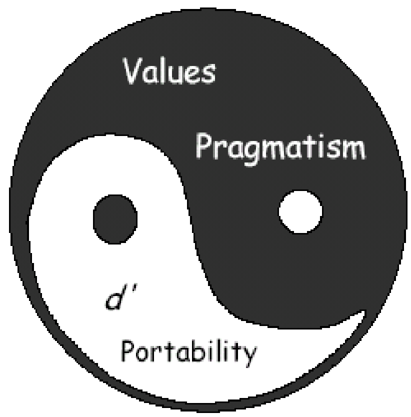
Behavior and Philosophy, *31*, 251-258 (2003). © 2003 Cambridge Center for Behavioral Studies

THE YINS AND THE YANGS OF SCIENCE

Peter R. Killeen
Arizona State University



ABSTRACT: Like the Tai Chi of the title, effective action is a blend of two modalities, the yang of science, the yin of desire. Without yang we are blind, without yin we don't look. Signal Detection Theory takes a line from this ancient book of changes. It clarifies why scientists—mavens of society's knowledge—must resist the temptation to arrogate their community's desire; and politicians—personifications of society's will—must resist the temptation to corrupt their community's data. Each can help the other understand, on the one hand what is knowable, on the other, what is worth knowing. Bias does not discredit truth; bias complements truth. Their conflation confounds knowledge when goals change, discomfits desire when facts change. Their preservation as disjunct modalities preserves the community.
Key words: bias, complementarity, criteria, data, *d'*, science, truth, values

Knowledge and desire are intimately related. Science is a systematic way of acquiring knowledge. Of what use is such knowledge? As a guide to action. Why do we act? To achieve goals. What are goals? Things we desire. Unmotivated action occurs for no reason; it is irrational. Generous, wise, altruistic, and synoptic acts are not unmotivated; they are motivated by desires we like. We like them because they typically have good consequences for our community. Mean, stupid, selfish, and narrow actions have consequences we dislike. They often hurt our

AUTHOR'S NOTE: This article is a response to the editor's call for commentary on the place of facts and values in science. Please address all correspondence to Peter Killeen, Department of Psychology, Box 1104, McAllister Street, Arizona State University, Tempe, AZ 85287-1104.

community. Science is a tool of desire. Desire for things that we value motivates the community to subsidize scientific research—but sometimes scientific results thwart desire.

We say a statement or belief is true to the extent there is correspondence between the *ding an sich* and the way we represent it.[1] Truth is valued in science, as it is nominally valued by all civilized humans. Getting correspondence between the phenomenon and your representation of it might be difficult, but it is often easier than getting correspondence between your representation and my representation—getting correspondence between what you call *true* and what I call *true*. True things have cachet. One can silence an argument by showing a premise to be false or a syllogism invalid. Truth is power: If you have the truth it forces antagonists to either (A) disavow truth telling, or (B) discredit the truth of your idea, perhaps by impugning your algorithm for allocating truth value.

Plan A. No matter how vicious and indifferent the mother; no matter how unpalatable and indigestible the apple pie; no matter how counterproductive and misleading the truth—these are three things we must be on the correct side of, or we shall quickly lose our audience. Big Brother may have convinced the masses that Truth is a Lie, and Jesuits may believe things *because* they are impossible; but without ample time to condition our audience, we must at least pay lip service to truth. We may call for a higher truth, but we cannot disavow truth. Plan A is out.

Plan B. *Fact* is the name of a category into which we place data that have been validated. Data in that category are called *true*. Some facts are as obviously true as "the dusk follows the day"; others as apparently true as "the earth is flat"; and others as personally and temporally true as "this cheese is good." Facts, like cheese, have flavors and shelf lives. What makes them facts is that a community has a taste for them, and their shelf life is longer than that of non-facts.[2] This paper is an exploration of the uses of Plan B.

To be put in a category such as *false*, however refined one's criteria, is a binary act, whereas data exist on an evidence continuum. "It's hot in here." Is that true? What centigrade does it take to make it a fact? Tobacco causes cancer, punishment is counterproductive, whites are smarter than blacks, sexual abuse leads to mental illness. Where do we draw the line?

Signal Detection Theory (SDT) Draws the Line

Signals can be clear or vague. Was that "I love you" she whispered? Or was it "I loathe you"? If the datum was irrelevant to your desires, it would be a mere Freudian curiosity that your partner lisped that crucial verb. But it is not irrelevant;

---

[1] Flexner (1987). Some philosophers reserve *true* for relations between sign and significant that are always and everywhere valid, a Platonic act of faith that gives comfort, if not confidence, in any particular assignment of truth value. All such assignments of factual status must suffer continual review.

[2] Many enduring and popular superstitions, alas, also have these features. The expropriation of *fact* by the scientific community, with their complex higher criteria, was a heavy blow to communities who value other ways of knowing (i.e., other beliefs).

there is an urgent need to know. The best thing to do is to get a confirming message: “Uhh, could you please repeat that, Honey?” But that will not always work. To kiss or not to kiss? Consider the operative payoff matrix. The columns label the states of nature, to which the suitor (and perhaps even the speaker) is not privy. The rows list the actions contemplated. The cells are the payoffs—the consequences of value.

Act Said:	*Love*	*Loathe*
Kiss	1. Sleep with	2. Slapped face
Do not kiss	3. Lose sleep	4. Saved face

Repeating this scenario many times with an unsophisticated suitor or with a population of similar suitors permits us to estimate the probability of an action appropriate to the state of nature (Cell 1 or Cell 4). That probability will vary as a function of the probability of an inappropriate action (Cell 2 or Cell 3) because the probability of both outcomes in the top row is greater for bold partners who require less evidence to advance. Precisely which outcome eventuates depends on the state of nature, which is shown in the top of Figure 1. Next, change how important the relative payoffs are to the individual, depending on his personality and current situation (some are humbled by a slap, others stimulated; absence makes the heart grow fonder, alternatives make the heart grow fleeter). Given a fixed signal strength, the probability of a true positive action (Cell 1) can range from 0 (for a very timid suitor with a high criterion for action) to 1.0 (for a libertine), as a function of the probability of a false positive (Cell 2) This is effected by an adjustable criterion. Slide the criterion *C* to the left (liberal bias) in Figure 1 or to the right (middle panel, conservative bias), and then take action if the evidence exceeds it—and hope for the best.

A wise suitor will attempt to increase signal strength. Perhaps he will observe body language, or review his history of reinforcement, or delay action until more ardent signals are received, or manage “I waited for ever for that word; say it again darling.” With that extra information, whatever his motivation, his expected payoff will increase, as shown by the bottom panel in Figure 1.

The relation between the probability of a true positive (the area under the right distribution to the right of the criterion) can be plotted as a function of the probability of a false positive (the area under the left distribution to the right of the criterion) as the criterion is moved from the far right to the far left. This is shown in Figure 2. When signal strength is weak, performance will follow the lower curve. When it is stronger, it will follow the upper curve. Given a particular level of motivation, we can calculate the optimal placement of the criterion. This is the operating point on the relative operating characteristic (ROC) curve that will maximize the expected value—the probability of the various outcomes weighted by their value. The optimal point is marked by the left hash mark on the lower curve (weak signal) and by the left hash mark on the upper curve (in the case of a stronger signal).

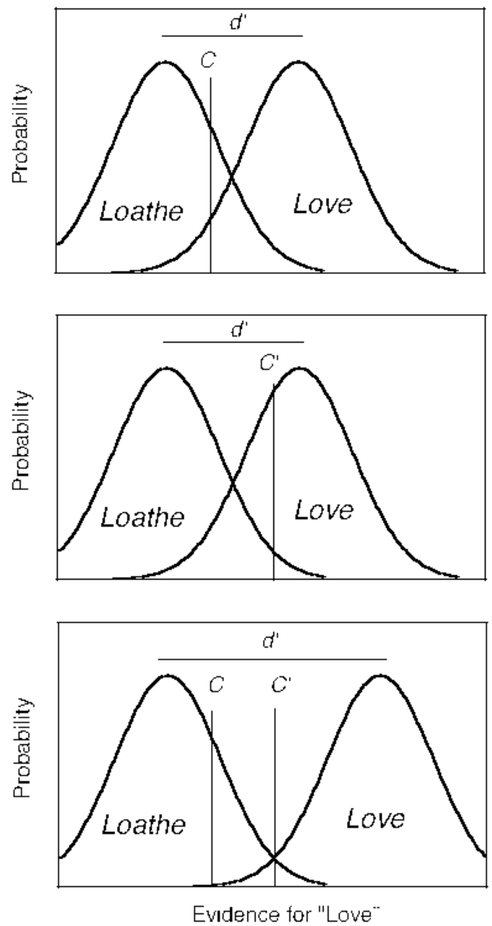


Figure 1. The evidence for the state of nature can be dispersed along a continuum. The distance between the modes of the distributions is called *d'*. A decision must be made in light of a new datum lying somewhere along the axis. An optimal strategy is to evaluate the costs and benefits of various actions and to set a criterion that then optimizes the expected value. This is shown as the line *C*; evidence falling above *C* should instigate positive action, that below it complementary action. In the middle panel the criterion is shifted to a more conservative position, *C'*, reducing false positives at the price of also reducing true positives. In the bottom panel additional evidence has been adduced to improve the diagnostic accuracy. Now the probability of missing a positive signal given the original criterion *C* is reduced almost to zero. An unbiased observer would place his criterion at *C'*.

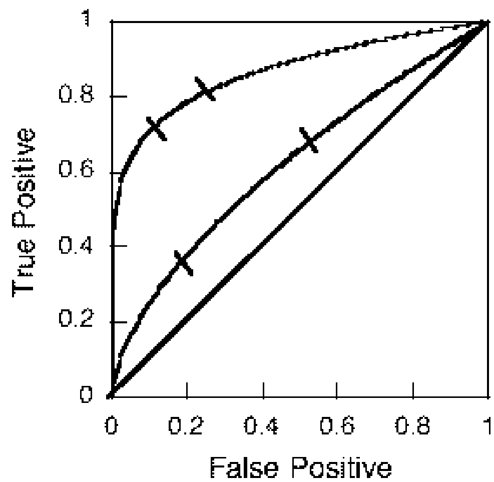


Figure 2. Relative operating characteristic (ROC) curves. As the criterion shown in Figure 1 is moved from right to left, the probability of true positives increases from 0.0 to 1.0 (the ordinates) along with the probability of false positives (the abcissae). The flatter ROC corresponds to the top panel in Figure 1, and the more peaked ROC corresponds to the bottom panel, evidencing a larger *d'*. The hash marks show the effects of a 2-to-1 change in motivation for each scenario.

There ensues a 2-week absence from the beloved, doubling the motivation. This will move the ideal operating points to the right marks. In the case of the indistinct signal, the change in motivation causes a large increase in the probability of a positive response. It increases both the probability of a true positive and a false positive, but in the case of the strong signal (top curve) it mandates a much smaller increase in the probability of a positive response. If the signal were stronger yet, with a curve hugging the left and top axis, motivation—bias—would play a minuscule role in informing the optimal response. In the case of a very weak signal, with a curve hugging the positive diagonal, the optimal response should be predicated almost entirely on the motivational state and outcomes.

Politics complements science. We seldom know the state of nature exactly. Arguments about truth often miss this point. All knowledge is corrigible. The question should be not "*What is true?*," but rather "*What to do?*" The first thing to do is to get as clear a signal as possible. This is the job of science. Science reduces the ambiguity of the signal, separating the distributions (Figure 1), and moving the ROC from close to the diagonal toward the upper left corner (Figure 2).

Then what to do? That depends on your motivations and those of others affected by your decision. Actions at that point become political—that is, concerned with the impact on the polity and its diversity of payoff matrices. The politic thing to do is to take into consideration the interests of all who have standing in the issue. This consideration sets the criterion, moving it to a point on the ROC that optimizes the objective, and counseling action whenever signal strength exceeds that criterion. This move is necessary to inform action, but whenever a continuum such as the evidence axis is reduced to a discrete category, information is lost. Scientists should not be precipitate in that reduction by rushing to categorize. Unfortunately, such categorizations—"true facts"—are often called for by unsophisticated audiences.

Truth is not in opposition to value. Truth is orthogonal to value. It is the task of science to increase the probability that what we think we know is true. That is, it is the task of science to increase *d'* (Figure 3).[3] This does not mean that scientists should tell us what to do. Scientists should tell us what *is*, qualified as best they can, and what the outcomes of relevant actions are likely to be. Finding the criterion that optimizes outcome might require technical expertise, but the involvement of scientists as technical consultants does not give them license to dictate society's values.

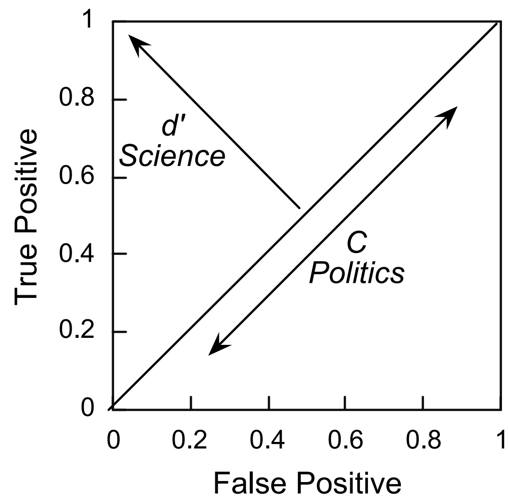


Figure 3. It is the task of science to increase *d'*—the truthfulness of things we believe to be facts. It is the task of politics to weigh costs and benefits and set a criterion for action in light of that knowledge: that is, to set *C*.

No matter what we know, at some point we must take action. It is the task of politics to direct that action toward agreed-upon communal goals. That is, it is the task of leaders to assess the expected values of outcomes for their community and to take action in light of the best available information. Political bias (criterion setting) is not bad; it is the only way to optimize expected outcomes. It only seems bad when the criterion setter does not have your values as the first priority (i.e., when her bias is different from yours).

---

[3] *d'* is strictly independent of *C* only when the ROC curve is symmetric around the negative diagonal. In Figure 2 a slight liberal bias maximizes the probability of a correct decision. This does not undermine the importance of distinguishing between fact and value; maximizing the expected utility of a decision requires a separate accounting of evidence and assignment of value and their eventual recombination in a Bayesian equation.

Decision makers need not assert that absolute truth is on their side. There is no absolute knowledge of truth. They need to need to convince their audience that the selected course of action optimizes the expected value for the community. This rational position permits the politician to change course as new data—or new members of the constituency—become available without having to assert a revelation about the new absolute truth.

Conclusions

Neither science nor politics is easy, but each has its own domain. A critical part of politics is comparing often incommensurable outcomes—is it better to build the bomber or the thousand new schools? Is it better to incarcerate the potential terrorist or defend the principles of civil liberties? Parties will make their best case that their scheme for weighting outcomes should be adopted. One tactic for achieving this is to attempt to discredit the validity of the opposition's facts. Such a challenge is often useful because all facts are corrigible.[4] Such inquiries can be as scientific—or more so—than those that birthed the original facts.

A less useful tactic is to corrupt the original data or their analyses—that is, to introduce bias in the process of inquiry rather than its proper place in criterion-setting. Although this practice is often effective in achieving short-term goals, such myopic pragmatism undermines the long-term utility of science. By degrading the process of inquiry it reduces the expected payoff in all other applications of science. Good science is portable: Its techniques work independently of one's political persuasion. When this ceases to be the case, science dies and communities are blinded. The introduction of political bias into scientific process—Lysenkoism—can bankrupt a society.

Scientific research improves decision-making. Executives must decide what decisions most need the light of science. Allocation of resources for science is, and should rightfully be, a political process. Because elected officials value immediate benefits more heavily than other citizens might, there is an enduring conflict of interests in such decisions. Here is where the forceful scientist can make a difference—as an informed and passionate player in the public arena. Rachel Carson, Carl Sagan, and Stephen Jay Gould advanced science by putting its good face on the stage. They did so both to conserve the environment and to display the beauty, complexity, and power of scientists' ways of knowing.

Figure 3 is admittedly idealistic; value and fact are more intimately intertwined, more like the Tai Chi symbol of yin and yang. Assertion of truth requires criteria for classification of facts, which are often driven by political expedience. What counts as an act of aggression depends on how much a country wants war; what counts as trade infringement depends on whose companies will suffer. But ideals such as Figure 3 are useful as a paragon. Science constitutes a

---

[4] With hindsight we can select which facts have had staying power, but we lose the utility of the word if we deny factual status to reasonably well-validated phenomena because they have not stood the test of centuries. Are gravitational and inertial mass in fact equivalent? Is cold fusion really a canard? *Yes* and *yes* will do for now.

commons; a little pollution of a commons often benefits the polluter because he accrues benefit while the damage is diffused among all members of the community. When exploitation starts, however, it can precipitate a run on the bank: Why should my valid needs be denied when your selfish ones benefit from such chicanery? The community must therefore be vigilant in policing the intellectual commons, in keeping separate how it knows and what it wants. Failure to do so will not only corrupt its facts, it will ultimately impoverish its actions. When *d'* is 0, only the loud prevail.

References

Flexner, S. B. (1987). *The Random House dictionary of the English language*. New York: Random House.