Superstition: A Matter of Bias, Not Detectability

Abstract. *Pigeons discriminated between stimulus changes dependent on their pecking and stimulus changes occurring independently of their behavior. Their performance was accurate, and when the payoffs for "hits" and "correct rejections" were varied, their response bias varied in a fashion similar to that of human observers detecting signals in a background of noise.*

Many animals, when fed periodically, will engage in stereotyped patterns of behavior that have little apparent relation to the acquisition of food. Skinner (*1*) noted that such experiments "might be said to demonstrate a sort of superstition. The bird behaves as if there were a causal relation between its behavior and the presentation of food, although such a relation is lacking." Skinner suggested that the accidental coincidence of any behavior with food would increase its probability; the ensuing feedback loop might perpetuate one or several behavior rituals indefinitely. That explanation depends on automatic reinforcement through temporal contiguity—a principle known as the "Law of Effect." However, Staddon and Simmelhag (*2*) replicated Skinner's experiment and found that many of the stereotyped behavior patterns were never contiguous with reinforcement—the law of effect could not account for them. Staddon and Simmelhag hypothesized that these "interim" behaviors were elicited by the periodic feeding, and formed a pool of behaviors from which some might be selected and strengthened by a law of effect mechanism.

One of the reasons for belief in the law of effect, despite such limitations, is its consistency with traditional criteria for causality: contiguity in space, precedence in time, and logical necessity. When reinforcement is arranged by experimenters, all of these criteria are usually satisfied. When spatial contiguity is varied, strong "sign-tracking" effects (*3*) affirm the importance of physical proximity to reinforcers or signs of them. When precedence in time is varied, steep delay of reinforcement gradients are often found, attesting to the importance of temporal contiguity (*4*). When logical necessity is reduced to contingent probability and then varied, the strength of conditioning varies with it (*5*). These findings suggest that the nature of learning has evolved to keep animals en rapport with the causal structure of their environments.

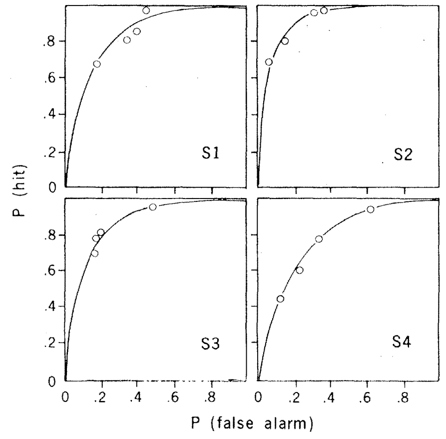

Fig. 1. Receiver operating characteristics for the four subjects. P(hit) is the proportion of times that a pigeon indicated "yes" when a stimulus-change was response-contingent; P (false alarm) is the proportion of times that a pigeon indicated "yes" when a stimulus change was computer-contingent.

Yet the putative existence of superstitions suggests a gross breakdown in an organism's ability to detect causality. In the present experiment, I asked whether animals are as badly off as Skinner's analysis suggests, or whether the ritualistic behaviors he and Staddon and Simmelhag noted can all be ascribed to "hypothesis testing" elicited by food. The distinction hinges on the ability to separate the organism's sensitivity to behavior-environment correlations from its willingness to act on those discriminations. For humans, when the stakes are high (for example, rain after a lengthy drought) or the response cost low (for example, carrying a charm) superstitions are understandable, often having as much the character of "playing a long shot" as of being duped by a coincidence of nature. Do animals share a similar relativity of judgment about causality?

Let us consider the following experiment: A pigeon pecks a central white disk, with each peck having a probability of .05 of darkening that disk and illuminating two side disks. While the pigeon is pecking the center disk, a computer is generating "pseudo-pecks" at the same rate (*6*), with each pseudo-peck also having a probability of .05 of extinguishing the center disks and lighting the side disks. It is the pigeon's task to decide whether the stimulus change was caused by his last peck on the center disk, or whether it was independent of that peck. He votes by responding on the side disks, right for "response dependent" and left for "response independent." A correct decision is rewarded with food, an incorrect one punished with a brief time out. The data from such signal-detection tasks are typically arrayed in a two-by-two conditional probability matrix, with the probability of a signal (response-dependent stimulus change) along one margin and the probability of voting "signal" along the other (*7*). We focus on two cells of the matrix: the probability of saying "signal" given that a signal occurred (hits) and the probability of saying "signal" given that no signal occurred (false alarms).

When operating near the limit of discriminability, an organism may increase its percentage of hits only by also increasing its percentage of false alarms; conversely, when more conservatively motivated, the organism may decrease its percentage of false alarms but only with a corresponding decrease in the percentage of hits. These relations are portrayed by plotting one percentage against the other along a curve called a receiver operating characteristic, or ROC curve (see Fig. 1). The most discriminable signals and the most sensitive observers generate curves that lie close to the upper left corner, where the probability of a hit, P (hit), = 1 and P (false alarm) = 0. Discriminability may be measured by the percentage of the unit area under the curve; the closer the curve lies to the upper left corner, the more discriminable the signal and the more acute the observer. Motivational variables are measured by the distance of the data from the negative diagonal; as the subject becomes more highly motivated to indicate "yes," the data follow the curve toward the upper right corner (*8*).

Four pigeons were trained in this paradigm and were given unequal amounts of food for correct responses to the left

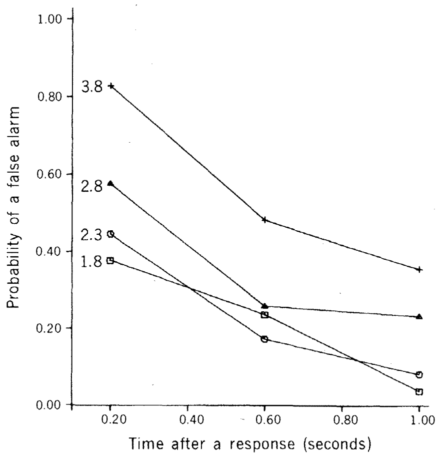

Fig. 2. The probability of a false alarm as a function of the time elapsing between a response and a subsequent noncontingent stimulus change, averaged over subjects. The parameter is the amount of reward (eating time, in seconds) for a correct "yes" response.

0036-8075/78/0106-0088$01.00/0 Copyright © 1977 AAAS

Table 1. Detectability (A) and bias (B) indices for each subject at each reinforcement condition. Values for R_Y and R_N indicate the amount of food contingent on "yes" and "no" responses. Positive values of B indicate the attribution of causality to an internal locus. Values of A estimating the percentage correct that would be obtained in a bias-free paradigm were relatively constant at about 85 percent.

R_Y	R_N	B				A
		1	2	3	4	
3.8	1.8	+95	+86	+80	+65	86
2.8	2.8	+33	+75	+ 1	+12	86
2.3	3.3	+18	− 7	−12	−16	85
1.8	3.8	−20	−50	−20	−39	84

and right disks. Each made approximately 1500 choices at each of four payoff ratios, and data from the last 500 choices were used in the subsequent calculations. Figure 1 shows that the pigeons' behavior is well captured by an ROC analysis: As the payoff for indicating "caused" increased from 1.8 seconds of eating time to 3.8 seconds, the probability of saying "caused" increased for all birds, resulting in both a higher hit rate and a higher false-alarm rate. Under all conditions sensitivity to the contingencies was high, with the birds being correct on about 80 percent of the trials—about the same percentage as that scored by humans watching them. Most of the false alarms occurred when a response-independent stimulus change occurred within 0.5 (subject 2) to 2.0 (subject 4) seconds of a response. In Fig. 2, false alarms are averaged over subjects and plotted against delay. The gradients are steep, approximately parallel, and elevated as a function of the amount of reinforcement for a "caused" response. These brief temporal limits for the attribution of causality are consistent with other estimates, even though organisms may hold some stimuli, such as novel foods, responsible for malaise occurring hours later (*4*). Whether the limits are relatively fixed, with a few notable exceptions, or whether they are a continuous function of both the ponderousness of the response and the salience of the resulting stimulus change, are questions that may be addressed with the present experimental paradigm (*9*).

Within the above limits of sensitivity, the animals maximized their reinforcement by biasing their responses as a function of the relative payoff. Table 1 shows the bias with a measure (B) that ranges from +100 (all signals are perceived to be caused) to −100 (all signals are perceived to be independent of behavior). In the present experiment, B ranged from +95 to −50, indicating a large responsivity in the allocation of votes. The measure of sensitivity, A, remained constant over experimental conditions (right column of Table 1), and varied slightly between subjects.

Although I chose a decision-theoretic framework for the present experiments, the data are relevant to both more behavioristic analyses and more cognitive phenomena. The data may be replotted as the log ratio of "caused" to "uncaused" decisions versus the log ratio of amount of reward (seconds of access per session, the product of hopper time, and number of hopper operations) for the two decisions. Figure 3 shows that the data fall along a straight line, a finding consistent with the effects of motivational variables on choice behavior in traditional animal learning paradigms, where there is no ambiguity or uncertainty about the signal (*10*). The present research is also related to Rotter's work on the attribution of causality to internal versus external loci by humans (*11*).

Hume (*12*) held that knowledge of cause and effect arose from experience, not reason, and in animals other than humans was augmented by instinct. Reinforcement was a factor neither in reasoning nor in the perception of causation, upon which that reasoning was based, but biased judgment subsequently, "after all circumstances and relations are laid before us" (*12*, p. 294). By these words Hume is associated with the school of "threshold psychophysicists," who hold that motivational variables affect a judgmental stage subsequent to perception (*13*). Unlike many of Hume's speculations, this may be testable. If pigeons first discriminate either "causality" or its absence and then bias their choice as a function of the payoff, the data in Fig. 1 should fall along two straight lines, one originating from (0, 0) and intersecting with one originating from (1, 1). If, however, there is a continuous interaction between motivation and perception—then the data should fall along one of a variety of smooth curves. The present data seem to favor the latter alternative, although a statistically adequate decision between the two requires a large data base (*14*).

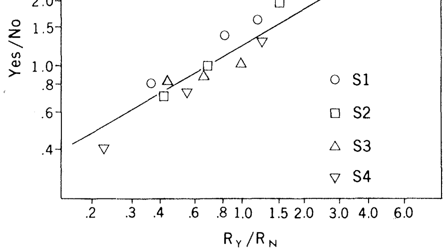

Fig. 3. The ratio of "yes" to "no" responses as a function of the ratio of reinforcements (seconds eating time per session) for those decisions, plotted on logarithmically spaced axes. The positive intercept of the regression line ($\log y = .58 \log x + .075$) indicates a slight bias toward internal locus of control.

These data indicate that pigeons are accurate in their perception of contingencies between behavior and environment, and that the discrimination in the present experiments was probably based on the delay between a response and its effects. Superstitions arise not from failures of discrimination but from biases due to differential reinforcement and, perhaps, instinctive predispositions. Viewed in this light, flexible criteria for attribution of causality, which may often promote superstitious behaviors, are more adaptive than fixed criteria that cannot shift to optimize expected payoff.

PETER R. KILLEEN
Department of Psychology,
Arizona State University, Tempe 85281

References and Notes

1. B. F. Skinner, *J. Exp. Psychol.* **38**, 168 (1948).
2. J. E. R. Staddon and V. Simmelhag, *Psychol. Rev.* **78**, 3 (1971).
3. E. Hearst and H. M. Jenkins, *Sign-tracking: The Stimulus-Reinforcer Relation and Directed Action* (Psychonomic Society, Austin, Tex., 1975); B. Schwartz and E. Gamzu, in *Handbook of Operant Behavior*, W. K. Honig and J. E. R. Staddon, Eds. (Prentice-Hall, Englewood Cliffs, N.J., 1977), pp. 53–97.
4. C. N. Uhl, *Learn. Motiv.* **5**, 511 (1974); G. R. Grice, *J. Exp. Psychol.* **38**, 1 (1948). For the traditional use of a 1.5-second "change-over delay," see R. J. Herrnstein, *J. Exp. Anal. Behav.* **4**, 267 (1961); A. C. Catania, in *Operant Behavior: Areas of Research and Application*, W. K. Honig, Ed. (Appleton, New York, 1966), pp. 213–270. H. Rachlin and W. M. Baum, *J. Exp. Anal. Behav.* **18**, 231 (1972). Shallow delay of reinforcement gradients have been reported by J. Garcia and R. A. Koelling, *Psychon. Sci.* **4**, 123 (1966); P. Rozin and J. W. Kalat, *Psychol. Rev.* **78**, 459 (1971); R. V. Krane and A. R. Wagner, *J. Comp. Physiol. Psychol.* **88**, 882 (1975).
5. R. A. Rescorla, *Psychol. Rev.* **74**, 71 (1967); J. Gibbon, R. Berryman, R. L. Thompson, *J. Exp. Anal. Behav.* **21**, 585 (1974); W. N. Schoenfeld and J. Farmer, in *The Theory of Reinforcement Schedules*, W. N. Schoenfeld, Ed. (Appleton-Century-Crofts, New York, 1970), pp. 215–245.
6. This rate was determined as an exponentially weighted moving average of the pigeon's peck rate, with a rate constant of 0.1. The average was updated with every response and every second without a response.
7. D. M. Green and J. A. Swets, *Signal Detection Theory and Psychophysics* (Wiley, New York, 1966); J. P. Egan, *Signal Detection Theory and ROC Analysis* (Academic, New York, 1975); J. Levitt, in *Human Communication: A Unified View*, E. E. David, Jr., and P. B. Denes, Eds. (McGraw-Hill, New York, 1972), pp. 114–174;

J. A. Nevin, K. Olson, C. Mandell, P. Yarensky, *J. Exp. Anal. Behav.*, **24**, 355 (1975).

8. The nonparametric measures of sensitivity (A) and bias (B) are from J. B. Grier, *Psychol. Bull.* **75**, 424 (1971); they are independent of the nature of the underlying hypothetical distributions. The smooth curves through Fig. 1 are isosensitivity contours based on Gaussian distributions of signal and noise. Data suggest that exponential distributions may be more appropriate.
9. K. A. Lattal [*J. Exp. Anal. Behav.* **23**, 241 (1975)] employed a similar technique and reviews the current experimental work on the discrimination of reinforcement contingencies.
10. W. M. Baum, *ibid.* **22**, 231 (1974); for an analysis of human signal detection performance using similar coordinates, see J. A. Nevin, *ibid.* **12**, 475 (1969); for the application of signal detection analysis to visual generalization gradients, see D. S. Blough, *Science* **158**, 940 (1967).
11. H. H. Kelley, *Attribution in Social Interaction* (General Learning Press, New York, 1971); R. de Charms, *Personal Causation* (Academic, New York, 1968); E. J. Phares, *Locus of Control in Personality* (General Learning Press, Morristown, N.J., 1976).
12. D. Hume, *Enquiries Concerning Human Understanding and Concerning the Principles of Morals* (Clarendon, Oxford, 1777/1975); see also A. Michotte, *The Perception of Causality* (Basic, New York, 1963); J. E. R. Staddon, *Behaviorism* **1**, 25 (1973); T. J. Testa, *Psychol. Rev.* **81**, 491 (1974); H. M. Jenkin and C. W. Ward, *Psychol. Monogr.* **79** (No. 594), 1 (1965).
13. R. D. Luce, *Psychol. Rev.* **70**, 61 (1963).
14. The decision is made more difficult by the existence of plausible threshold theories that invoke multiple thresholds [for example, D. H. Krantz, *Psychol. Rev.* **76**, 308 (1969)] and are therefore more flexible in fitting theory to data. Conversely, the conventional assumption of Gaussian distributions of signal and noise is robust, and it is difficult to discriminate between them and other continuous distributions that might have given rise to the data [I. G. Abrahamson, H. Levitt, L. Landgraf, *J. Acoust. Soc. Am.* **42**, 1195 (1967)]. Such decisions will probably be made on the basis of convergent experimental paradigms, including but not limited to yes-no ROC analysis.
15. This work was supported by grants BMS 74-23566 and BNS 76-24534 from the National Science Foundation to Arizona State University. I thank S. Osborne for his help in conducting these experiments.

7 March 1977; revised 15 August 1977

Phenobarbital: Effects of Long-Term Administration on Behavior and Brain of Artificially Reared Rats

Abstract. *Two doses of phenobarbital were given daily for 2 weeks to infant rats fed by intragastric cannulas. The larger dose (60 milligrams per kilogram of body weight) resulted in decreased spontaneous activity and increased responses to novel stimuli. The smaller dose (15 milligrams per kilogram) resulted in increased spontaneous activity and also an increase of responses to novel stimuli. The larger dose produced a 12 percent reduction in brain growth, while the smaller dose was associated with a 3 percent reduction in brain growth.*

The rapid growth the brain undergoes early in its development causes it to be vulnerable to exogenous insults (*1*). Although a number of centrally acting drugs are commonly used for treating human infants, there is little experimental information regarding the effects of such drugs on brain or behavioral development.

Exposure to drugs during infancy in laboratory animals has been shown to result in behavior and brain alterations later in life (*2, 3*). Drug-induced undernutrition, however, often accompanies neonatal drug treatments, and since early undernutrition itself causes behavioral changes (*4*), it is difficult to interpret these studies. The purpose of the study described here was to examine directly the developmental effects of phenobarbital administered chronically during infancy to artificially reared pups.

Male Wistar rats (4 to 5 days old) were selected for body weight within a range of 6 to 11 g. The animals ($N = 46$) were lightly anesthetized with ether, and intragastric cannulas were permanently implanted by means of a technique similar to that refined by Hall (*5*). Once the cannulas were implanted, the pups were placed in circular plastic dessert cups (12 cm in diameter and 8 cm deep), which floated in a warm water bath (40°C). The cannulas were connected to syringes filled with milk formula (*5*) and mounted on an infusion pump (*5*). The room that housed the water bath was on a 12-hour reverse light cycle. Every morning the animals were disconnected from the pumps, the syringes were washed and refilled, and the cannulas were flushed with saline. Each plastic washer securing the cannula at the stomach was checked and loosened whenever necessary to accommodate the animal's growth.

On day 5, the pups were assigned by weight to control and experimental groups, and for the next 13 days the experimental group ($N = 27$) received daily subcutaneous injections of phenobarbital, while the control group ($N = 19$) received subcutaneous injections of the vehicle. In the first of two studies, the experimental group ($N = 15$) received "high" doses (60 mg/kg) of phenobarbital; in the second study, the experimental group ($N = 12$) received "low" doses (15 mg/kg) of phenobarbital. Immediately after the injections, the animals were placed back in their cups and returned to the water bath. All the cannulas were reconnected to the infusion pump at least 1 hour after the last injection and the pump speed was adjusted to infuse approximately 0.5 ml more milk formula than the previous day.

At 18 or 19 days of age, all the animals were tested in a circular open field enclosure (60 cm in diameter) with the floor divided into 10-cm squares. Each animal was gently placed in the center of the field and immediately covered with a plastic cup (12 cm in diameter, 8 cm deep). After 10 seconds, this cup was lifted and the number of locomotions—that is, squares centered (both forepaws placed into an adjacent square), and the time spent rearing (both forepaws off the floor)—was recorded by two trained observers, one of whom was unaware of the animal's group assignment. The first 5 minutes of the test were conducted under dim illumination (one overhead filtered fluorescent light) with a masking white noise present (approximately 45 db). After 5 minutes, intermittent flashing lights (one overhead 75-watt bulb) and noises (two electromagnetic relays clicking on and off quickly, resembling teeth chattering) were present for an additional minute.

After the last animal was tested in the open field, all the animals were decapitated and their brains were quickly removed. The cerebellum was separated and weighed, and the remaining brain sample was divided by a transcollicular cut into cerebrum and brainstem and weighed. At the time of decapitation,

Table 1. Effects of phenobarbital on brain growth. Data are expressed as means ± standard error. Weights are expressed in grams.

Tissue weighed	Low doses of phenobarbital (15 mg/kg)			High doses of phenobarbital (60 mg/kg)		
	Control ($N = 7$)	Experi mental ($N = 12$)	Difference from control (%)	Control ($N = 12$)	Experimental ($N = 15$)	Difference from control (%)
Body	32.1 ± 1.7	30.7 ± 2.7	−4	34.2 ± 2.9	33.8 ± 2.4	−1
Brain*	1.34 ± 0.05	1.101 ± 0.07	−3	1.213 ± 0.05	1.077 ± 0.06†	−12
Cerebrum	0.831 ± 0.03	0.820 ± 0.05	−1	0.907 ± 0.04	0.808 ± 0.04†	−11
Cerebellum	0.146 ± 0.01	0.144 ± 0.01	−1	0.160 ± 0.01	0.147 ± 0.01†	−9

*The brainstem is included. †$P < .01$, t-test.