

# Low false-alarm rates in signal detection by marine mammals

Ronald J. Schusterman

Department of Psychology, California State University, Hayward, California 94542  
(Received 5 September 1973)

Psychophysicists attempting to obtain acoustic threshold estimates in marine mammals have been relatively unaware of setting up reinforcement contingencies which have resulted in their animals using a Neyman-Pearson objective as their response criterion. Low false-alarm rates strengthen the investigator's idea that he has acquired good stimulus control over the animal's behavior but may also lead to underestimates of sensitivity. Where possible, psychophysical procedures with animals should be amenable to analysis of response bias independent of sensitivity.

Subject Classification: 80.45.

The purpose of this paper is to point out that experimentalists in animal psychophysics, particularly those working on acoustic "thresholds" in marine mammals, have been relatively unaware of setting up reinforcement contingencies which have resulted in their animals having a strong response bias. Another point to be made is that such bias sometimes leads to arbitrary threshold estimates which may underestimate the sensitivity of the animal. It is not my intention to discuss the relationship between signal detectability and the data of classical threshold measurement procedures nor to review aspects of signal detection theory (SDT) as it applies to animal psychophysics in general or to selected species other than marine mammals. The former has been elegantly accomplished by Treisman and Watts (1966) and the latter is a relatively new discipline in which some of the more important work is being done by J. S. Nevin and M. Terman (Nevin, 1970; Terman and Terman, 1972).

In several recent studies dealing with visual acuity thresholds and sound detection thresholds, sea lions (*Zalophus californianus*) were trained to emit a vocalization underwater if a particular visual or acoustic signal was seen or heard and to inhibit such vocalization for a given period when such signals were not detected (Schusterman and Balliet, 1970; Schusterman, Balliet, and Nixon, 1972). In the parlance of signal detection theory vocalization was considered the "yes" report and the lack of vocalization during the choice period was considered the "no" report. Both the hearing and the vision experiments used modified psychophysical methods such that stimulus magnitudes or signal strengths relative to the background noise were varied within blocks of trials so that both tasks were amenable to analyzing response bias independently of sensitivity. Thus, although the data from these experiments were presented in the classical psychophysical framework in which ogival psychometric functions were shown, they could also be analyzed in terms of probabilities of correct "yes" responses or "hits" ( $Y/sn$ ) and incorrect "yes" responses or "false alarms" ( $Y/n$ ). In the visual acuity experiments a very definite response bias toward reporting "no" was noted. Even though the likelihood of a sea lion vocalizing to a comparison target (hits) decreased directly as a function of decreasing resolution of visual detail, the probability of false alarms remained relatively constant (Schusterman, 1972). Inspection of several sound detection "thresholds" in *Zalophus* re-

vealed the same relationship between hits and false alarms as the resolution of  $n$  and  $sn$  became increasingly difficult (Schusterman, Balliet, and Nixon, 1972).

Since response criterion is reflected in the rate of false alarms obviously these results (in which a conditioned vocalization was used as an indicator response) from visual and auditory detection tasks suggest that the response criterion of *Zalophus* was not only quite high or conservative but also relatively invariant even with unfavorable signal to noise ratios. The experiments with *Zalophus* were repeated using signals at three different levels of magnitude. One animal served in a visual acuity task and the other animal's task was to report the presence of a 48-kHz tone. Figure 1 shows the results of these experiments in terms of the relation of hits to false alarms.

A check of the literature as well as correspondence with several other investigators revealed that, in those cases where the psychophysical procedures permitted, separation of response criterion and auditory sensitivity invariably resulted in marine mammals maintaining a low and constant false-alarm rate. The marine mammals included two pinniped species of the family *Phocidae*, the harbor seal (*Phoca vitulina*) and the harp seal (*Pagophilus groenlandicus*), and two small odontocete cetacean species, one from the family *Platanistidae*, the Amazon river porpoise (*Inia geoffrensis*) and the other from the family *Delphinidae*, the Atlantic bottlenosed porpoise (*Tursiops truncatus*). The tasks included passive aerial and underwater hearing in the seals (Möhl, 1968; Terhune and Ronald, 1972) and active sonar by the porpoises (Penner and Murchison, 1970; Murchison and Penner, in preparation). Presence or absence of signals was indicated by touching the nose (seals) or rostrum (porpoises) to one of two manipulanda and receiving a fish reward for "correct" responses and either nothing (porpoises) or an air-blast punishment as a consequence for "incorrect" responses.

In each of the above described experiments, including those with *Zalophus*, the investigators were attempting to estimate a sound detection threshold of one sort or another. In all experiments the probability of signal presentation was approximately 0.50.

Data from these experiments have been reanalyzed and Figs. 2 and 3 show the relation between hits and

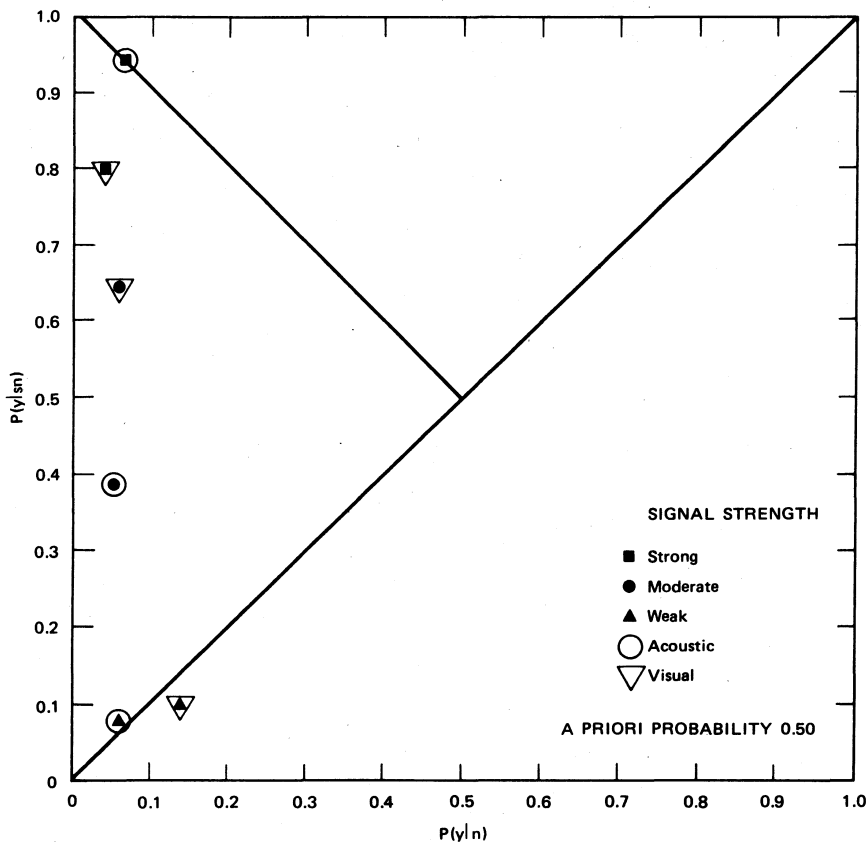


FIG. 1. Relation of hits to false alarms for acoustic and visual psychometric functions in the California sea lion.

false alarms for several different marine mammals under a variety of psychophysical tasks and reinforcement contingencies. The data points in Figs. 1–3 lie between the upper left-hand corner (perfect detection) and the major diagonal (chance detection). Points lying along the minor diagonal (line drawn from the upper left-hand corner to the major diagonal) would represent no response bias, since the two possible types of “error” (false alarms and misses) would be equally likely.

Since response bias is defined as a tendency to make one type of error rather than another, a constant bias can be seen (almost independent of detectability) in all animals under all conditions as a constant departure of data points away from the minor diagonal. In all instances the data points gradually approach the major diagonal. This reflects the fact that the animal’s task of resolving  $n$  and  $sn$  became more and more difficult. However, in general the data points are almost parallel to the ordinate, and so represent a tendency to keep the proportion of false alarms constant regardless of the degree of detectability. For example, when the sea lion was attempting to detect a 16-kHz tone it consistently maintained a false-alarm rate of  $< 0.10$  even when its hit rate was between 0.20 and 0.30. In the case of the river porpoise we have an even more dramatic illustration of this phenomenon. In attempting to detect copper wires (with diameters ranging from 0.6 to 2.6 mm) by means of echolocation, the river porpoise consistently maintained a false alarm rate of  $< 0.05$  even when its hit rate was  $< 0.20$ . Two male bottlenosed porpoises (Fig. 3) were trained on an open-water echolocation or sonar task in which they were to report the presence or ab-

sence (by activating a spatially arranged yes/no manipulanda) of a solid 2.54-cm sphere (task 1) and a 7.6-cm water filled sphere (task 2) at various distances from the animal. The data in Fig. 3 show that when the probability of detections or hits fell below 0.40, false-alarm rates did increase, but even when the animals’ performances were near chance there was still a relatively strong response bias in favor of reporting the absence of the target.

It would appear then that at an *a priori* signal presentation probability of 0.50, the strategy of seals, sea lions, and porpoises in a variety of psychophysical tasks (including echolocation for the porpoises), using a variety of indicator response, is to maximize the proportion of hits while holding the proportion of false alarms at a constant low value. Such a criterion for responding “yes” is quite conservative. In human psychophysics Treisman and Watts (1967) point out that using such procedures as the method of constant stimuli and the method of limits, this criterion (originally proposed by Neyman and Pearson in 1933) is defined as the value on the decision axis which restricts the false alarm rate to some acceptable limiting value. In classical human psychophysics trained subjects were used in order to avoid “error” and the subjects’ strict adherence in maintaining low and constant false alarm rates was the decisive factor in determining whether the person remained as a subject in the experiment. Modern psychophysicists working with humans frequently incorporate procedures comparable with SDT, thus separating the person’s decision criterion from the person’s sensitivity. For

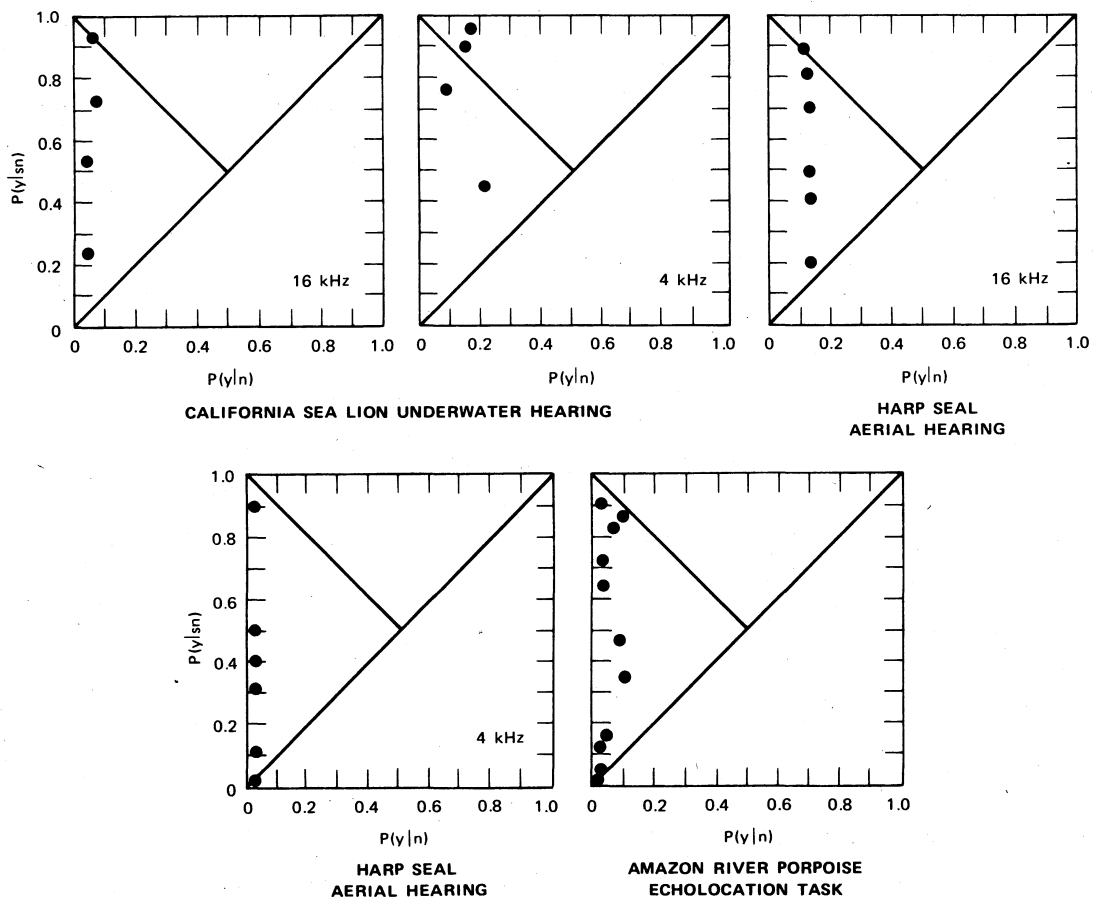


FIG. 2. Relation of hits to false alarms for a variety of acoustic detection tasks in several different marine mammals.

example, even a false-alarm rate as high as 0.30 with a hit rate of 0.70 would be perfectly acceptable as an estimate of acute signal detectability or sensitivity along some physical dimension. However, the determinations of various sound detection and visual discrimination "thresholds" in marine mammals were all accomplished by traditional psychophysical procedures and analysis. The question arises as to why such a group of mammals in such diverse circumstances used the Neyman-Pearson objective as their response criterion. Was it the "nature" of the beasts? Hardly likely, since animal psychophysicists have traditionally limited false-alarm rates to a figure of 0.10 or less by arranging a specific set of reinforcement contingencies including severe shocks, long "time-outs," etc., or elimination of the animal from the experiment if the response criterion were such that some acceptable limiting level of false alarms was exceeded (see Gourevitch, 1970). However, one searches in vain when attempting to ferret out some of the reinforcement contingencies that were used to train these marine mammals to use a Neyman-Pearson criterion in obtaining a variety of signal detection "thresholds." One clue comes from the work of Jacobs and Hall (1971) on determination of hearing thresholds in the freshwater porpoise, *Imia geoffrensis*. They state that false alarms occurred only during the initial stages of training and were "extinguished with 'time-outs.'" Personal communication with C. Scott Johnson regarding his pioneer-

ing effort on hearing thresholds in the bottlenosed porpoise (1966) using the "staircase" method confirms the use of time-out as an aversive event contingent on one type of error—false alarms and not the other type of error—misses. Personal communication with Ralph Penner and my own experience in acquiring stimulus control over vocalizations emitted by sea lions suggests that many investigators doing psychophysical experiments with nonhuman animals, particularly with marine mammals, may have inadvertently trained their animals to have a high or conservative response criterion. Frequently this has been accomplished by giving the animal

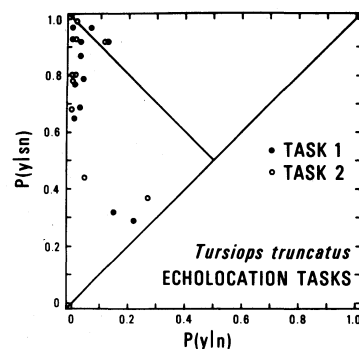


FIG. 3. Relation of hits to false alarms for two echolocation tasks in the bottlenosed porpoise, *Tursiops truncatus*.

a "warm up" series of trials with termination of the experimental session as a consequence of "too many false alarms." Thus the animal rapidly learns to bias its report in favor of "no" since these responses (if not given in abundance to a strong signal) are not likely to result in a 24 h period of time out. A low rate of false alarms by the animal usually strengthens the researcher's notion that he or she has acquired good stimulus control over the animal's behavior.

Typically these experimentors use an arbitrary 50% correct detection threshold or a 75% correct response threshold and with false-alarm rates consistently <0.20 it is likely that sensitivity of the animal is underestimated. For example, Penner and Murchison (1970) found that the 50% threshold of detection for a copper wire by echolocation in *Inia geoffrensis* was 1.12 mm. However, if one uses hits and false alarms to estimate  $d'$  as an index of sensitivity then a  $d'$  value of 1.02 would be associated with a copper-wire diameter of 0.8 mm. Recent findings by Seaton (1973) support the notion that thresholds are higher when animals are trained to maintain a low false-alarm rate. He found that critical ratio measures in the chinchilla are somewhat lower when the animals are trained to have higher false-alarm rates than when the animals had been trained to keep the false-alarm rate low.

Two problems have emerged from this discussion. First, the need on the part of the bioacoustician for maintaining stimulus control over his animal has frequently created a "Neyman-Pearson animal." In the case of marine mammal psychophysicists this phenomenon has often occurred inadvertently. Secondly, without consideration of response bias, extremely low false-alarm rates sometimes lead to underestimates of an animal's sensitivity. Both problems can at least be minimized if strong signal "probe" trials are used in order to determine degree of stimulus control when using classical threshold measurement procedures and by varying reinforcement contingencies in an attempt to generate iso-sensitivity curves.

#### ACKNOWLEDGMENT

Preparation of this paper was supported by ONR contract N00014-73-C-0186. I thank Earl Murchison and Ralph Penner of the Naval Undersea Research and De-

velopment Center, Kailua, Hawaii, for allowing me to reanalyze some of their as yet unpublished data. Thanks are also due to Dr. Frank Beach III for his helpful comments on the original manuscript.

- Gourevitch, G. (1970). "Detectability of Tones in Quiet and Noise by Rats and Monkeys," in *Animals Psychophysics: The Design and Conduct of Sensory Experiments*, edited by W. C. Stebbins, (Appleton-Century-Crofts, New York), pp. 67-98.
- Jacobs, D. W., and Hall, J. D. (1972). "Auditory Thresholds of a Fresh Water Dolphin, *Inia geoffrensis* Blainville," *J. Acoust. Soc. Am.* **51**, 530-533.
- Möhl, B. (1968). "Auditory Sensitivity of the Common Seal in Air and Water," *J. Aud. Res.* **8**, 27-38.
- Murchison, A. E., and Penner, R. H. (in preparation). "Open Water Echolocation in the Bottlenose Dolphin (*Tursiops truncatus*): Metallic Sphere Detection Thresholds as a Function of Distance."
- Nevin, J. A. (1970). "On Differential Stimulation and Differential Reinforcement," in *Animal Psychophysics: The Design and Conduct of Sensory Experiments*, W. C. Stebbins, Ed. (Appleton-Century-Crofts, New York), pp. 401-424.
- Penner, R. H., and Murchison, A. E. (1970). "Experimentally Demonstrated Echolocation in the Amazon River Porpoise, *Inia geoffrensis* Blainville," *Naval Undersea Res. Develop. Ctr.* TP 187, 26 pp.
- Schusterman, R. J. (1972) "Visual Acuity in Pinnipeds," in *Behavior of Marine Animals: Vol. II*, edited by A. E. Winn and B. L. Olla (Plenum, New York), pp. 469-492.
- Schusterman, R. J., and Balliet, R. F. (1970). "Conditioned Vocalization As a Technique for Determining Visual Acuity Thresholds in the Sea Lion," *Sci.* **169**, 498-501.
- Schusterman, R. J., Balliet, R. F., and Nixon, J. (1972). "Underwater Audiogram of the California Sea Lion by the Conditioned Vocalization Technique," *J. Exp. Anal. Behav.* **17**, 339-350.
- Seaton, W. H. (1973). "A Comparison of a Direct and Indirect Measure of Critical Bands in the Monaural Chinchilla," Master's Thesis, University of Illinois, Urbana-Champaign, Illinois, 59 pp.
- Terhune, J. M., and Ronald, K. (1972). "The Harp Seal, *Pagophilus groenlandicus* (Erxleben, 1777). III. The Underwater Audiogram," *Cand. J. Zool.* **50**, 565-569.
- Terman, M., and Terman, J. S. (1972). "Concurrent Variation of Response Bias and Sensitivity in an Operant Psychological Test," *Percept. and Psychophysics* **1**, 428-432.
- Treisman, M., and Watts, T. R. (1966). "Relation Between Signal Detectability Theory and the Traditional Procedures for Measuring Sensory Thresholds: Estimating  $d'$  from Results Given by the Method of Constant Stimuli," *Psychol. Bull.* **66**, 438-454.