identifying an identical or nearly identical sequence and using this as the template.

In animal social systems, social relationships can be damaged by conflict. A common mechanism of repair that restores the quality of relationships to baseline (pre-conflict) levels, is reconciliation[**?**]. Reconciliation typically involves fight participants showing increased social investment in each other immediately following conflict. Presumably this increased social investment serves as a costly behavioral signal the function of which is to reestablish mutual trust.

5.10. **Relationship to Niche Construction and Adaptive Traits More Generally.** A major challenge faced by biological and social systems is environmental uncertainty. There are two classes of strategies for dealing with environmental uncertainty. When the environment is largely unpredictable– meaning the timeseries of environmental states is characterized by minimal regularity, organisms make themselves impervious to environmental fluctuations by canalizing development. In other words, a developmental program

evolves that unfolds along a rigid trajectory, producing an invariant phenotype, regardless of environmental fluctuations. Another way to say this is that randomness is countered with a near deterministic buffer.

Alternatively, when an environmental process is characterized by sufficient regularity, organisms can evolve means for intervening directly into the environmental process and modify it. This involves changing the rate or trajectory of environmental events in order to make them more predictable and compatible with an existing phenotype. A benefit of this strategy is that it permits an organism's fit to the environment to be more finely tuned, resulting in a higher exchange of information between the organism and environment (increased adaptability). Strategies of this type include *building extended phenotypes* and *niche construction*, whereby individuals by modifying the environment, to some extent control the selection pressures to which they are subject. These would include bird nests, termite mounds, and spider webs[**?**], and can also include social structures like power structures [**?**]

Canalization and niche construction, both of which are strategies for reducing variance represent two extremes for reducing uncertainty on an ontogenic continuum. Whereas robustness mechanisms might appear to resemble more closely canalization, most robustness mechanisms observed in nature fall between the two extremes. Niche construction, where elements of the environment are encoded in the organism's genome, can be shown in some limit to lead to the evolution of an extended developmental program. And it is often through the requirements of robustness that these mechanisms evolve.

If robustness mechanisms include uncertainty reduction mechanisms from niche construction to canalization, what kinds of adaptive strategies are not instances of robustness? Or what distinguishes a robustness mechanism from adaptive mechanisms more generally? All robustness mechanisms are certainly adaptive traits but not all adaptive traits are robustness mechanisms. The critical distinction is that robustness mechanisms are adaptive traits that promote system persistence by reducing variance in performance in the face of entropic factors. On the other hand, for a trait to be considered adaptive it need only meet the requirement that over some timescale it increases mean performance in some well defined functional context.

6. Robustness in Social Systems

The question of how and why social systems persist despite being composed of individuals that die and possess partially aligned interests, has been addressed qualitatively by structuralist anthropologists and to some extent sociologists. These would include Emile Durkheim and Levi Strauss pursuing questions on the functional value of the division of labor and ubiquitous features of culture such as the incest taboo, or Ruth Benedict addressing the role of the individual in the production of culture and the role of culture as a

constraint on individual behavior. However, as anthropology devolved into post-modernism, the issue of robustness was largely abandoned although small communities of anthropologists and archeologists interested in the rise and decline of civilization continued to pursue related concepts [**?**, **?**].

With the rise of sociobiology, behavioral ecology, and eventually evolutionary psychology, beginning in the late 1960's and early 1970's, attention turned again to social behavior, this time with biologists and cognitive scientists with a training in evolutionary theory taking the lead. Rather than focusing on the mechanics of social organization, the focus has largely been "economic" emphasizing the cost and benefits that support one behavioral strategy over another. The overwhelming goal in this research trajectory has been to give a compelling account for when selection favors group formation and cooperative interactions in the framework of game theory.

Unlike evolutionary theory which has developed in parallel with mechanistic biology –that is with an empirical emphasis on the genetic foundations of population structure – the theory of social evolution has suffered from a relative absence of mechanical understanding of how societies arise. And both biological and social evolution have tended to underemphasize ontogenetic regulatory dynamics, such as metabolic, physiological and developmental processes.

Hence despite a growing body of research on the cognitive and neural bases for the cooperative strategies studied in game theoretic models and experimental games, there has been relatively little quantitative work on how collective properties, social structure, and other aggregate features of groups are produced collectively through the interaction of individuals using different strategies and decision-making rules, or on how these aggregate properties persist in the face of high level of uncertainty and noise. Important exceptions to this include research on the rules underpinning fish schools and bird flocking behavior[**?**], studies of colony formation in social insects[**?**], the emergence of social power structure, collective conflict dynamics, and the implications of robustness mechanisms for social niche construction in primate societies[**?**, **?**].

Our goals in the final sections of this chapter are to discuss how robustness can be studied in social systems and what we can learn about social evolution by considering how societies and the decision-making rules underpinning them persist despite frequent perturbations.

6.1. **Collective properties: The analog to the phenotype.** If we are to retain the formal definition of robustness we provided earlier in the MS –*system property S is robust if it remains invariant under a perturbation P that removes, disables, or otherwise impairs the performance of adaptive component C.* – we have to ask what we mean by system property S in the social case.

In the biological domain, system property S is typically some aspect of the phenotype such as or the number of spines or the number of cells in

the endomesoderm of a sea urchin. Or it can be a highly conserved trait across lineages like the ability to metabolize a novel substrate. Each of these examples can be thought of as a collective feature of a system in so far as they arise from the interaction of genes and proteins through a complicated regulatory process. The dynamics can be summarized in terms of causal networks or regulatory circuits[**?**].

In the social case, the analogs to these phenotypic properties are aggregate properties that arise out of the collective behavior of individuals in regulatory social networks. Examples range from the statistically simple to the complex, including the trajectory and / or volume of schools of fish, the distribution of fight sizes in a primate society, the degree of assortativity or reciprocity characterizing social interactions, and the distribution of power in animal societies[**?**, **?**]. Whereas some of these aggregate properties arise from simple rules, others require integration over complex dynamics at the microscopic level.

Perturbations relevant at the social level include chronic conflict, and environmental, physiological, or cognitive deficits that result in anomalous behavior or the failure of individuals to perform critical social roles. Other common perturbations are acts of predation that remove key individuals. When roles and functions are institutionalized or supported by "pool" mechanisms, perturbations can take the form of a sudden lack of funds required to support the function. An example would be a market crash that leads towns to disband their police departments. The key is to show that these perturbations remove individuals or subgroups that have critical social roles with respect to some system function, or directly impair the functions themselves. Hence, as with the study of robustness in biological systems, we must demonstrate that the component affected by the perturbation makes a critical contribution to a measure of system performance and we must demonstrate that the perturbation is non-trival and of a relevant scale.

To make this concrete, consider the following example. In pigtailed macaque societies, third-party policing is a primary conflict management strategy. This social function is performed by a small subset of individuals at low cost (quantified in terms of aggression received by them from conflict participants in response to their intervention). These individuals pay low cost because they sit in the tail of a heavy tailed power distribution, meaning they are perceived by group members as disproportionately capable of using force successfully during fights and so are rarely challenged when intervening. Hence the high variance power distribution supports the policing mechanism. By temporarily disabling the policing mechanism using "behavioral knockout" methods, it has been shown that the policing mechanism is a robustness mechanism. The policing mechanism checks the escalation of conflict, thereby allowing individuals to interact at relatively low cost and so build more connected and integrated social networks, which in turn facilitates the extraction of social resources (alliance partners, information about location of food and sleeping sites, etc.). Hence policing,

a robustness mechanism, changes the economics of behavior by modulating the cost of interactions and allowing individuals to invest more in beneficial, socio-positive interactions.

To summarize, the question being asked in the policing study was, "is the policing mechanism a robustness mechanism?". The system property, S was the degree of integration in social networks and this property was quantified using a variety of biologically meaningful social network metrics. The perturbation was chronic conflict, and the component or feature of the system being disrupted by the perturbation was social niche construction, or the ability of individuals to build edges in their social networks.

6.2. **Robustness at the Behavioral Scale.** In the policing study the system feature being studied –degree of network integration– is a collective feature defined over nodes in the network. It is also possible to ask about the robustness of the underlying behavioral strategies themselves. These are the decision-making rules individuals use when, for example, deciding to fight or to help an alliance partner, or invest in an edge in a social network.

In the case of robustness mechanisms, such as third-party policing which buffer against conflict, the "system" is the social system. The "system" in the case of decision-making rules or strategies is the individual's cognitive system which encompasses its neuro-physiological and behavioral activity. The strategy is the output of the cognitive system and the system property S is some statistical feature of the strategy. The perturbation could include environmental factors –either ecological or social –or factors internal to the organism that increase, for example, entropy in neural firing patterns. The component disrupted by the perturbation might be the firing pattern of a single neuron, the firing pattern defined over a network of neurons, the degree of blood flow to a brain region, or any other neural-physiological behavior that makes a functional contribution to the cognitive / behavioral output of interest.

6.3. **Methods of Experiment and Analysis.** Studies that explicitly aim to determine the robustness of decision-making rules are uncommon, but key insights and methods from biological and social robustness studies can be generalized into the decision making area.

Studies of robustness at the societal scale will require creative and novel approaches, as integrated social systems are hard to manipulate systematically, and human social systems present important ethical limitations. When such experiments can be conducted, they can be hard to replicate as similar factors have to be controlled to make comparison worthwhile. Nonetheless, with caution, it is possible to proceed, and some progress has already been made[**?**].

6.3.1. *Behavioral knockout.* When the robustness mechanism is localized in a single or small subset of individuals it is possible to conduct "behavioral knockout" experiments, as illustrated by the policing study. Behavioral

knockout studies are methodologically similar to those conducted at the genetic and molecular levels (see for an example, Flack references in Recommended Reading). In these studies, the focal mechanism is disabled by temporarily removing individuals responsible for it. There are obviously many potentially confounding issues, such as how to ensure that only a single function disabled when individuals typically perform multiple roles. But these can sometimes be resolved through careful and creative experimental design.

Similarly it is possible to ask whether a system is robust to some specified perturbation if the components (individuals or behavioral mechanisms) effected by the perturbation can be determined. Typically obtaining rigorous and trustworthy results will require the use of animal society model systems. These involve animals that live in large, captive social groups, allowing highly resolved time-series and network data to be collected during the experiment as well as pre and post the experiment.

Down the road it might also be possible to find creative ways to conduct knockout studies on social systems with distributed robustness mechanisms. This has also been a challenge in the genetics and molecular cases, but recent studies designed to disrupt the distributed microtubule networks of cells suggest that this is also achievable.

6.3.2. *Natural experiments.* When a controlled experiment is not possible, but time-series and network data are available before and after a documented "natural perturbation" , it is possible with the help of appropriate null models to draw preliminary conclusions about whether the system is robust to the observed event.

6.3.3. *Simulated experiments.* Finally, if time-series data are available on social interaction patterns, the decision-making rules individuals use can be extracted from data using a variety of non-parametric and machine learning approaches such as those used in Inductive Game Theory (see DeDeo reference in Recommended Reading). These decision-making rules can then be used to build an empirically grounded virtual environment that allows for simulated and controlled virtual experiments.

7. Summary

The study of robustness represents an attempt to explain the continued function of complex systems in the face of frequent errors and perturbations. Historically, questions related to robustness have emphasized either functional frameworks (economic and game theoretic) or mechanical frameworks (developmental and engineering). The challenge is to unite these so as to explain the evolution of robust mechanisms. A critical element of this objective is to formalize the logic of robust systems and present systematic experimental approaches to their analysis. With the appropriate experimental logic we can go about isolating the many mechanisms that yield

robust functions, such as error correction, conflict management, developmental canalization and niche construction. Decision making rules represent an exemplary study system for robustness, as errors in individual judgement will need to be compensated through mechanisms of consensus formation at the collective level. In highly coupled systems that are intractable to experiment, simulated experiments, building upon frameworks such as agent-based models and inductive game theory, will prove to be important.

8. Recommended Reading

Ancel, LW. and Fontana, W. 2000. Plasticity, evolvability, and modularity in RNA. Journal of Experimental Zoology, 288: 242-283.

Ay, N., Flack, J.C. and Krakauer, D.C. Robustness and complexity in co-constructed multimodal signaling networks. Phil Trans, Series B. 382: 441-447. *The above paper presents a formal definition of robustness.*

Ay, N. and Krakauer, D.C. 2007. Geometric robustness theory and biological networks. 125: 93-121. *Mathematics of formal definition of robustness more fully developed in this paper.*

DeDeo, S., Krakauer, D.C. and Flack, J.C. 2009. Inductive game theory and the dynamics of animal conflict. PLoS Computational Biology, 6(5): e1000782. *Paper presents methods for extracting decision making rules from timeseries data.*

Felix, M.A. and Wagner, A. 2008. Rogustness and evolution: Concepts, insights and challenges from a developmental model system. Heredity. 100: 132-140.

Flack, J.C., Girvan, M.C., de Waal, F.B.M. and Krakauer, D.C. 2-006. Policing stabilizes the construction of social niches in primates. Nature, 439: 426-429.

Flack, J.C., Krakauer, D.C., and de Waal, F.B.M. 2005. Robustness mechanisms in primate societies: A perturbation study. Proceedings of the Royal Society, Series B. 272: 1091-1099.

Krakauer, D. 2005. Robustness in biological systems: A provisional taxonomy. In E. Jen (ed) Robust design: A repertoire for biology, ecology and engineering. Oxford University Press.

Jen, E. 2002. Stable or robust: What's the difference? In E. Jen (ed) Robust Design: A repertoire of biological, ecological, and engineering case studies. Oxford University Press.

Jen, E. 2002. Robust Design: A Repertoire of Biological, Ecological, and Engineering Case Studies, Oxford University Press.

Lansing, S. 2006. Perfect Order: Recognizing Complexity in Bali. Princeton University Press.

Wagner, A. 2007. Robustness and Evolvability in Living Systems, Princeton University Press. *Strong molecular bias, and focus is on redundancy versus neutrality.*

Waddington, C.H. 1942. Canalization of development and the inheritance of acquired characters. Nature 150: 563-565.