

SIGNAL, DECISION, ACTION*

Consider an inner state guiding a particular behavior in an idealized animal, and an external condition to which this behavior is an appropriate response. The inner state is also linked to the idealized animal's senses. What kinds of statistical relations will obtain among the action-steering inner state, the animal's senses, and the world when the animal is wired up in an optimal way?

The behavior I am discussing is best performed when, and only when, the important external condition obtains. So in a sense, the optimal inner state is one that is tokened when and only when the external state obtains. But I shall assume that this kind of optimality is impossible for the animal, because it does not live in an epistemically benign world. Although the important environmental state does have characteristic effects on the animal's senses, it is possible for random "noise" in the environment to produce some of these sensory states as well. So, if the inner state is thought of as a belief about the state of the world, it is not possible for the animal always to believe truly. The topic of this paper is the animal's optimal wiring, where this optimality is understood as relative to a set of constraints imposed by imperfect senses and an unfriendly environment.¹

I suggested that the inner state could be thought of as a belief about the world. This idea will have heuristic value at a number of points in the discussion, but it is not essential to the model. Some might prefer to think of the inner state as a decision to act, others might refuse to interpret it in this kind of way at all. As one of the products of this discussion will be an outline of a naturalistic semantic theory, it is important that everything in the model can be described nonintentionally. But there are other ways to put the model to philosophical work, and many of these projects are more easily

* This paper has benefitted from audiences at the University of Pennsylvania, UC Berkeley, the 1990 conference of the Australasian Philosophical Association, the University of Maryland, and Rijeka and Zadar (Yugoslavia). I am especially grateful, however, to Philip Kitcher. I would also like to thank the University of Sydney for financial support.

¹ For discussions of other aspects of the adaptive wiring of simple model animals, see Valentino Braitenberg, *Vehicles: Experiments in Synthetic Psychology* (Cambridge: MIT, 1984); and Paul Churchland, "Cognitive Neurobiology: A Computational Hypothesis for Laminar Cortex," *Biology and Philosophy*, 1, 1 (1985): 25–51.

presented if the question about reducing semantic properties is ignored much of the time.

There are five schematic letters that must be introduced at this point. Four of them refer to state types, and one to a variable. They are most easily remembered with the aid of a diagram (figure 1).

The important environmental condition is *S* (for "signal"). The other possible state of the world is *N* (for "noise"). *X* is a variable describing the animal's sensory input. *C* (for "cause") is the action-guiding inner state of the animal, and *M* is the movement or behavior which *C* produces. In this discussion, I shall assume that the link between *C* and *M* is fixed: *C* always causes *M*, and nothing else does. I shall also assume certain fixed relations between input *X* and the two-world states *S* and *N*; these will be described shortly. The connection that can be tuned is the link between *X* and *C*; the question is about which sensory events should cause *C* in an optimally wired animal.

The first outlines of the answer are obvious. The animal wants to produce *M* when and only when *S* obtains, and *M* results when and only when *C* is tokened, so the link between *C* and the senses should be the one that maximizes the correlation between *C* and *S*. Because the animal labors under constraints, this correlation cannot be perfect, but the animal should seek as good a correlation as it can manage.

The idea that valuable or virtuous cognition consists in a high degree of correlation between thoughts and the states of the world that the thoughts are about is familiar and well-explored in recent philosophy. This is the central idea of reliabilism in epistemology.² The appeal to reliable correlation has also spread from epistemology to the philosophy of mind, and it has become popular to claim that the semantic properties of internal representations depend, at least in part, on some kind of reliable correlation between thought and the world.³ The relevance of mind/world correlation to various philosophical problems is not the chief concern of this paper. Rather, the aim is to develop a reasonably precise theory of what a valuable mind/world correlation is.

² See D. M. Armstrong, *Belief, Truth, and Knowledge* (New York: Cambridge, 1973); Alvin I. Goldman, "Discrimination and Perceptual Knowledge," this JOURNAL, LXXIII, 11 (November 1976): 771-91, and *Epistemology and Cognition* (Cambridge: Harvard, 1986); Fred Dretske, *Knowledge and the Flow of Information* (Cambridge: MIT, 1981); and Robert Nozick, *Philosophical Explanations* (Cambridge: Harvard, 1981).

³ See Dretske, *op. cit.*; Robert Stalnaker, *Inquiry* (Cambridge: MIT, 1984); and Jerry A. Fodor, *Psychosemantics* (Cambridge: MIT, 1987).

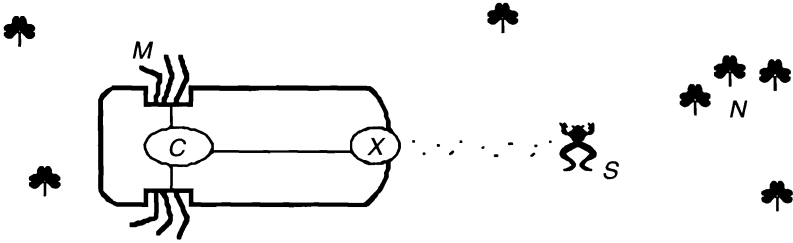


Figure 1

The first essential point is that there are two distinct kinds or “directions” of correlation to consider, and one of these is far more prominent in philosophical discussion than the other. Following Hartry Field,⁴ I shall call the currently favored direction of correlation *head-world reliability*. A perfect degree of head-world reliability exists when an inner state is tokened only if a certain external condition obtains. It is not required that the inner state always be produced by the external condition. So C is sufficient, but perhaps not necessary, for S . The second direction of correlation is simply the converse. A perfect degree of *world-head reliability* obtains if C is always tokened when S obtains, whether or not C obtains at other times as well. It is not hard to remember which direction is which, as Field’s labels follow the direction of a conditional with an arrow ‘ \rightarrow ’. Extreme head-world reliability obtains when a state of the head implies a state of the world; strictly, it is head \rightarrow world reliability. The converse is world \rightarrow head reliability. It does confuse matters slightly when we leave the extreme case of perfect correlation, and express the degree of correlation as a conditional probability. Then a high degree of head-world reliability is a high value of $Pr(S|C)$, and a high degree of world-head reliability is a high value of $Pr(C|S)$. But many people find that thinking of Field’s labels as abbreviated expressions using an arrow is a useful mnemonic.

It is a central contention of this paper that head-world reliability has been wrongly elevated to a central position in explanations of virtuous cognition. This is explicit in the work of Dretske, Armstrong, and Goldman. Dretske⁵ has built a detailed theory of knowledge, perception, and belief content around a notion of correlation

⁴ “‘Narrow’ Aspects of Intentionality and the Information-theoretic Approach to Content,” in *Information, Semantics and Epistemology*, E. Villanueva, ed. (New York: Blackwell, 1990).

⁵ See *op. cit.* and also his *Explaining Behavior* (Cambridge: MIT, 1988).

that is purely a matter of head-world reliability. Armstrong's 1973 account of knowledge and rational belief makes similar demands. In Goldman's magnum opus, "reliability" is central to his account of good cognition, and for Goldman "reliability" is simply head-world reliability.⁶ Goldman also sometimes discusses "power," which is admitted as another valuable property of cognition, and one that must be traded off against reliability. Power is a relative of world-head reliability, but they are not the same thing. Although this is not the place for detailed scholarship on the matter, I think it is clear that these are not isolated cases, that head-world reliability is central to many analytic and naturalistic conceptions of good, responsible, duly cautious cognition.⁷

World-head reliability has had a little more attention in the philosophy of mind. Although Dretske seeks to explain meaning with head-world reliability, the versions of indicator semantics proposed by Stalnaker (*op. cit.*) and Fodor (*op. cit.*) require both kinds of correlation. Nozick (*op. cit.*) has a similar two-way position in epistemology. Although this is a step in the right direction, I shall argue, the accounts given by Fodor, Stalnaker, and Nozick neglect or distort the most important fact about the relations between the two kinds of reliability; for they apparently require that thinkers simply bring their thoughts into both kinds of correlation with the world. But the critical fact here is the impossibility of pursuing both kinds of reliability simultaneously. Except in the most epistemically benign and friendly environments, the two kinds of reliability cannot be bought as a package deal; rather, one must be traded off against the other. This is seen readily in extreme cases; one way to achieve excellent head-world reliability is to believe almost nothing, but this sacrifices almost all of one's world-head reliability. World-head reliability can be pursued by believing everything that enters one's head, but this also leads to many false beliefs. The remainder of this paper will present a quantitative theory of the reliability trade off, and describe how to strike the balance optimally in particular situations.

I. OPTIMALITY AND RELIABILITY

My discussion of virtuous correlations between mind and world is based on signal detection theory. Signal detection theory originated

⁶ See his *Epistemology and Cognition*, esp. chs 1–6.

⁷ The work of Isaac Levi is an important exception; see his *Gambling with Truth* (New York: Knopf, 1967). In a much longer work (in preparation) I discuss in detail the place of the two kinds of reliability in contemporary epistemology and philosophy of science.

in the 1950s, combining statistical decision theory with electrical engineering to give a general description of both ideal and real signal-receiving devices. It has been used extensively in psychophysics, and has more recently been discovered by foraging theory in biology.⁸ I shall introduce the apparatus in an anthropomorphic, biologically unrealistic way at first, and then show how a core directive of the theory can be mechanized.

Suppose you are a carnivore, hunting in long grass, and you see a movement in one spot in this grass. This might be something edible creating the disturbance, or it might be just the wind. The decision problem is: Should you bound aggressively to that spot or not?

Assume first that your sensory input can be represented with a single variable X . A high value of X corresponds to the input from a large disturbance of the grass. The mathematics are simpler if we assume that there is only a finite number of values this variable can take. You also know a function relating various values of input X probabilistically to the presence of prey (signal), and a corresponding function for noise. That is, you know how likely each value of X is, given the presence of prey, and you know how likely that same input is to occur when there is no prey there. So what you know can be represented as in figure 2.

Earlier I said that the model would apply to cases where the animal is operating under constraints imposed by imperfect senses and an uncooperative environment. Figure 2 represents this fact in the overlap of the two functions. Degrees of input between 7 and 11, inclusive, can be produced by both signal and by noise. It is important not to read the value of X in the middle of the overlap ($X = 9$) as an input equally likely to have resulted from signal or noise; this depends on the unconditional, or prior, probabilities of signal and noise, as well as on the information in figure 2. In this discussion, the prior probability of signal, $Pr(S)$, will always be 0.2, so the prior for noise, $Pr(N)$ will be 0.8.⁹

A central measure in signal-detection theory is the *likelihood ratio*, $l(x)$, for an input. Let us avoid a profusion of subscripts in what

⁸ A standard reference is D. M. Green and J. A. Swets, *Signal Detection and Human Observers* (New York: Wiley, 1966). For psychophysical applications, see J. A. Swets, ed., *Signal Detection and Recognition by Human Observers: Contemporary Readings* (New York: Wiley, 1964). For applications in biology, see J. A. Staddon, *Adaptive Behavior and Learning* (New York: Cambridge, 1983); and D. Stephens and J. Krebs, *Foraging Theory* (Princeton: University Press, 1986).

⁹ The units of probability on the vertical axis are 0.0277 recurring.

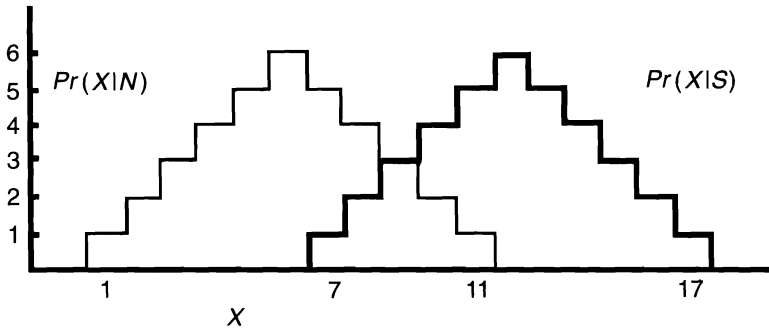


Figure 2

follows by reading an unembellished lower case 'x' as standing for some particular value of the variable X , some x_i .

$$(1) \quad l(x) = \frac{Pr(x|S)}{Pr(x|N)}$$

It is a consequence of my discrete example that the likelihood ratio is undefined for values of X to the far right, where the value of $Pr(X|N)$ is zero. For these values of X , which do not matter much in this discussion, $l(x)$ will be set artificially at infinity.

With the information in figure 2, and the prior probabilities for signal and noise, we can use Bayes's theorem to tell us what the probability is, given any input x , that there is prey in the grass.

$$(2) \quad Pr(S|x) = \frac{Pr(x|S)Pr(S)}{Pr(x)}$$

Similarly,

$$(3) \quad Pr(N|x) = \frac{Pr(x|N)Pr(N)}{Pr(x)}$$

The unconditional probability of x , $Pr(x)$, is equal to

$$Pr(x|S)Pr(S) + Pr(x|N)Pr(N),$$

but this is not important for the moment, as $Pr(x)$ cancels when from (2) and (3) we arrive at

$$(4) \quad \frac{Pr(S|x)}{Pr(N|x)} = \frac{Pr(x|S)Pr(S)}{Pr(x|N)Pr(N)}$$

Readers who find this opaque need not be concerned, as formulas (2) through (8) are only important in their justification of decision

rule (9), which we will come to in a moment. Much of the algebra can be skipped, and rule (9) taken on trust, if preferred.

Figure 2 and the prior probabilities of signal and noise tell you a lot about the likely states of the world behind your inputs, but none of this information tells us what you should *do*. To decide what to do, you must take into account the various ways of being right and being wrong, and their consequences. There are in fact four possible outcomes:

- (i) You decide there is prey, and there is. This is called a *hit*. You attack and get fed, and the value of a hit, or $V(H)$, here is 16 units, let us say.
- (ii) You decide there is prey, but there is none. This is a *false alarm*.¹⁰ You waste energy attacking, and maybe scare away other animals. The cost of a false alarm $V(FA) = -2$.
- (iii) You decide there is no prey, and indeed there is not. This is a *correct rejection*. The value of a correct rejection $V(CR) = 0$.
- (iv) You decide there is no prey, but there is. You miss out. This is called a *miss*. You have not actually spent anything, so the value of a miss here $V(M) = 0$.

Some might think a miss should have a definite cost. In fact it does not matter, for we are envisaging a limited number of opportunities for hits, and a competitive process between different decision rules. Misses are then effectively penalized. These pay-offs can be represented in a simple matrix.

The Foraging Matrix

	<i>S</i>	<i>N</i>
yes	$V(H) = 16$	$V(FA) = -2$
no	$V(M) = 0$	$V(CR) = 0$

S = prey; N = no prey
yes = attack; no = ignore

The expected pay-off from deciding to attack will then be:

$$(5) \quad Exp(yes|x) = V(H)Pr(S|x) + V(FA)Pr(N|x)$$

and the expectation from deciding no:

$$(6) \quad Exp(no|x) = V(CR)Pr(N|x) + V(M)Pr(S|x)$$

¹⁰ From the standpoint of the theory of statistical testing, if noise is a "null hypothesis," a false alarm is a type 1 error and a miss is a type 2 error. The relations between signal detection theory and other aspects of statistical theory are discussed in the longer work referred to in note 7.

Assuming that you are interested in simply maximizing your expected pay-offs, case by case, you should say yes if and only if:

$$(7) \quad V(H)Pr(S|x) + V(FA)Pr(N|x) \geq V(CR)Pr(N|x) + C(M)Pr(S|x)$$

This rearranges to give:

$$(8) \quad \frac{Pr(S|x)}{Pr(N|x)} \geq \frac{V(CR) - V(FA)}{V(H) - V(M)}$$

Substituting (4) into (8) gives, with some rearranging:

$$(9) \quad \frac{Pr(x|S)}{Pr(x|N)} \geq \frac{Pr(N)}{Pr(S)} \times \frac{V(CR) - V(FA)}{V(H) - V(M)}$$

The expression on the left of the inequality we know as the likelihood ratio $l(x)$, and the expression on the right is known as β . Within β , the quantity $V(CR) - V(FA)$ is the *importance of noise*. It combines the outcomes of the different actions you might perform as a response to mere noise. The quantity $V(H) - V(M)$ is the *importance of signal*; it combines the pay-offs for actions in the presence of signal. The quantities $l(x)$, β , and, within β , the importance of signal and noise, are the important parts of signal detection theory for us, and (9) is the important formula. It directs that you say yes, attack, if the likelihood ratio, calculated from the particular input you are now receiving and its relation to the probability functions for signal and noise, is greater than or equal to β , which is determined by the prior odds of signal and the pay-off matrix. (Strictly, it does not matter what you do when $l(x) = \beta$.) In this particular case:

$$\begin{aligned} l(x) &\geq \frac{.8}{.2} \times \frac{0 - (-2)}{16 - (0)} \\ &\geq 1/2 \end{aligned}$$

The value of $l(x)$ beyond which you should say yes determines in turn the critical value of X , the intensity of sensory input that you should consider sufficient to convince you to act. This is shown in figure 3. The critical value of X is labeled x_c . For any value of X at or higher than x_c , you should say yes.

We can see that in this graph the ratio between $Pr(X|S)$ and $Pr(X|N)$, where the value of X is x_c , is $1/2$. As the critical value of $l(x)$ delivered by formula (9) increases, the cutoff x_c moves to the right, and it moves to the left as $l(x)$ decreases.

We are now in a position to answer the questions that began this paper. We asked if there is a general principle describing the opti-

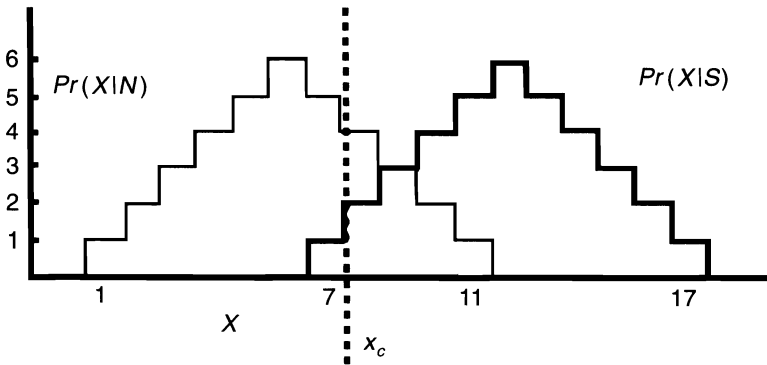


Figure 3

mal way for an animal to have an action-guiding inner state wired to its senses in order to achieve a desirable correlation between its thoughts and the world. Formula (9) is just such a principle. Given that M is caused by C , and the animal wants to produce M if and only if $l(x) \geq \beta$, the optimal C is a state tokened if and only if $l(x) \geq \beta$. So the animal in figure 1 is optimal if the link between X and C is such that C only fires if $l(x) \geq \beta$.

Crucially, the optimal animal does not need to represent to itself any of the algebra through which we have gone. It just has to have connected to M a state C that, as a matter of physical contingency, fires when the animal's sensory input is such that $l(x) \geq \beta$. It does not know that these are the inputs such that $l(x) \geq \beta$; it just happens to be physically tuned to that set of inputs.

Two objections can be raised and answered at this point to clarify the proposal. First, how much depends on the assumption that X , the sensory input, varies in one dimension only? Secondly, the optimal animal described above apparently does not really need C at all. The signal detection apparatus only describes the relations between S , X , and M . C is an epiphenomenal way station on the road to M . These objections can be answered simultaneously. It is not essential that the input X be one-dimensional; it can be as multidimensional as you please. It can extend over time, and can combine the states of a number of sense modalities along with memory. All that matters is that it make sense to assign a likelihood ratio to each possible total input. It must be possible to compare the probability of receiving that input as a consequence of signal and as a consequence of noise. Further, once we leave the idealized one-dimensional case, C does

not look so unimportant in the architecture. C , the inner state, is then the place where the different components of the input are combined.

Although formula (9) describes the optimal wiring between the senses, described by X , and the inner state C , it has not yet told us about the relations between C and S . We have not been told how an optimal inner state is correlated with the world. But this is easy to calculate from the information we have, for formula (9) is a description of how best to strike a balance between the competing demands of head-world and world-head reliability. Once formula (9) has established a cutoff point on the X axis, simple algebra tells us the head-world and world-head reliability values that result. The fastest way to do the calculation is to note first that the world-head reliability of C , $Pr(C|S)$ is simply the proportion of the area under the $Pr(X|S)$ function, to the right hand side of x_c . This area is the probability that S will produce a value of X that the animal will regard as sufficient to trigger C . In this case, $Pr(C|S)$ is about 0.972.

Once we have this figure, the head-world probability falls out of Bayes's theorem (formula (2) above, with C substituted for x). We first work out the unconditional probability of C , as the overall probability that a value of X will be produced, by signal or noise, which exceeds the cutoff x_c . This is done by summing the value of $Pr(x|S)Pr(S) + Pr(x|N)Pr(N)$, for each value of X over x_c . When this figure is divided into the product of the world-head reliability and the prior for signal, the head-world reliability is the result. In this case it is about 0.467. The animal has sacrificed head-world reliability in order to obtain a very high value of world-head reliability; it is tolerating a large number of false alarms in order to achieve as many hits as it can.

Generally, the further the cutoff point is shifted to the left, the more head-world reliability is being sold off in favor of world-head reliability. It is easy to see how this trade is driven by formula (9). When the cutoff x_c is shifted to the left, this is because a low value of $l(x)$ is being demanded by β in formula (9). Ignoring perturbations in the prior probability of signal, a low value of β in formula (9) will be due to a large "importance of signal" in the denominator, compared to the "importance of noise" in the numerator. Perhaps hits are very valuable, false alarms are very cheap, misses are very expensive, or some combination of these holds. On the other hand, as false alarms become expensive, or the importance of noise increases in other ways, the value of $l(x)$ demanded by formula (9) will rise. Then the cutoff x_c will be forced to the right, and head-world reli-

ability will increase at the expense of world-head reliability. In sum, as the importance of signal increases relative to noise, it becomes rational to trade head-world reliability for world-head reliability. As the importance of noise increases relative to signal, it becomes rational to trade world-head for head-world reliability.

So far we have only considered an example involving foraging for food. By shifting our attention to other goals, other ways of making the trade off can be illustrated quantitatively.

II. LOVE AND DEATH

The next case will be one in which the basic form of the trade off is the same, but for different reasons. This is the problem of avoiding one's own predators. A pay-off matrix for interactions with predators might be as follows.

Predation Matrix

S = predator in the vicinity; N = noise

yes = flight or concealment response; no = no behavior

Here the single most important case is the miss; if you fail to detect a predator, you risk injury and death. Hits and false alarms cost -2 units alike. Correct rejections have no consequences. Leaving the prior probabilities of S and N as they were in the previous case, the value of β becomes 0.2 (see figure 4).

In this case the average head-world reliability of C drops to about 0.375 . World-head reliability, on the other hand, is forced all the way to 1 . Because of the cost of a miss, the importance of signal is so high in this case that every possible degree of world-head reliability is purchased.

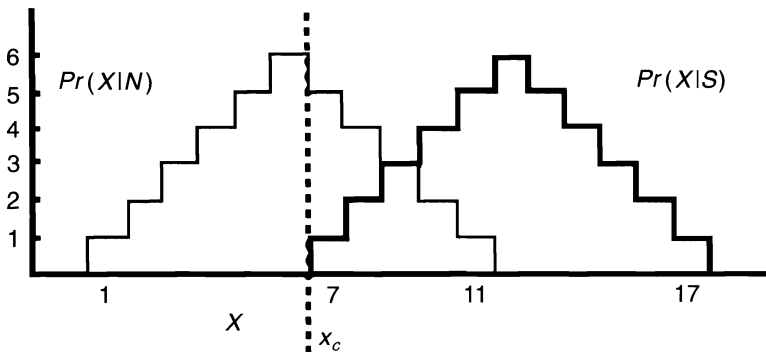


Figure 4

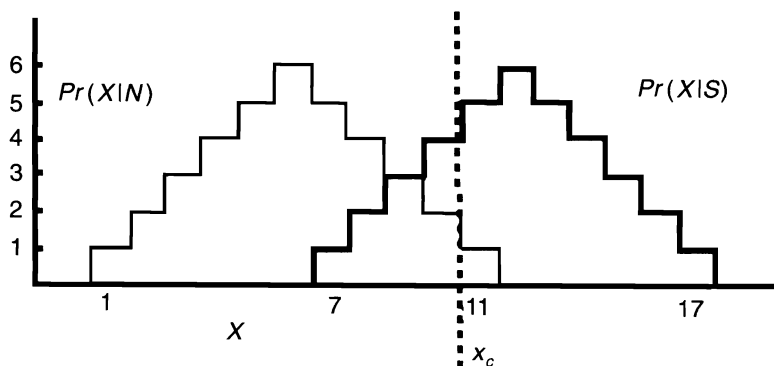


Figure 5

Finally, we might look at a case where the situation is reversed, and head-world reliability should be sought. A suitably ecological case might be one's relations with the opposite sex. We assume the animal is female, and in an environment where there are many males around who are members of different but closely related species to her. Her decision problem is the detection of true conspecifics with which to mate. Here is a matrix:

The Mate Choice Matrix

	S	N
yes	40	-50
no	0	0

S = conspecific male; N = male of another species

yes = mate; no = don't mate

In this situation both hits and false alarms are important. A hit results in pregnancy that costs -50 units, but also in offspring that is worth 90 units. A false alarm is getting pregnant to a male of a different species. The result will be inviable or infertile offspring, so the animal bears all the costs of pregnancy with no reward. Correct rejections and misses have no consequences; there is no shortage of males. The importance of noise is higher than in the previous examples, though the importance of signal is significant as well. Leaving the prior probability of signal at 0.2, the value of β is 5 (see figure 5).

The high value of β pushes x_c to the right and, as a consequence, the head-world reliability of C is now about 0.867. The animal has

paid for this with a world-head reliability of 0.722. It is tolerating more misses in its attempt to avoid false alarms.¹¹ Only in this last case does the optimal animal have the kind of head-world reliability that many philosophers have prescribed for good cognition. If a loosely reliabilist perspective on knowledge is assumed for a moment, then only in this last case might the optimal animal *know* that *S* obtains, when it tokens *C*. As this is largely a consequence of the pay-off matrix operating, a rough slogan suggests itself: the importance of knowledge is the importance of noise.

III. APPLICATIONS

Philosophers have focused on head-world reliability as an important raw ingredient in analyses of justification, meaning, and related concepts. But signal detection theory puts head-world reliability in its place, as a commodity to be sought ardently in some situations, but to be traded off in favor of world-head reliability in others. There are many philosophical projects that might be overhauled in the light of this fact.

In competition with the reliability-based approaches of Dretske, Stalnaker, Fodor, and others, a number of philosophers have recently proposed analyses of the content of mental representation which rely entirely on the biological functions of the representations and the devices that make use of them.¹² A version of this idea can be developed within the signal detection framework. Suppose that the history of an animal's interactions with some state of the world *S* has produced a state *C*, whose location in the animal's wiring is as close to optimal as could be managed in the circumstances. That is, some selective force has led the animal to approximate the optimal trade off between the two kinds of reliability in the wiring of *C*. If *C* has been wired as a detector of *S* in this way, it might have the function to represent *S*, and hence have *S* as its content. This is certainly too crude as it stands—we should have stopped expecting to hear four-line theories of meaning by now. But it suggests a way in which the teleological/functional view of meaning can be made

¹¹ Interestingly, it has not paid that much in world-head reliability to generate this very high head-world value. This fact is also discussed in my longer treatment, along with other features of the signal detection theory apparatus that I have simplified or skipped in this exposition.

¹² See Ruth G. Millikan, *Language, Thought and Other Biological Categories* (Cambridge: MIT, 1984), and "Biosemantics," this JOURNAL, LXXXVI, 6 (June 1989): 281–97; David Papineau, *Reality and Representation* (New York: Blackwell, 1987); and Mohan Matthen, "Biological Functions and Perceptual Content," this JOURNAL, LXXXV, 1 (January 1988): 5–27.

more precise than it is presently, or perhaps a way to steer a middle course between reliabilist and teleological approaches. Dretske, in his recent work, is already moving closer to such a middle course, and is held back only by a preoccupation with head-world reliability.¹³

Whether the explanation of meaning can benefit from signal detection theory or not, the theory has important consequences for the program of naturalistic epistemology. Most immediately, the signal detection perspective issues a challenge to those philosophers who claim special status for head-world reliability. This challenge has obvious similarities to pragmatist rebellions against orthodox epistemic values, sharing an emphasis on action and the practical consequences of cognition. Note, however, that while I assume a biological, behavior-linked determination of the pay-off matrices in this discussion, someone who assigned pay-offs on the basis of pure "epistemic utility" could make use of the same formal apparatus. Further, in contrast to many central strands of pragmatist thought, correspondence truth has a definite place in the signal detection approach. For, if *C* is regarded as representing *S*, hits and correct rejections are truths, while misses and false alarms are errors. So truth remains a definite goal of cognition, though the truth-linked virtue of "reliability" is taken apart and overhauled.

PETER GODFREY-SMITH

Stanford University

¹³ See especially his *Explaining Behavior*.