

# Natural selection and the regulation of defenses A signal detection analysis of the smoke detector principle

Randolph M. Nesse\*

*The University of Michigan, Room 5261, 426 Thompson Street, Ann Arbor, MI 48104, United States*

Initial receipt 3 November 2003; final revision received 6 August 2004

---

## Abstract

Many of the body's adaptive responses, such as pain, fever, and fear, are defenses that remain latent until they are aroused by cues that indicate the presence of a threat. Natural selection should shape regulation mechanisms that express defenses only in situations where their benefits exceed their costs, but defenses are often expressed in situations where they seem unnecessary, with much resulting useless suffering. An explanation emerges from a signal detection analysis of the costs and benefits that shaped defense regulation mechanisms. Quantitative modeling of optimal regulation for all-or-none defenses and for continuously variable defenses leads to several conclusions. First, an optimal system for regulating inexpensive all-or-none defenses against the uncertain presence of large threats will express many false alarms. Second, the optimum level of expression for graded defenses is not at the point where the costs of the defense and the danger are equal, but is instead where the marginal cost of additional defense exceeds the marginal benefit. Third, in the face of uncertainty and skewed payoff functions, the optimal response threshold may not be the point with the lowest cost. Finally, repeated exposures to certain kinds of danger may adaptively lower response thresholds, making systems vulnerable to runaway positive feedback. While we await quantitative data that can refine such models, a general theoretical perspective on the evolution of defense regulation can help to guide research and assist clinical decision making.

© 2005 Elsevier Inc. All rights reserved.

*Keywords:* Defenses; Natural selection; Evolution; Signal detection theory; Pharmacology; False alarms; Control theory; Cybernetics; Systems theory

---

\* Tel.: +1 734 764 6593; fax: +1 734 647 3652.

*E-mail address:* nesse@umich.edu.

## 1. Introduction

Understanding how natural selection shaped the systems that regulate defense responses is obviously important because we want to avoid and alleviate aversive experiences such as pain, nausea, and anxiety. This importance grows as we must decide how to use increasingly powerful drugs that block bodily defenses. Negative emotions such as fear, anger, and depression are also defenses that are ever more susceptible to pharmacological manipulation. Understanding the systems that regulate defenses is a crucial foundation for clinical decision making.

The systems that regulate defenses are distinctive subtypes of the general control systems that make life possible. Control systems shaped by natural selection use negative feedback to maintain stability in the multitude of components and processes at various levels of organization that constitute an organism. From gene expression to biochemical pathways, development, physiological variables, behavior and social relationships, organisms are dynamic systems whose stability is maintained by control mechanisms that regulate their components.

The understanding of control systems is a remarkably recent development. Most introductory biology books mention [Bernard's \(1872\)](#) notion of the “milieu intérieur” and then explain “homeostasis” ([Cannon, 1929](#)) using examples such as the regulation of respiration and body temperature. In simple homeostasis, the deviation of a controlled variable from its set point initiates responses that return the variable back towards the set point in a process of stabilizing negative feedback. For instance, increased body temperature initiates sweating, flushing, and decreased activity. When these responses lower the temperature to the set point, the homeostatic system turns them off, and stability is thus maintained.

This simple fundamental principle has provided the foundation for an increasingly sophisticated perspective on living systems. The classic book *What is Life?* by [Schrödinger \(1944\)](#) emphasized how organisms avoid entropy by using energy to create and maintain order. Shortly thereafter, [Wiener \(1948\)](#) expanded the basic principle of feedback control into cybernetics, and [Shannon and Weaver \(1949\)](#) codified information theory. Grand syntheses followed for systems theory, in general ([von Bertalanffy, 1969](#)), and biology, in particular ([Miller, 1978](#)). In an elaboration of cybernetics, the “perceptual control theory” of [Powers \(1973\)](#) emphasized that the behavior of organisms is not controlled directly by external variables but by the regulation of internal perceptions that convey data from sensory receptors. Because these perceptions are the focus of regulation, anything that distorts them will disrupt the system. For instance, after taking a drug that makes you feel subjectively hot, you might remove your sweater even in a cool room. In yet another rapidly advancing area, computer modeling is revealing how the behavior of myriad individual agents can give rise to complex systems that exhibit patterns of order that cannot be predicted from their individual control systems ([Holland, 1992](#)).

Modern systems theory formalizes the notion of feedback regulation in standardized control system diagrams and nomenclature (see [Fig. 1](#)). Such descriptions are applied routinely in physiology, biochemistry, and developmental biology. For behavior, they have been the basis for much work in learning and behavioral ecology ([Krebs & Davies, 1997](#)).

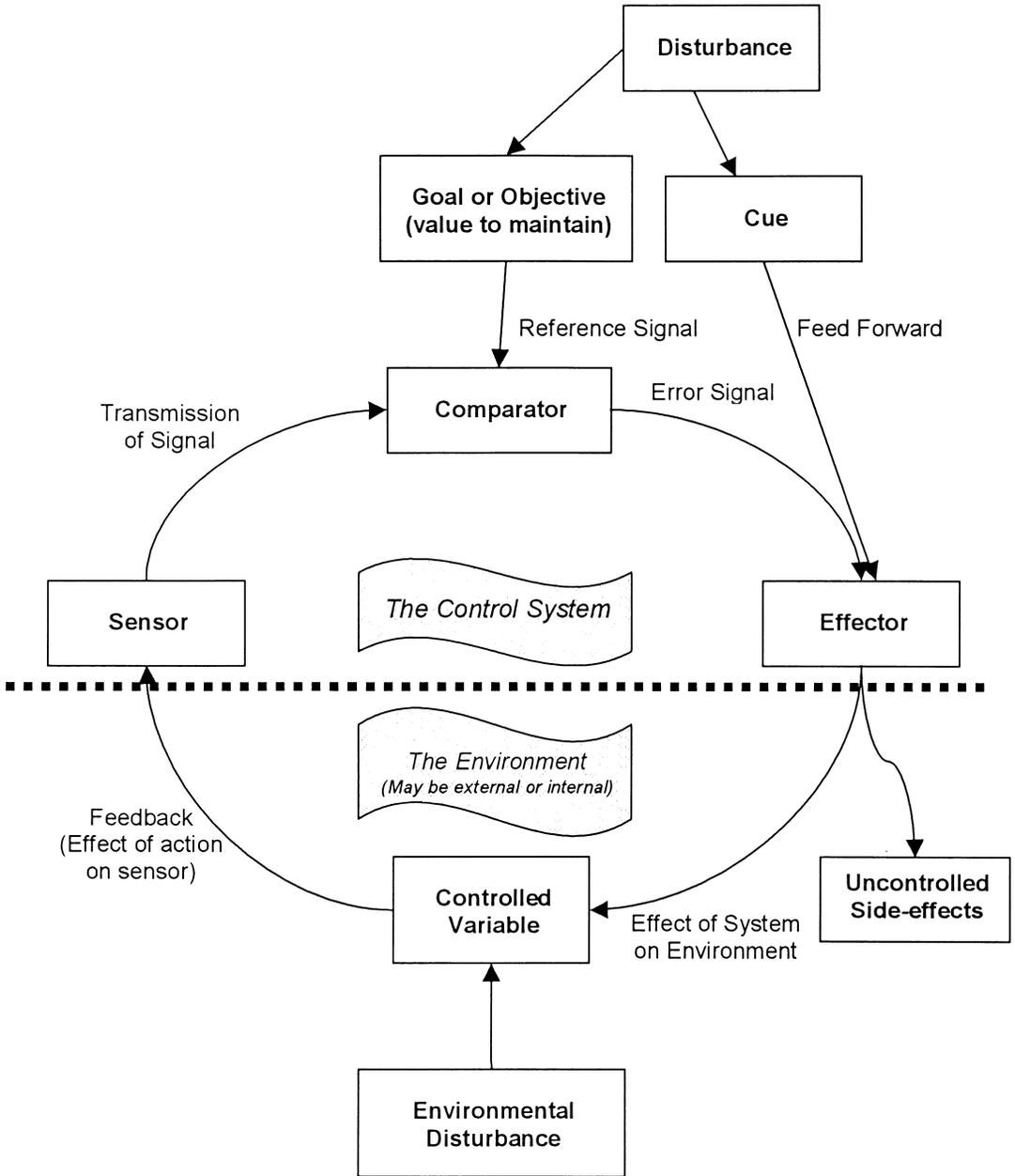


Fig. 1. Control system components and nomenclature (after Cziko, 2000).

They are just now being applied systematically to the problem of how natural selection shaped the mechanisms that regulate defenses.

Maintaining homeostasis is, however, only the most fundamental regulation task for life. To survive and compete, organisms must monitor outer and inner environments to not only

maintain stability but also to be able to respond effectively to changing circumstances. For instance, as the sun increases the temperature of desert animals, they move to cooler locations, a standard example of homeostasis. However, selection also shapes systems that use other cues that allow an organism to respond before the controlled variable changes. An animal exposed to sunlight may move to the shade even before its body temperature rises. Such “feed-forward” systems give an advantage by initiating adaptive behavior more quickly. For instance, sweet tastes initiate a neural reflex that releases insulin from the pancreas even before blood glucose levels rise, thus speeding sugar utilization and minimizing glucose fluctuations. If the taste is from an artificial sweetener, however, this mechanism will lower glucose levels inappropriately. Far more complex anticipatory capacities have been observed, such as butterflies that use the anticipated heat of each day to forage in the morning hours for the amount of water that they will likely need during that day.

The value of responses that anticipate future events helps to explain why natural selection shaped classical and operant conditioning. In operant conditioning, behaviors followed by a reward become more frequent, and those followed by a punishment become less frequent, with obvious benefits. A cue that indicates the availability of the reward can become a secondary reinforcer whose presence elicits the behavior sooner, or in complex sequences that are necessary to get the reward.

In classical conditioning, an unconditioned reflex response gradually comes to be expressed in response to previously neutral cues that have been paired repeatedly with an unconditioned reflex response. In Pavlov’s classic demonstration, dogs that repeatedly heard a bell just before being fed salivated when they heard the bell, even in the absence of food. Classical conditioning gives a selective advantage by expressing useful responses sooner. Animals that salivated only after they tasted food must have been at a selective disadvantage, although the ubiquity of classical conditioning in many orders of organisms suggests its early origins and diverse utility.

Many defensive responses are regulated, in part, by classical conditioning. For instance, anxiety can be conditioned readily by cues that precede a painful stimulus; some people have a conditioned anxiety response to sitting in a classroom, and others become nauseated by the smell of the peppermint schnapps that they once consumed in excess. Such conditioning is more rapid for certain cues, such as snakes and spiders, for which anticipatory defensive arousal was likely especially useful in our evolutionary past (Mineka, Keir, & Price, 1980; Öhman & Mineka, 2001).

Control systems do not necessarily return a system to the previous level; different set points are adaptive in different situations. In fever, for example, a higher set point for body temperature facilitates fighting infection (Kluger, 1979). To describe the many regulation mechanisms that adjust set points adaptively, Mrosovsky (1990) coined the word “rheostasis.”

Not all defenses are regulated. Some, such as the skin, are always present (although callus formation adjusts skin thickness to the local degree of protection needed). Other defenses are aroused or adjusted by early life environmental cues that permanently alter developmental trajectories. In a well-studied example, *Daphnia* (water fleas) grow sharp protective helmets only if they are exposed early in development to chemicals present when predators are in the vicinity (Tollrian & Harvell, 1999). Recent work shows that the chemical is not from the

invertebrate predator itself but consists of products released when it digests *Daphnia* (Stabell, Ogebebo, & Primicerio, 2003). In bacteria, heat stress proteins and the pathways that regulate them have been studied extensively. In many higher animals, adult body size is influenced by fetal levels of available nutrition. In humans, early exposure to cues that indicate scarce resources initiates a whole suite of possibly protective physiological changes such as weight gain and other components of “Syndrome X,” including hypertension, diabetes, and atherosclerosis (Neel et al., 1998). In many species, early exposure to stress, even in utero, can lower the stress response threshold and result in stress-associated diseases later in life (Sapolsky, 1992). In the behavioral realm, early loss of the father has been suggested to induce different female mating strategies (Belsky, Steinberg, & Draper, 1991).

Many defenses, such as pain, vomiting, fever, cough, and the fight–flight response, respond on a much shorter time scale. In the absence of a relevant stimulus, there is no way to know that some of these defenses even exist. You would not guess that mammals shiver, but reducing body temperature reliably elicits this useful defense. You might guess that organisms would have a way to clear foreign matter from the digestive and respiratory systems, but the exact mechanisms of vomiting and coughing would be hard to predict. The existence of many such systems that are useful only in certain situations may help explain why many lines of genetically manipulated mice appear to be normal although they are completely missing certain knock-out genes.

Some of the body’s protective responses are general. For instance, Selye’s “general adaptation response,” later called stress, adjusts a variety of physiological parameters in ways that facilitate coping with threats or opportunities that require action (Nesse & Young, 2000). Interestingly, the hypothalamic pituitary adrenal responses that are often taken as synonymous with stress may exist, in part, to protect the body from other aspects of emergency responses (Munck & Narayfejestoth, 1994). Natural selection has also shaped more specific defensive responses to cope with distinct kinds of threat. These are superb examples of domain-specific mechanisms (Cosmides & Tooby, 1987). Threats of injury or infection arouse specific physiological defenses such as pain and inflammation. A threat to the welfare of kin arouses worry. Threats to mating fidelity, friendships, or social status arouse yet more complex special states of emotion and motivation, such as jealousy and anger. Table 1 summarizes some defenses and the situations that arouse them.

While not the main focus here, it is worth noting that natural selection seems to have partially differentiated certain defenses from more generic precursors so that organisms can cope more effectively with particular kinds of challenges. This suggests a possible solution to debates about whether the mind is composed of discrete modules. It also obviates arguments about whether the subtypes of a response like anxiety are all separate or all fundamentally the same—they are neither, they are partially differentiated from a common precursor (Marks & Nesse, 1994). Advances in molecular biology may well help to confirm this if variations in receptors and other regulators do indeed map onto a proposed phylogeny of different subtypes of anxiety.

Defenses against severe threats that are hard to detect reliably pose special regulation challenges that are the focus of this article. Because such defenses have costs, they tend to be held invisibly in reserve until they are needed. Because they tend to be much less expensive than are the dangers they defend against, false alarms are inexpensive compared with the

Table 1  
 Defensive responses and situations in which they give an advantage

Events and situations in which defenses are useful	Defensive responses that improve ability to cope with these events and situations
Physical trauma	Pain, startle, withdrawal reflexes, anxiety, learned avoidance
Environmental extremes	Protective behavior, seeking shelter, shivering, sweating
Deficiencies	Suffocation, thirst, hunger, etc., and fear of these responses
Infection	Disgust, expulsive defenses including, spitting, vomiting, diarrhea, coughing, sneezing, rhinorhea, immune responses, inducible immune responses, immune memory, inflammation, fever
Predator attack	Anxiety, fight flight response, avoidance learning
Conspecific attack	Suspiciousness, threat and submission communications, alliances, avoidance, gossip
Loss of territory or material resources	Worry, monitoring resources and protecting them from theft or loss, suspicion
Death or harm to mate or kin	Worry about any indication of possible threat to kin, and attempts to prevent harm or loss, bereavement
Loss of mate's fidelity	Jealousy
Loss of allies	Anger, anxiety, guilt, other social emotions
Loss of status	Social fear, embarrassment, anger

possibly dire consequences of failure to express the defense if the threat is present. Because the cues to such dangers tend to be only probabilistic indicators of the presence of the threat, an effective regulation system needs to assess the likelihood that a stimulus does or does not indicate the presence of the danger.

As already mentioned, there are practical reasons why the regulation of defense responses is worth a detailed investigation. Most of life's suffering arises from defense responses. Pain, aversive in its very essence, is the exemplar. People born without the capacity for pain die in early adulthood, tragically demonstrating the utility of a capacity for suffering (Sternbach, 1963). Cough, nausea, vomiting, diarrhea, and anxiety are all also unpleasant, almost certainly because their aversiveness motivates escape and future avoidance of situations that could have caused the response (Williams & Nesse, 1991). Most of the problems for which people seek medical help are also defenses. In fact, one of the most common and useful interventions offered by general physicians is prescribing drugs that block defenses and their associated suffering. Life is vastly better now that pain, vomiting, cough, diarrhea, fever, anxiety, and other defensive responses can usually be controlled.

The second reason is intimately related to the first. Optimal clinical practice—whether in general medicine, psychotherapy, or public policy—requires understanding how defenses are regulated as a basis for deciding when it is safe to block them, and when it is not. The relative safety of blocking defenses often leads to the “clinician's illusion” of viewing defenses as the problem. While most clinicians realize that cough is useful, fewer recognize fever as a defense, and many have difficulty seeing any utility in anxiety and sadness. Such ignorance about the utility of defenses is usually of little consequence, but on occasion, blocking a defense can be harmful or even fatal, such as when codeine suppresses cough, and pneumonia

results. A more rigorous framework for assessing when defenses can and cannot be safely blocked is an important missing foundation for clinical practice.

A comprehensive treatment would be based on a detailed analysis of the specific mechanisms that regulate each defensive system. We have little data to support such an understanding. However, the “phenotypic gambit” of treating behavioral tendencies as if they were components directly shaped by selection has been useful in behavioral ecology and offers a way to proceed (Alexander, 1975; Grafen, 1984). It is especially apt in this case because physiologists have long pursued similar strategies in their assumption that the body’s internal regulation mechanisms are somewhere near optimal (Schmidt-Nielsen, 1990). Often, such approaches are conducted under the rubric of homeostasis, rheostasis, symmorphosis (Weibel, Taylor, & Bolis, 1998) or optimization theory (Alexander, 1996). But there is a surprising dearth of systematic applications of optimization to the problem of defense regulation.

This article extends a control systems approach to the problem of predicting how optimal defense regulation systems should function. Several somewhat different situations require different models. The simplest model predicts when an all-or-none defense of known cost (such as vomiting at an average cost of, say, 500 kcal) should be expressed given how much it is expected to reduce harm. That model is expanded to consider responses to uncertain cues. A related model covers responses such as fever that are expressed in continuous gradations. A final model considers expectations about how metaregulation mechanisms should adjust thresholds for expression of particular defenses as a function of prior experience and knowledge about the current environment.

## 2. All-or-none defenses

Many defensive responses are expressed either fully or not at all. Vomiting and panic attacks are examples. These offer a good starting point because, for the purposes of the model, the cost of such defenses can be considered fixed. Expressing such responses will give a net payoff when the cost of the defense is less than the benefit (amount of harm reduction). If the defense offers perfect protection and the potential harm is 100% certain, then an optimal mechanism will express the defense whenever the cost of the defense (CD) is less than the cost of the harm (CH) that will occur when the danger is present and no defense is expressed. In a more typical situation, CH will not be eliminated but only reduced, thus, the defense should be expressed whenever  $CD < (CH \text{ w/o defense}) - (CH \text{ w/ defense})$ . For simplicity, CH will be treated here as equivalent to the reduction in harm. When mapped onto a traditional decision table (Table 2), CD is the cost of a false alarm, while CH is the cost of a missed response.

Table 2

The costs of responding versus not responding as a function of whether the harm is present or absent

	Harm present (signal plus noise)	Harm absent (noise only)
Response	Hit (true positive) cost = CD	False alarm (false positive) cost = CD
No response	Miss (false negative) cost = CH	Correct rejection (true negative) cost = 0

The difficulty comes on the detection and decision side of the system. What concentration of toxin in the stomach should set off the vomiting reflex? Most food will contain some toxin molecules, thus, too sensitive a system will result in vomiting after each meal. But an insensitive system runs the risk of exposing the individual to serious or fatal infection. The optimum is somewhere in between.

A very sophisticated system would monitor not only the momentary concentration but also the direction and rate of change in toxin concentration. Just such a system responds to the rate of change in cortisol to regulate the hypothalamic–pituitary axis (Young & Vasquez, 1996). Having acknowledged this possibility, with a note that experimental tests to look for such sophisticated control mechanisms in other bodily systems would be worthwhile, we will set it aside for now to keep the model simple.

To decide whether to flee in response to a stimulus, an organism needs to assess whether it indicates a danger. For instance, an impala hears a noise from behind a bush that is too loud to have been made by a mouse, but not so loud that it was definitely made by a lion. From prior experience with such cues and knowledge of the relative prevalence of noises at that volume made by predators versus other animals (the likelihood ratio), the impala assesses the probability that the noise came from a potential source of real harm  $p(H)$ . An optimal system uses this information to express the defense whenever  $CD < p(H) \times CH$  or, equivalently, whenever  $p(H) > CD/CH$ . Obviously, the organism does not actually calculate probabilities, it needs only to have a system that assesses the relative likelihood that the danger is present.

As an example, assume that the fixed cost of expressing a panic flight response is 300 kcal, and the damage that results from an unprotected attack by a predator is, on average, 300,000 kcal. The ratio of CH to CD is 1000:1; hence, if the expression of the defense offers complete protection, an optimal regulation mechanism would express the flight response whenever the magnitude of the cue indicates that  $p(H) > 0.001$ . The implications are somewhat surprising. Out of 1000 episodes of flight at this response threshold, on the average, 999 will be false alarms and only 1 will be flight from an actual predator. The false alarm flight responses each are unnecessary and wasteful in the specific instance but are nonetheless completely normal products of an optimal system. This preponderance of normal false alarms in an optimal defense regulation system has been called “the smoke detector principle” (Nesse & Williams, 1994). This is a biological version of Pascal’s Wager.

One common method for analyzing decision systems is to graph the receiver operating characteristic curve that shows the percentage of false alarms (responses to stimuli that are not real signals of danger) plotted on the  $x$  axis and the percentage of hits (correct responses to signals that indicate danger) plotted on the  $y$  axis (Green & Swets, 1966). The curves in Fig. 2B represent systems with different sensitivity. If the means and variances of the two distributions are essentially the same, stimuli arising when the danger is present (signal plus noise) cannot be distinguished from other (noise) stimuli, hence, all points fall on the diagonal where the rate of false alarms equals the rate of hits. In a more sensitive system, however, the mean of values from signals is different from the mean for the values that arise from noise only (as in Fig. 2A). This difference offers information for distinguishing when the danger is present or not. As the means for the two distributions in Fig. 2A become more separated, the

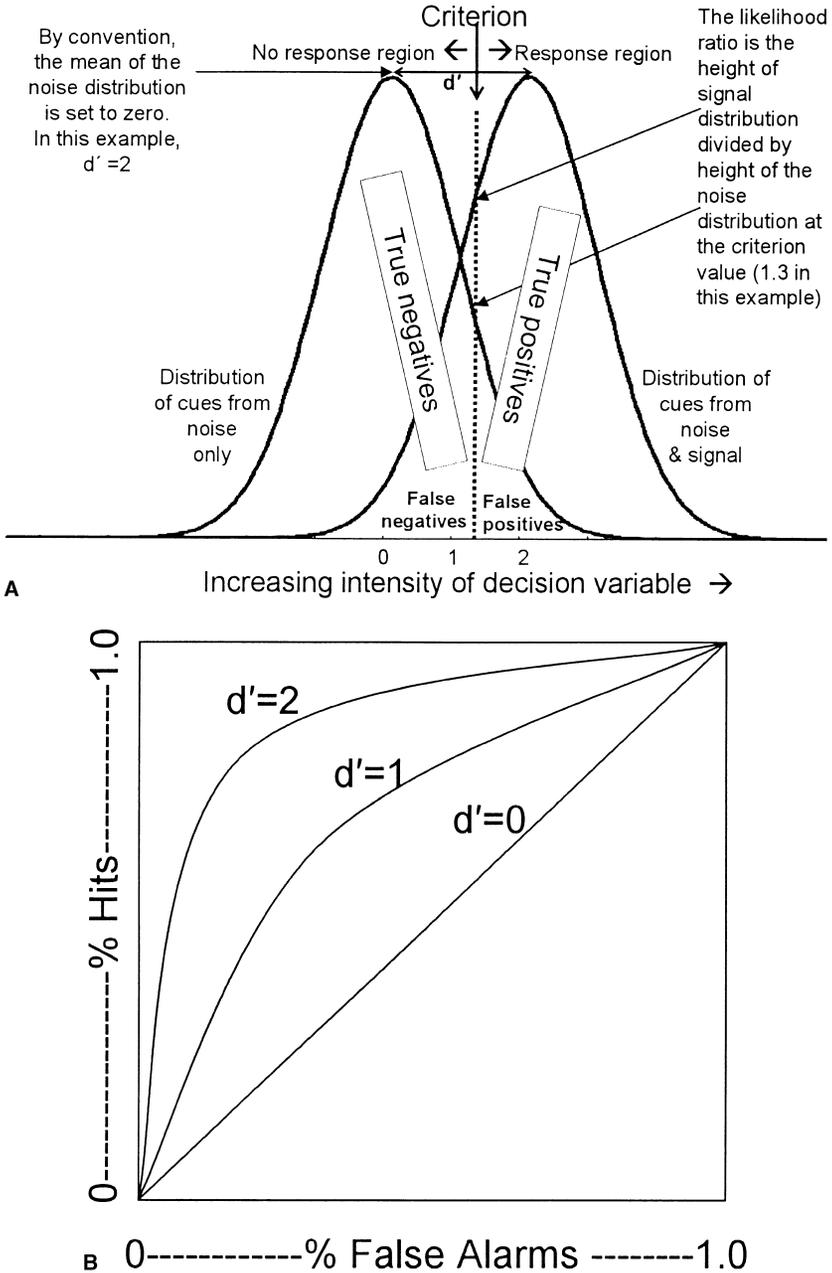


Fig. 2. (A) The four categories of response from a criterion that partitions overlapping distributions of noise-only cues and noise plus signal cues. (B) Receiver operating characteristics for  $d' = 0$ ,  $d' = 1$ , and  $d' = 2$ .

corresponding curves in Fig. 2B bow out further towards the upper left corner, and the proportion of hits increases relative to the proportion of false alarms. Each sensitivity line defines the proportion of hits depending on the proportion of false alarms deemed acceptable.

The optimal response criterion,  $\beta$ , is the threshold for the decision variable above which expressing a response gives a net gain. The decision variable might be the intensity of a noise that could indicate either a dangerous predator or a harmless rodent, or it could be the concentration of toxin molecules in the stomach. If the base rates of signal and noise are equal, and the costs of errors are equal, then the optimal criterion is at the intersection of the distributions where the likelihood ratio is 1.0. Things are rarely so simple, however. Noise may be much more common than situations with real dangers, and costs differ for different mistakes.

The likelihood ratio  $p(x|s)/p(x|n)$  is the density at stimulus intensity  $x$  of the distribution for signal divided by the density of the distribution with noise only. Utility is maximized if the organism responds whenever  $\frac{p(x|s)}{p(x|n)} \geq \frac{p(n)}{p(s)} \times \frac{v(\text{correct rejection}) + v(\text{false alarm})}{v(\text{hit}) + v(\text{miss})}$ . This standard formula includes the baseline probabilities that a stimulus will be noise  $p(n)$  or signal  $p(s)$ , and the value  $v$  for each outcome. In our simplified example, there is no cost for a correct rejection, the values of a hit and a false alarm are both CR, and the value of a miss is CH, hence,  $\beta = \frac{p(n)}{p(s)} \times \frac{\text{CR}}{\text{CR} + \text{CH}}$ . This is consistent with intuition; the defense should be expressed less readily (at a higher criterion) as noise becomes more prevalent than signal, as the costs of false alarms increase, and as the cost of a missed response decreases.

To take a simple illustration, if cues arise equally often from danger and noise ( $p(s) = p(n) = 0.5$ ) and we assume  $\text{CD} = 1$  and  $\text{CH} = 10$ ,  $\beta$  will be  $0.5/0.5 \times 1/11 = 0.091$ . If the means of the noise and signal distributions are two standard deviations apart, that the discriminability index ( $d'$ ) equals 2.0, and we assume standard normal curves with a mean of zero for the noise distribution, this likelihood ratio would be located just to the left of the mean of the noise distribution, at a signal intensity of  $-0.2$ . Calculating from the  $z$  scores, this criterion will result in an expression of a response to 98.6% instances of actual danger but will fail to express a response in 1.4% of dangerous situations. This criterion will correctly not respond to 42.1% of noise stimuli, but will give false alarms to 57.9% of noise stimuli. Given the assumption that events from signal and from the signal plus noise distributions are equally common, 49.3% of overall events will result in hits, 0.7% in misses, 29.0% in false alarms, and 21.0% in correct lack of response. About 37% of all responses will be false alarms.

As  $\text{CH}/\text{CR}$  increases, the optimal threshold moves to the left, where false alarms are more frequent. If  $\text{CH}$  is  $1000 \times \text{CR}$ , then the optimal likelihood ratio of 0.001 corresponds to a criterion of  $-2.46$ , where a defense will be expressed 99.9996% of the time that danger is present and will fail to respond to only four signals out of a million. This comes at a price, however: 99.3% of responses to noise stimuli will be false alarms, as will 50% of total responses.

The optimum is also influenced by the baseline prevalence of signals relative to noise. If, as is often the case, stimuli from noise are more common than those from danger, the optimum shifts back towards the right. If  $\text{CH}/\text{CR}$  stays at 1000, but the danger is present for only 10% of cues, then the optimal likelihood ratio shifts from 0.001 to 0.009, where the optimal criterion is a signal intensity of  $-1.36$ . Under these assumptions, 99.96% of actual dangers would arouse a response. Responding to any value from the noise distribution above the criterion would give a false alarm 91.3% of the time. Because such a high proportion of stimuli is from noise only, 89% of all responses will be false alarms.

Improving the system's sensitivity improves things only a little. If all else is held the same as in the previous example, but  $d'$  increases from 2 to 4, then the optimal criterion and the false alarm rate to cues from the noise distribution stay the same, but the percent of dangers that fail to elicit a response moves from 0.04% to 0.07%, and the overall false alarm rate is reduced to 65% of all responses.

I have been unable to find actual data for distributions of noise and signals that are detected by biological defense systems, but our defense alarms are not going off constantly, thus, noise and signal must usually be separable. On the other hand, considering the amount of anxiety many people experience when there is no apparent danger, one could also argue that the distributions for certain kinds of danger cues may overlap considerably with the distributions for noise. One can also conclude that some defenses, such as anxiety, must be inexpensive compared with the harms that they protect against (Rabin, 2000). The important points are the following: (1) normal systems express many false alarms, especially if the cost of the defense is low relative to the cost and prevalence of the harm, and the base rate of events from noise is high, and (2) the optimal threshold for expression of a defense depends not only on the means and standard deviations of signal and noise, but also on the relative costs of false alarms and missed responses and the relative prevalence of stimuli associated with actual danger.

Everyday examples are ubiquitous. Birds flee from backyard feeders when any shadow passes overhead. Most people tolerate many false alarms from their home smoke detectors. Wearing a seatbelt is unnecessary 999 times out of 1000, but sensible people do it. Most recently, the very remote possibility of a chemical attack has led some towns in the U.S. central plains region to advise all citizens to purchase tape to seal their windows. The prevalence of such examples may help to explain the tendency of humans and other animals to be more averse to risk and loss than seems warranted by rational choice theory (Rabin, 2000; Tversky & Kahneman, 1974).

Some people consult a doctor for any small symptom that could conceivably indicate cancer or heart disease. And doctors, confronted with such patients, may order many more tests than are in a patient's best interests to protect themselves from the rare but devastating possibility of a lawsuit. Likewise, many communities prohibit parking adjacent to fire hydrants although the chance that a fire truck will use that hydrant on a given day is less than 1 in 100,000. This is an instance where the cumulative costs may not be worth the rare benefits. A quantitative analysis might well reveal that the vast resource of all that lost parking over the decades is so many-fold larger than the benefits of easier access to fire hydrants on the rare occasions when they are needed, that communities would be safer if they added parking meters adjacent to hydrants and gave the extra revenues to the fire department.

### 3. Continuously variable responses

A different approach is needed to model defenses that can be expressed at continuously variable intensities, such as fever, rate of cough, and amount of anxiety. The costs increase across a range of increasing investments in defense. A model can be based on the reasonable

assumption that greater expression of the defense offers greater protection from the threat, with initial incremental increases usually providing more reduction in harm than identical increments at higher levels. For instance, coughing twice an hour is likely to substantially decrease the costs of a bronchial infection; coughing four times an hour would cost twice as much energy, but would not likely decrease the infection twice as much. Likewise, even small amounts of pain effectively inhibit actions that are damaging tissues; pain that is twice as disabling would not likely offer twice as much protection.

Each actual system will be somewhat different. For the purpose of illustration, the most basic model is presented here, with the simplifying assumptions that defense costs increase linearly with increasing express of defense intensity and that the costs of harm decrease exponentially with increasing levels of defense intensity. The sum of these two costs across a range of defense intensities defines a curve whose nadir is the optimum amount of defense expression. In Fig. 3A, the nadir is where  $CH=CD$ . It seems intuitively sensible that the cost of defense would be justified up to the cost of the harm. However, the optimum solution is not necessarily where the investments are equal, but is wherever the marginal benefit of additional defense investment is no longer positive. If there is a minimum CD, or if CH decreases very quickly with increased defense, then increasing levels of defense can continue yielding net benefits even when CD is greater than CH, as in Fig. 3B. Conversely, if investments in defense offer substantial payoffs only at higher levels, the optimum may be where CD is much less than CH. Such nonintuitive situations may be important in understanding certain disease states in which the defense seems worse than the disease is.

#### 4. Skewed payoff curves

Yet, another factor complicates the picture. Because information is limited and prediction is imperfect, the level of defense expressed will usually deviate somewhat from the optimal level. If the curve that describes net payoffs is symmetrical, then errors on either side have identical consequences. However, if the curve for net cost is much steeper to the left than to the right, as in Fig. 3A, then expressing three units less than the optimal defense will have much worse consequences than expressing three units more than the actual optimal level of defense. This will shape tendencies to express somewhat more defense expression than would otherwise be explicable.

On the other hand, high levels of some defenses expose the individual to the possibility of catastrophic costs. With fever, for instance, above a certain level, the likelihood of seizures creates a discontinuity. Graphically, the CD, at first increases gradually and then skyrockets at the point where seizures become likely. For payoffs with this pattern, errors of excessive defense expression are far more expensive than are errors of deficient expression, hence, the optimum level of defense under uncertainty will not be at the minimum net cost, but somewhat to the left. With diarrhea, for instance, more may well be more effective, but this may be limited because of the danger of death from excessive fluid loss. Behavioral examples are plentiful, such as the demonstration that small birds monitor the variability of payoffs at alternative feeding patches. They usually prefer a steady source of food, but when it decreases

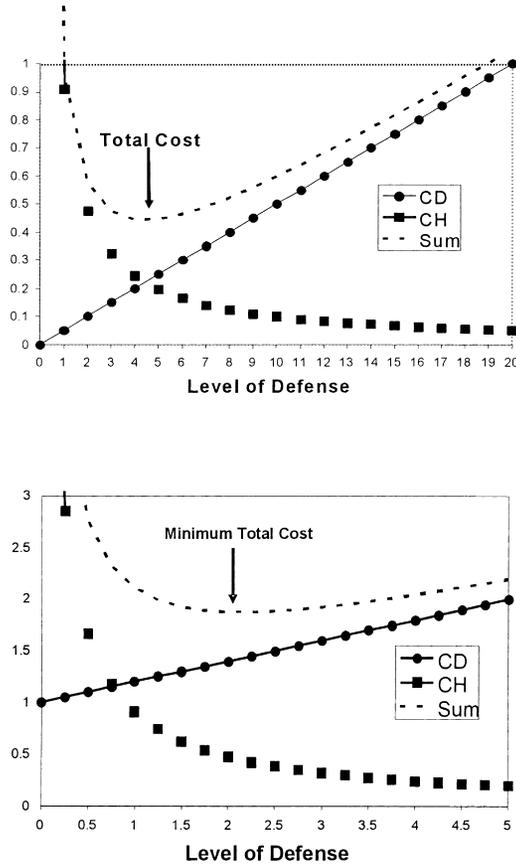


Fig. 3. Total cost at different levels of defense. (A)  $CD = LD \times 0.05$ ,  $CH = 1 / (LD + 0.1)$ . (B)  $CD = 1 + LD \times 0.05$ ,  $CH = 1 / (LD + 0.1)$ .

below the minimum needed to survive the night, they turn instead to a riskier source that offers some possibility of survival (Real & Caraco, 1986).

In humans, much research has confirmed the principle that the sex with higher variance in reproductive payoff (males) takes more risks (Daly & Wilson, 2001; Kruger & Nesse, 2004). Along related lines, Haselton and Buss (2000) have used the “error management theory” to explain the male bias to interpret almost any ambiguous cue from a woman as indicating sexual interest as a possibly adaptive cognitive distortion.

### 5. Facultative adjustment of defense thresholds

So far, we have treated the thresholds for expression of a defense as if they are fixed. But experience in a particular environment yields information about the actual CD and CH, how prevalent the harm is, how well its cues can be discriminated from noise, and how effective the organism’s defenses are. Organisms use this information to adjust the response threshold

and intensity. Such adjustments have been best studied in the capacity for sensitizing or desensitizing responses to certain cues. If repeated exposure to a danger or a cue of danger is followed by no actual harm, or only an easily avoidable threat, then the response threshold can safely be raised or the intensity of the response can be reduced. In its simplest, form this is habituation, a readily observable characteristic of many systems, from the decreased withdrawal reflex of *Aplysia* after repeated stimulation to the calm with which experienced dogs react to a stimulus that precedes shock that they can easily escape (Seligman, Maier, & Geer, 1968). This adjustment of the response criterion is also the basis for the extinction of conditioned responses. In the clinic, repeated exposure to cues that set off phobic responses is the basis for behavioral exposure therapy, the mainstay for clinical treatment of phobias (Marks & Tobena, 1990). Interestingly, brief exposures do not work, but exposures lasting many minutes to hours are reliably effective in curing simple phobias.

In other situations, however, each successive exposure indicates increasing prevalence of danger, and each narrow escape indicates a smaller margin of safety, thus making a lower response threshold advantageous. For instance, every successive attack by a predator indicates increasing danger that should lower the threshold for panic/escape responses. This is exactly what happens not only with actual exposure to danger, but also in response to episodes of panic that arise from internal dysregulation, whether at the neurochemical or cognitive level. Individuals who experience a panic attack react as if they had encountered a real danger, subsequently becoming not only more cautious but also more susceptible to additional attacks (Nesse, 1987). Panic disorder may be an example of a clinical syndrome that results from a runaway positive feedback process acting on an adjustable defense threshold. Medications that block expression of panic may thus not just block symptoms but may, by preventing attacks, help to reset the threshold to a higher level.

## 6. Novel cues and defense dysregulation

Defense regulation mechanisms were shaped in an environment very different from the one in which we live. As a result, novel cues acting on normal regulation mechanisms can result in pathology. One important example is the current obesity epidemic that results from the failure of normal body weight limiting mechanisms in our current environment. A more specific example has been suggested by Jennifer Weil (personal communication), who notes that when the heart cannot move enough blood, the resulting change in pressure stimulates increased fluid retention by the kidneys. This is exactly the opposite of what is needed. Heart failure causes fluid retention, which worsens the condition in a classic vicious circle. The system was never designed to cope with hearts that do not pump properly; it was designed to deal with the overwhelmingly likely cause of low pressure in the natural environment, dehydration. When doctors prescribe diuretics to decrease the fluid load, they are trying to compensate for the body's maladaptive defense against apparent dehydration.

Anxiety disorders offer another example. Panic attacks are essentially the same as what Cannon (1929) called the fight–flight response. Nearly every aspect of this response seems well tuned to foster effective escape from life-threatening danger of the sort posed by a

predator. There are even specific brain loci and tracts specialized for this task (Nutt & Ballenger, 2003). However, people rarely experience life-threatening encounters with carnivores in modern life, and the vast majority of panic attacks are false alarms. The difficulty is that people respond to these false alarms as if they had really encountered a predator, with subsequent lowering of the panic threshold, and frequent development of agoraphobia with its strong tendency to stay at home, if at all possible. Modern life makes staying at home possible, thus, such patients may not go out for weeks. When they finally do venture out, the least cue of danger, or no discernable cue at all, can set off another attack, illustrating how positive feedback can create a medical disorder.

Specific phobias are yet another example (Poulton & Menzies, 2002). Many snake phobics live in cities, where they almost never see a snake. In a more natural environment, experience and social learning lead to the ability to discriminate dangerous from nondangerous snakes, but for many people in a modern setting, the primal image of a snake retains its ability to arouse extreme fear for years. A few sessions of exposure therapy is usually effective in desensitizing the reaction (Marks & Tobena, 1990).

## 7. Research implications

Just as quantitative models that predict hypothetical optima have been useful in assessing anatomical structures, physiological regulation, and animal behavior (Alexander, 1975), they can also be useful in examining the regulation of defense responses. That such methods have not already been used extensively may well be because of the difficulty of getting actual quantitative data. For instance, while the costs of an episode of panic flight can be quantified, estimating the cost of not expressing a flight response is difficult. It depends not only on the predation risk, but also on the distribution of the costs from different degrees of exposure to the risk and the internal characteristics of the individual that determine how easy escape will be. Extensive data on alarm calls in prairie dogs and birds may be relevant, but I have not found experiments that systematically examine responses to different intensities of cues. Physiological defenses are more amenable to study. For instance, responses to stimuli at different intensities can be measured, such as fever responses to incremental doses of endotoxin or vomiting responses to different doses of enteric toxins. Relevant data may well be available from drug screening trials.

## 8. Clinical implications

Even in the absence of specific quantitative data, an evolutionary perspective on defense regulation is useful in the clinic. The simple conclusion that many instances of defense response are perfectly normal but not useful in the specific instance provides a theoretical framework for physicians who are trying to decide whether it is wise to use a drug to block a defense. In many cases, such as when redundant defenses offer adequate protection or when the defense is being unnecessarily elicited, this perspective offers good reasons for

intervening. The fever and discomfort accompanying colds are an example. In other situations, however, such as cough that helps to clear pulmonary secretions after surgery, this perspective will arouse appropriate cautiousness, providing a reminder that even many prior safe prescriptions of treatment do not guarantee that the defense can necessarily be safely blocked on this specific occasion.

Perhaps, the most useful impact of this perspective for clinicians should be recognition that the studies needed to make many clinical decisions about blocking defenses have not yet been performed. We know that *Shigella* infections should not be treated with antidiarrhea drugs (DuPont & Hornick, 1973), but we do not know if the same is true for ordinary traveler's diarrhea. We know that blocking fever only slightly prolongs chickenpox, but we do not know its effect on influenza. Whether nasal sprays that reduce runny nose associated with upper respiratory infections prolong illness is unknown. The needed studies are simple to carry out, and their results will be of great practical importance.

Finally, an evolutionary perspective on defenses can help to ensure that the prevalent clinician's illusion not be replaced by a simplistic belief that defenses are always useful. In fact, because so much suffering associated with defenses is unnecessary in the specific instance, it is essential that clinicians not withhold treatment out of commitment to a simplified version of an abstract theory, however correct it may be. Instead, to safely reduce human suffering, we need far more knowledge about when and how it is safe to block specific defenses.

## Acknowledgments

Thanks to Richard Alexander for his enduring inspiration and for his extensive comments on a 1982 draft of these ideas. Edward Rothman offered valuable statistical consultation. Thanks also to Nicholas Mrosovsky and Miles Kimball, who offered useful suggestions, and to the editors and two referees whose comments were very helpful.

## References

- Alexander, R. D. (1975). The search for a general theory of behavior. *Behavioral Science*, 20, 77–100.
- Alexander, R. M. (1996). *Optima for animals*. Princeton, NJ: Princeton University Press.
- Belsky, J., Steinberg, J., & Draper, P. (1991). Childhood experience, interpersonal development, and reproductive strategy: an evolutionary theory of socialization. *Child Development*, 62, 647–670.
- Bernard, C. (1872). *De la Physiologie Générale*. Paris: Hachette.
- Cannon, W. B. (1929). *Bodily changes in pain, hunger, fear, and rage. Researches into the function of emotional excitement*. New York: Harper and Row.
- Cosmides, L., & Tooby, J. (1987). From evolution to behavior: evolutionary psychology as the missing link. In J. Dupre (Ed.), *The latest on the best: essays on evolution and optimality*. Cambridge, MA: MIT press, 277–306.
- Cziko, G. (2000). *The things we do*. Cambridge, MA: MIT press.
- Daly, M., & Wilson, M. (2001). Risk-taking, intrasexual competition, and homicide. *Nebraska Symposium on Motivation*, 47, 1–36.

- DuPont, H. L., & Hornick, R. B. (1973). Adverse effect of Lomotil therapy in shigellosis. *Journal of the American Medical Association*, 226, 1525–1528.
- Grafen, A. (1984). Natural selection, kin selection, and group selection. In J. R. Krebs, & N. B. Davies (Eds.), *Behavioural ecology: an evolutionary approach* (pp. 62–84). Sunderland, MA: Sinauer.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psycho-physics*. New York: Wiley.
- Haselton, M. G., & Buss, D. M. (2000). Error management theory: a new perspective on biases in cross-sex mind reading. *Journal of Personality and Social Psychology*, 78, 81–91.
- Holland, J. H. (1992). *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. Cambridge, MA: MIT Press.
- Kluger, M. J. (Ed.) (1979). *Fever, its biology, evolution, and function*. Princeton, NJ: Princeton University Press.
- Krebs, J. R., & Davies, N. B. (1997). *Behavioral ecology: an evolutionary approach*, (4th ed). Oxford: Blackwell.
- Kruger, D., & Nesse, R. (2004). Sexual selection and the male: female mortality ratio. *Evolutionary Psychology*, 2, 66–85.
- Marks, I. M., & Nesse, R. M. (1994). Fear and fitness: an evolutionary analysis of anxiety disorders. *Ethology and Sociobiology*, 15, 247–261.
- Marks, I. M., & Tobena, A. (1990). Learning and unlearning fear: a clinical and evolutionary perspective. *Neuroscience and Biobehavioral Reviews*, 14, 365–384.
- Miller, J. G. (1978). *Living systems*. New York: McGraw-Hill.
- Mineka, S., Keir, R., & Price, V. (1980). Fear of snakes in wild- and laboratory-reared rhesus monkeys (*Macaca mulatta*). *Animal Learning and Behavior*, 8, 653–663.
- Mrosovsky, N. (1990). *Rheostasis*. New York: Oxford University Press.
- Munck, A., & Narayfejestoth, A. (1994). Glucocorticoids and stress—Permissive and suppressive actions. *Annals of the New York Academy of Sciences*, 746, 115–130.
- Neel, J., Julius, S., Weder, A., Yamada, M., Kardia, S., & Haviland, M. (1998). Syndrome X: is it for real? *Genetic Epidemiology*, 15, 19–32.
- Nesse, R. M. (1987). An evolutionary perspective on panic disorder and agoraphobia. *Ethology and Sociobiology*, 8, 73S–83S.
- Nesse, R. M., & Williams, G. C. (1994). *Why we get sick: the new science of Darwinian medicine*. New York: Vintage.
- Nesse, R. M., & Young, E. A. (2000). The evolutionary origins and functions of the stress response. In G. Fink (Ed.), *Encyclopedia of Stress*, (vol. 2, pp. 79–84). San Diego, CA: Academic Press.
- Nutt, D. J., & Ballenger, J. C. (2003). *Anxiety disorders*. Malden, MA: Blackwell.
- Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: toward an evolved module of fear and fear learning. *Psychological Review*, 108, 483–522.
- Poulton, R., & Menzies, R. G. (2002). Non-associative fear acquisition: a review of the evidence from retrospective and longitudinal research. *Behavior Research and Therapy*, 40, 127–149.
- Powers, W. T. (1973). *Behavior: the control of perception*. Chicago: Aldine.
- Rabin, M. (2000). Diminishing marginal utility of wealth cannot explain risk aversion. In D. Kahneman, & A. Tversky (Eds.), *Choices, values and frames* (pp.202–208). New York: Cambridge University Press.
- Real, L., & Caraco, T. (1986). Risk and foraging in stochastic environments. *Annual Review of Ecology and Systematics*, 17, 371–390.
- Sapolsky, R. M. (1992). *Stress, the aging brain, and the mechanisms of neuron death*. Cambridge, MA: MIT Press.
- Schmidt-Nielsen, K. (1990). *Animal physiology: adaptation and environment*. Cambridge: Cambridge University Press.
- Schrödinger, E. (1944). *What is life? The physical aspect of the living cell*. Cambridge: The University Press.
- Seligman, M. E., Maier, S. F., & Geer, J. H. (1968). Alleviation of learned helplessness in the dog. *Journal of Abnormal Psychology*, 73, 256–262.
- Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*. Urbana, IL: University of Illinois Press.

- Stabell, O. B., Ogbobo, F., & Primicerio, R. (2003). Inducible defences in *Daphnia* depend on latent alarm signals from conspecific prey activated in predators. *Chemical Senses*, 28, 141–153.
- Sternbach, R. A. (1963). Congenital insensitivity to pain. *Psychological Bulletin*, 60, 252–264.
- Tollrian, R., & Harvell, C. D. (1999). *The ecology and evolution of inducible defenses*. Princeton, NJ: Princeton University Press.
- Tversky, A., & Kahneman, D. (1974). Heuristics and biases. *Science*, 185, 1124–1131.
- von Bertalanffy, L. (1969). *General system theory: Foundations, development, applications*. New York: G. Braziller.
- Webel, E. R., Taylor, C. R., & Bolis, L. (1998). *Principles of animal design: the optimization and symmorphosis debate*. New York: Cambridge University Press.
- Wiener, N. (1948). *Cybernetics; or, control and communication in the animal and the machine*. Cambridge, MA: Technology Press.
- Williams, G. C., & Nesse, R. M. (1991). The dawn of Darwinian medicine. *Quarterly Review of Biology*, 66, 1–22.
- Young, E., & Vasquez, D. (1996). Hypercortisolemia, hippocampal glucocorticoid receptors, and fast feedback. *Molecular Psychiatry*, 1, 149–159.