The Smoke Detector Principle
Natural Selection and the Regulation of Defensive Responses

RANDOLPH M. NESSE
Department of Psychiatry, The University of Michigan, Ann Arbor, Michigan 48104, USA

ABSTRACT: Defenses, such as flight, cough, stress, and anxiety, should theoretically be expressed to a degree that is near the optimum needed to protect against a given threat. Many defenses seem, however, to be expressed too readily or too intensely. Furthermore, there are remarkably few untoward effects from using drugs to dampen defensive responses. A signal detection analysis of defense regulation can help to resolve this apparent paradox. When the cost of expressing an all-or-none defense is low compared to the potential harm it protects against, the optimal system will express many false alarms. Defenses with graded responses are expressed to the optimal degree when the marginal cost equals the marginal benefit, a point that may vary considerably from the intuitive optimum. Models based on these principles show that the overresponsiveness of many defenses is only apparent, but they also suggest that, in specific instances, defenses can often be dampened without compromising fitness. The smoke detector principle is an essential foundation for making decisions about when drugs can be used safely to relieve suffering and block defenses.

KEYWORDS: Evolution; Natural selection; Smoke detector; Defenses; Anxiety; Stress; Signal detection; Pharmacology; Darwinian medicine; Medicine

Some 15 years ago, in the middle of a summer afternoon while seeing patients in the anxiety clinic, it suddenly struck me that I might be harming my patients. I knew that anxiety is useful and the mechanisms that regulate it have been shaped by natural selection. The treatment of anxiety disorders is basically an exercise in downregulating those mechanisms by whatever means work: behavior therapy, psychotherapy, and medications. But if natural selection is efficient, and it is, then it must have set the regulation systems to a point somewhat near the optimal. Here I was tampering with those systems. My patients didn’t seem to be harmed, however. In fact, treatment relieved their suffering and allowed most to return to full functioning. The easy answer, of course, is that these patients were in the clinic precisely because their anxiety regulation mechanisms were awry. But so many of us feel so much anxiety all the time. Would it be wise to eliminate it?

Address for correspondence: Randolph M. Nesse, M.D., Department of Psychiatry, The University of Michigan, 426 Thompson St., Room 5057, Ann Arbor, MI 48104. Voice: 734-764-6593.
nesse@umich.edu
My thoughts quickly went to the rest of medicine. General practice consists largely of relieving suffering by blocking defensive responses. Fever, pain, nausea, vomiting, diarrhea, cough, and anxiety are all latent traits that protect us when they are aroused by cues associated with some danger. People who are born without a capacity for pain are dead by early adulthood. People who do not cough after surgery are likely to get pneumonia. Fever is a useful component of our defense against infection. Patients with shigella infection who take drugs that block diarrhea are likely to get complications. So, how is it possible for us to block such useful defenses so routinely with such apparent safety? It almost seems as if natural selection has made a mistake. As Schopenhauer put it in 1851, "If the immediate and direct purpose of our life is not suffering, then our existence is the most ill-adapted to its purpose in the world." Long before, the Buddha made his First Noble Truth, "Life is suffering." Darwin, with his usual insight, had a different take: "Pain or suffering of any kind, if long continued, causes depression and lessens the power of action; yet it is well adapted to make a creature guard itself against any great or sudden evil.

This question then, about the regulation of defenses, is intimately connected to the larger question of the origins and utility of suffering. Most kinds of suffering are associated with the expression of a defensive response. It is aversive to experience cough, diarrhea, or fatigue. Anxiety and pain are aversive in their essence. This association of defenses and negative affect makes great sense from an evolutionary point of view. What could be more useful than a system to motivate escape from and avoidance of situations that harm fitness? As demonstrated so tragically by those who lack the capacity for pain, the capacities for suffering are useful. As Edward Wilson put it in Sociobiology, "Love joins hate; aggression, fear; expansiveness, withdrawal, and so on; in blends designed not to promote the happiness of the individual, but to favor the maximum transmission of the controlling genes."

Not all defenses are inducible. Some, like the turtle's shell, are constantly expressed; others, such as flight, fever, vomiting, stinging, and playing dead, are inducible. Whether a defense is expressed constantly or only in certain situations depends on its costs and whether it can be induced quickly. A turtle's shell cannot. Other defenses are permanent but develop only in the presence of danger. Water fleas, for instance, grow a pointed hood that deters predators only if chemical from the predator is in the water while they develop. A thriving area of biology investigates such defenses and their regulation. Other defenses, such as tanning and callus formation, are persistent but temporary responses to relevant cues, in these cases, sun exposure and friction, respectively. Medicine is more interested in physiological defenses that are aroused temporarily in response to specific challenges. Whereas homeostasis describes the body's ability to maintain a constant interior environment, the term rheostasis better describes the many adjustments the body makes to changing situations. The mechanisms that regulate inducible defenses have been shaped and adjusted by natural selection. They should, in most natural situations, express defenses to a degree that is in the vicinity of the optimum.

It seems, however, that defense expression is often excessive. Birds at the feeder outside my window suddenly startle and flee every few minutes, for no reason I can detect. In humans, malaise from a minor infection prevents food gathering, while pain from a twisted ankle interferes with flight from a predator. (Although the endorphin component of the stress response may be designed precisely for such situations.)
tions.) Fear at the site of a fall on the ice may persist, even in summer. Many defenses seem especially overresponsive when one considers the safety with which they can be suppressed. A smashed finger heals even if a narcotic dulls the pain. Upper respiratory infections are eliminated even if aspirin blocks the fever and malaise. Lung infections resolve even if cough is suppressed, bowel infections usually get better even if we do take medication to stop diarrhea. Fatigue is readily and safely relieved by caffeine. And, anxiety can be substantially blocked without greatly increasing exposure to danger. Such experiences explain the prevalence of the "clinician's illusion" that defenses are defects.

Given that so much of medicine is devoted to suppressing such defenses and the aversive sensations that accompany them, it is surprising that there is no generally accepted method for deciding when this suppression is wise and when it is not. As a result, markedly divergent opinions exist. Some people believe it is generally unwise to interfere with nature's defenses, while others are quick to use drugs to relieve any discomfort. Surprisingly many doctors do not explicitly distinguish manifestations of disease that are defenses from those that are defects. The general approach to benefit/cost analysis of defenses that is described below does not always provide specific guidance, but it can clarify the relevant variables and provide a quantitative framework for bringing them to bear on decisions regarding a defense's utility. It was inspired by Cosmides and Tooby's phrase, "Darwinian algorithms specialized for predator avoidance, that err on the side of false positives in predator detection." (See Cosmides and Tooby, p. 296.)

While the term "defense" might apply to almost any organismic system, we will here restrict it to those inducible responses that protect against some relatively specific potential harm. Most such defenses tend to be relatively inexpensive compared to the harms they protect against. Vomiting, for example, may cost only a few hundred calories and a few minutes, whereas not vomiting may result in a 5% chance of death. An episode of flight may cost 100 calories, whereas not fleeing may result in being caught by a predator. Defenses against social threats pose such special complexities that they will not be considered here. For the sake of simple exposition, I will analyze two kinds of situations separately: (1) all-or-none defenses against threats that are either present or absent and (2) graded responses to graded threats.

ALL-OR-NONE DEFENSES

Vomiting and panic are examples of defenses that are expressed either fully or not at all. The apparatus that regulates such defenses must make a "go/no go" decision based on whatever information is available. This information is rarely definitive. For instance, when an animal hears a rustle behind a bush, it cannot tell for sure whether this is a predator or not, yet it must instantaneously decide whether or not to flee. (The option of gathering more information will be ignored for now.) Flight will pay off whenever the expected cost of harm times the expected probability of the harm is greater than the cost of expressing the defense. The organism does not calculate this parameter, but some estimate must be involved in whatever mechanism regulates the defense. A relatively inexpensive defense, like flight, should be expressed if there is even a small chance of a catastrophic harm, like attack by a predator. As
FIGURE 1. Optimal defense expression as a function of the cost of harm and cost of defense. At point A C(H) is two times C(D), so the defense should optimally be expressed whenever there is greater than a 5% chance that the predator is actually present. At point (B), C(H) is 20 times C(D), so the defense should optimally be expressed whenever there is greater than a 5% chance that the predator is actually present.

Lima and Dill put it, "Few failures are as unforgiving as failure to avoid a predator. Being killed greatly decreases future fitness." (See Lima and Dill,\textsuperscript{14} p. 619.)

This is illustrated by Figure 1, which shows the expected probability of harm \([P(H)]\) above which it becomes worthwhile to express a defense, as a function of the ratio of the expected cost of harm \([C(H)]\) to the cost of defense \([C(D)]\). In a situation described by point "A," for instance, the C(H) is two times C(D), so the defense should optimally be expressed whenever there is greater than a 50% chance that the predator is actually present. At point B, C(H) is 20 times C(D), so the defense should optimally be expressed whenever there is greater than a 5% chance that the predator is actually present. Thus, it is easy to see why newly hatched chicks should flee from even crude cardboard models of hawks. In many situations, of course, C(H) may be 1000 times greater than C(D), thus making expression of the defense worthwhile whenever \(P(H)\) is only one in a thousand. In this situation, 999 out of 1000 expressions of defense will turn out to be unnecessary, but they are, nonetheless, a normal and necessary price for ensuring that the defense is expressed when it is needed.

This approach is based on signal detection theory,\textsuperscript{15,16} but it could equally well be derived from any of the maximization principles: risk assessment theories, expectation × value theories in psychology, or utility maximization in economics. When signal (stimulus from a threat) and noise (stimulus from a nonthreat) have overlap-
ping distributions, the shapes and means of those distributions can be used, along
with information about the prior probabilities of signal versus noise and the relative
values of correct responses, false alarms, and missed responses, to determine the op-
timum stimulus threshold at which the response should be expressed. For instance,
whether an animal should flee from a sound from behind a rock depends on the loud-
ness of the sound, the relative frequencies of stimuli from predators and other sourc-
es, and the costs of a false alarm (unnecessary flight) versus the costs of not fleeing
if the sound was from a predator. Although signal detection theory has the virtue of
providing separate estimates of sensitivity (ability to distinguish signal from noise)
and bias (tendency to give excessive false positives or false negatives), its standard
mathematical description is awkward to apply directly to the problem of defense reg-
ulation, and therefore will not further be considered here.

The costs of unnecessary defense expression and the costs of not responding to
actual threats are opposing forces of selection that act, respectively, to increase and
decrease the threshold at which the defense is expressed. These forces also tend to
improve the sensors and regulatory algorithms to the point where incremental bene-
fits of further improvements are fewer than the incremental costs. To take a hypo-
thesical example, the eyesight of a songbird species that has recently moved to an
environment with predatory hawks should improve its ability to discriminate hawks
from harmless birds, but only up to the point where (a) the costs of further improve-
ments outweigh the benefits, (b) visual acuity is no longer a limiting factor in defen-
sive response, or (c) constraints make further improvements unlikely.

After these improvements, there will still be some distance at which the discrimi-
nation of hawks remains uncertain. The escape response will be optimally expressed
at that distance where the probability that the bird is a hawk equals the defense cost
divided by the expected cost of an attack of a hawk from that distance. Many false
alarms will result. I initially called this the "principle of defense over-responsiveness,” (see Nesse, p. 283), but when I was working with George Williams, he suggested the analogy with smoke detectors. No one would want a smoke detector that almost always detected a fire; we want one that goes off every single time here is a blaze of any kind. To get this reliability, we are willing to accept false alarms. Of course, when the alarm goes off every time you boil water, it soon becomes tempting to take out the battery, in which case less sensitivity would be desirable.

GRADED DEFENSES AGAINST GRADED DANGERS

Many defenses, for instance, fever, fatigue, and fear, are expressed to varying de-
gresses depending on the severity of the threat. A model of such defenses requires sev-
eal variables. The independent variable is level of defense \(L(D)\), for instance, fever
of 1, 2, or 3 degrees. \(C(D)\) increases with \(L(D)\). This increase may be linear, as in the
cost of walking 1, 2, or 3 miles per hour, or may curve upwards sharply, as in the cost
of fever of 1, 2, or 3 degrees. There is also a minimum cost of defense that includes
the costs of maintaining the defense and its regulatory system, and any constant low
level of expression. The main dependent variable is \(C(H)\). This may include factors
such as loss of foraging days because of illness and decreased mating competition
from an injury. Because \(C(H)\) can be indefinitely large (death), a cost/cost model is
more appropriate than a benefit/cost model. There may also be some minimum cost

of harm that cannot be avoided by any amount of defense, for instance, the infection that invariably follows sufficient exposure to certain pathogens. The total cost is the sum $C(D) + C(H)$. Fitness is maximized when the defense is expressed to the degree that minimizes total cost.

Like most resources, small amounts of defense offer substantial benefits, whereas increasing amounts have declining marginal utility. For instance, small investments in grooming offer substantial protection from parasites, whereas constant grooming offers little additional protection. Also, as noted above, many defenses are inexpensive compared to the potential harms they protect against. The requirements for curves that describe these relationships are quite general: the cost of expected harm must rapidly decrease in response to relatively low levels of defense and decrease more slowly with higher levels of defense.

Figure 2 uses arbitrary cost units to illustrate the relationships among $L(D)$, $C(H)$, and total cost. $C(H)$ is portrayed as an exponentially declining function, while $C(D)$ increases linearly. These portrayals are relatively arbitrary and are intended only to illustrate the model and some general principles that follow from it. In Figure 2A, the minimum defense cost and the minimum cost of harm are both set at 0.5. At a level of defense of 1.0, a $C(D)$ of 1.5 brings $C(H)$ down to 1.5, for a minimum total cost of 3.0. This graph could represent the benefits of increasing levels of anxiety in a strange place or of different degrees of inflammation around a foreign body in the skin.

Figure 2B illustrates a situation in which incremental investments in defense give large decrements in harm. Here, total cost is minimized where the $C(D)$ is more than twice as large as the $C(H)$. Intuition may suggest that the optimum is where $C(H) = C(D)$, but in fact, defense expression goes up so long as an incremental investment yields a greater reduction in harm. Note in the graph that the arrow at the point where $C(D) = C(H)$ is far to the left of the point that gives the minimum total cost. Thus, there may be many body systems in which the defense costs seem senselessly high, even when they are actually optimal.

Yet another reason why defenses may seem to be too easily elicited is because errors of insufficient defense expression tend to be more serious than errors of excessive expression (Kim Hill, personal communication). The reason is illustrated by Figure 2A. In this example, a level of defense that is 7 units below the optimum results in a total cost 1.6 units greater than the minimum possible total cost, whereas a level of defense that is 0.7 above the optimum gives a total cost only 0.3 greater than the minimum. Thus, when small investments in defense offer big payoffs, and when defense costs increase slowly, natural selection will tend to shape systems that err on the side of excessive defensive expression. If defense costs escalate rapidly, as with high fever and the possibility of seizures, this principle is reversed, and could explain a tendency to err on the side of deficient response.

Still another factor arises when the mean value of a trait at which fitness is maximized differs from the central tendency of the stochastic distribution of the values that maximize fitness. Yoshimura and Shields demonstrate that observed values for a trait that interacts with environmental uncertainty are better predicted by integrating the fitness expected from a statistical distribution of the trait, instead of using the mean value predicted by a deterministic approach. In the case of defense regulation, defenses that seem unnecessary in many environments might be so valuable in
FIGURE 2. Optimum expression of graded defense. (A) Minimum defense cost and minimum cost of harm are set at 0.5. (B) Incremental investments in defense give large decrements in harm. Abbreviations: C(H): cost of harm; C(D): cost of defense; C(Tot): total cost.

occasional harsh situations that observed defense expression might be much greater than otherwise expected. Our world is much safer now than it used to be, so the utility of extreme defenses, such as panic attacks, is hard to see.
STRESS AND ANXIETY

The stress system offers a good example. The costs of stress are obvious (now that we have studied them!). Resistance to infection is lowered, bones thin, cancer is more likely, and atherosclerosis progresses faster.\footnote{22} It would seem sensible to tone down the stress system. However, the benefits of the stress system are even more dramatic when they are absent; people with Addison's disease are very likely to die with even the minor stress that comes from a routine infection. Cortisol secretion and other changes adjust the body to a mode where energy can be readily metabolized and the body is prepared for action.\footnote{23} Some anomalous actions of cortisol, such as blocking inflammation, have been interpreted as a way to protect the body from other parts of the stress response.\footnote{24} If this is the case, then other effects of stress must be costly indeed. The difficulty in deciding whether the adrenal cortex is a direct part of the stress response or a defense against it underscores the difficulties in quantifying the costs and benefits of defenses.\footnote{25} One might, in fact, think of components of the stress response as defenses that are potent, but so costly that they must be reserved for situations in which they offer major benefits, and even then must be dampened. The evolutionary reason why the stress response damages tissues may be quite simple. Useful changes that do not damage the body can be expressed at any time. But those useful changes that damage tissues are best contained in an emergency kit that is opened only when the substantial costs are worth it. This is why chronic exposure to stress creates an allostatic load that slowly but inevitably damages the body.\footnote{26}

Anxiety is often thought to be just another aspect of the stress response, but it is actually quite a different system. Stress prepares the body for action, anxiety prepares for danger.\footnote{27} Preparation for action is often an appropriate part of preparing for danger, and cortisol is sometimes aroused by anxiety, but very often there will be no cortisol increase despite intense anxiety.\footnote{28} Conversely, the signal that initiates the stress response, secretion of corticotrophin-releasing hormone from the hypothalamus, reliably arouses anxiety and even panic.\footnote{29}

From the point of view of a regulated defense, it would seem that there would be many disorders of deficient anxiety, just as there are disorders of excessive anxiety. Almost certainly there are, but because of the "clinician's illusion" we recognize and treat only disorders of excessive anxiety. To be fair, it is also true that patients with hypophobia do not line up outside the door of the anxiety clinic demanding the latest drug to increase their anxiety. The pleasure principle is just that, a tendency to prefer positive and avoid negative states, irrespective of what is best for one's health. I have looked hard for evidence for the early death and disability expected in people with too little anxiety and have found remarkably few studies. One particularly strong one examined a sample of 18-year-olds, some of whom had been severely injured in falls in their first few years of life.\footnote{30} The expectation was that this would result in a high rate of height phobias; the result showed that these children were only one-sixth as likely to have height phobias as the children who had not had a severe fall in childhood. It appears that reckless toddlers had become fearless adolescents!

Both stress and anxiety seem to be good candidates for routine pharmacological blockade. When all is said and done to demonstrate that natural selection has indeed shaped defense regulation systems that are somewhere in the vicinity of optimal, it becomes clear that the vast majority of the time, especially in our very safe modern
environment, we benefit from only a tiny proportion of the anxiety and stress we experience. Why not block the rest so we can suffer less and live longer? This apparently theoretical question is rapidly becoming of great practical importance. New medications now can block anxiety reliably. The only difficulties are side effects and dependence with eventual withdrawal reactions. If these can be overcome, there will be a great temptation for many normal people to reduce their moderate levels of anxiety, and huge profits will be made providing them with medications. Will this be wise? Only a systematic assessment of costs and benefits will answer the question.

**CONCLUSION**

The smoke detector principle shows that the overresponsiveness of many defenses is an illusion. The defenses appear overresponsive because they are "inexpensive" compared to the harms they protect against and because errors of too little defense are often more costly than errors of too much defense. Uncertainties of information result in these defenses being normally expressed when danger is not actually present. False alarms are to be expected and accepted. The degree of expression of graded defenses may also seem excessive because fitness is determined not by absolute, but by marginal costs and benefits.

These conclusions have several clinical implications. The first is that much apparently excessive defense is actually normal and useful. Physicians already know that excessive blockade of cough can result in pneumonia. They need also to be alert to the possibility that the blockade of fever, pain, nausea, diarrhea, rhinorrhea, and anxiety can also be harmful. The second implication is more supportive of clinical intervention. Even though Mother Nature's overprotectiveness is only apparent, it may still be possible to block many of our discomforts without causing much harm. This is because (1) regulatory mechanisms tend to err on the side of excess defense, (2) a full defensive reaction is often not needed to avoid a particular instance of a threat, (3) we have multiple redundant defenses, and (4) our environment is much safer than it was at the time we evolved. Thus, many situations exist in which it is safe to use medications to dampen or block anxiety, stress, pain, and other kinds of suffering without decreasing individual fitness. Pharmacotopia may be possible.

Because the goal of this paper is to explain certain apparently excessive reactions, I have emphasized inexpensive defenses. A benefit/cost approach could also be applied to expensive defenses, like high fever or posttraumatic stress reactions. Such defenses might appear to be underresponsive. A benefit/cost analysis could also be applied to situations that are risky as well as expensive, like a challenge to a social superior, although the complexities of such an analysis would be formidable.

In practical terms, we remain far from knowing how to distinguish situations in which more or less of a defense would be beneficial. This knowledge is urgently needed, as we quickly gain the capacity to pharmacologically block defenses more specifically and more safely. In particular, arguments about the use of psychotropic drugs are now largely based on medical dangers and possible dependency, but as new agents have fewer such problems, such debates will be superseded by questions about the wisdom of blocking normal defenses. In order to answer such questions, we will need to know what each defense is for, how it is regulated, and the kinds and
amounts of costs and benefits it offers in different situations. The principles outlined above offer a framework for addressing such questions.

REFERENCES