MTR190014

MITRE TECHNICAL REPORT

The Watching—A Review of the Vigilance Research

Dr. Valerie Gawron

January 2019

McLean, VA

Center for Advanced Aviation System Development

Table of Contents

The Watching—A Review of the Vigilance Research 1
References 1

List of Figures

Figure 1. Two States of the World 2
Figure 2. λ as a Function of Neural Activity 3
Figure 3. ROC Curve 4

The Watching—A Review of the Vigilance Research

During WWII an unusual performance phenomenon was noted among radar operators: specifically, the probability of detecting a signal decreased dramatically over time-on-watch. This decrement was greatest when:

- The signal duration was short.
- The probability of a signal was low.
- The signal intensity was low.
- The signal was simple rather than complex

A number of theories were developed to explain these findings. The most well known of these is the Theory of Signal Detection (TSD). TSD was introduced over forty years ago (Swets, Tanner, & Birdsall, 1961). Since that time, it has been one of the most widely used theories of human behavior. There are two main reasons for the theory's popularity. First, TSD is a model of human detection in noise and, therefore, applicable to many real-world situations. Second, TSD is a normative model and thus describes how the human should act. As such, TSD is useful in developing criteria of optimum performance against which actual performance can be compared. From this comparison, suggestions can be made for improving a system.

TSD is a rather simple and straightforward theory. It proposes that there are two states in the real world: noise alone or a signal in that noise. Each of these states gives rise to neural activity. The activity from each of these states is the same at certain intensities, and the probability that a given intensity is due to either state is hypothesized to describe two overlapping normal curves (see Figure 1). In the area of overlap, detection becomes not only a sensory problem but a decision-making one, as well (Swets, 1964; Swets, Tanner, & Birdsall, 1961).

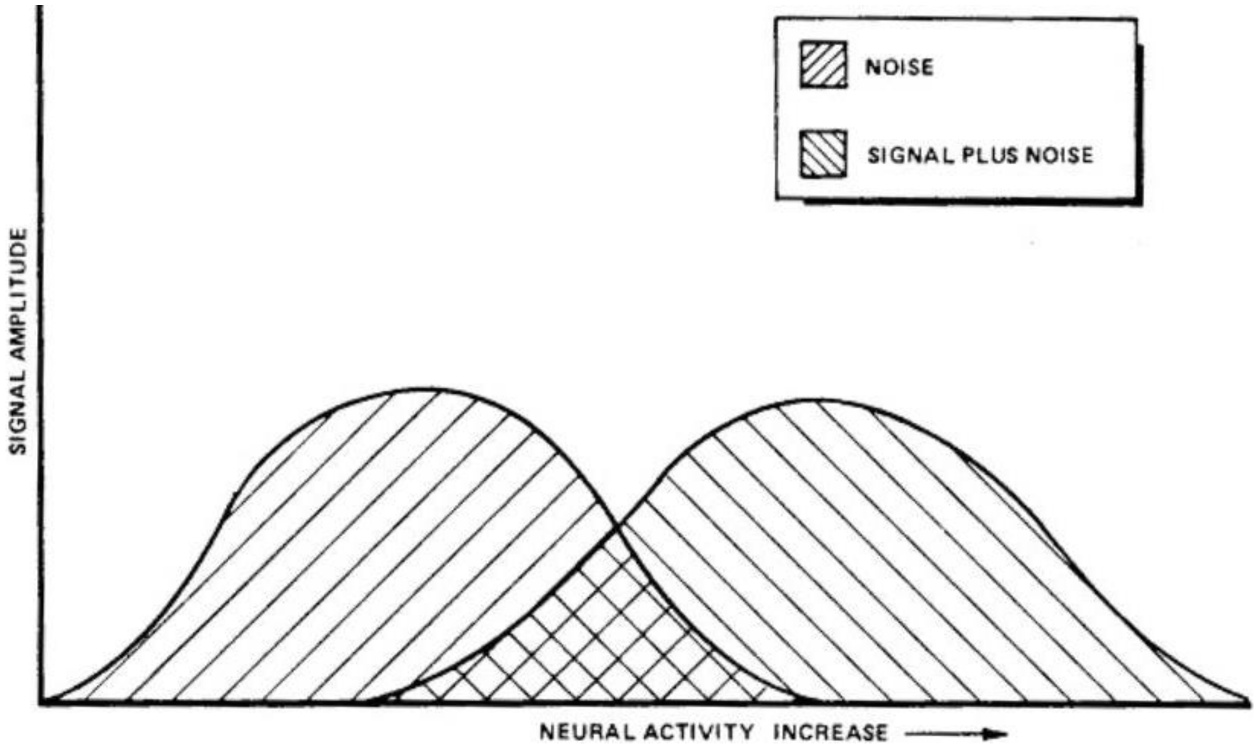

Figure 1. Two States of the World

The decision-making aspect of the detection process is described by the TSD concept of the likelihood ratio (λ). This is the ratio of the probability that a given intensity of neural activity resulted from the occurrence of a signal to the probability that the intensity resulted from noise alone. Above a certain value of λ, the operator will decide that the signal did occur and below this value he or she will decide that he neural activity is due to noise alone. The value of λ that serves as a yes-no decision demarcator is called β (see Figure 2).

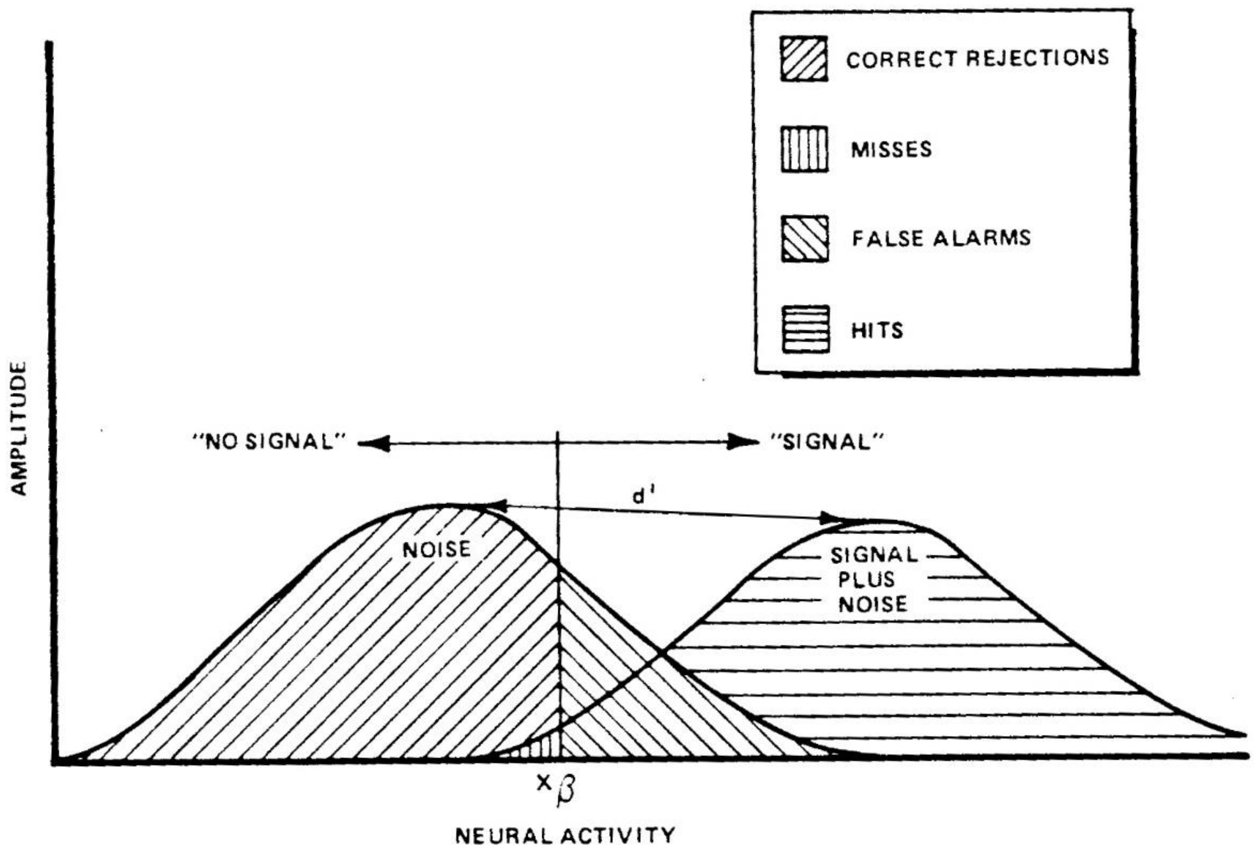

Figure 2. λ as a Function of Neural Activity

Given the decision scheme proposed by TSD, there are four possible outcomes (see Figure 2). First, an operator may decide that a signal occurred when one actually did—a hit. Second, he or she may decide that no signal occurred when in fact none did—a correct rejection. These two outcomes are clearly desirable. However, there are two other outcomes that are not desirable. An operator may decide that a signal occurred when one did not—a false alarm. Finally, he or she may decide that no signal occurred when in reality, one did—a miss.

The probability of each of these outcomes is described by the areas of the noise and signal-plus-noise curves allotted to each. As Figure 2 clearly demonstrates, the frequency of occurrence of these outcomes is determined by □. A high or conservative □ is associated with a large number of missed signals but a small number of false alarms. On the other hand, a low or risky □ will result in a large number of hits but also a decrease in correct rejections.

In both cases there is a tradeoff between desirable and undesirable outcomes. It is at this point that the normative aspect of TSD becomes useful since it can describe an optimum value. This value is defined in terms of the probabilities of the occurrence of noise (P(N)) or signal (P(S)), the values of correct rejections (V(CR)) and hits (V(H)) and finally, the costs of false alarms (C(FA)) and misses (C(M)). The equation defining the optimum □ value is presented below.

Optimum □ = P(N)   V(CR) + C(FA)

P(S) x   V(H) + C(M)

The operator's actual performance can be compared against this optimum and suggestions made for the improvement of performance. This comparison is made by manipulating the operator's perception of values and costs of the decision outcomes and plotting the resulting probabilities of hits (P(H)) versus false alarms (P(FA)). This function is called a receiver operating characteristic (ROC) curve. The curve is especially useful since the distance of the curve to the positive diagonal of the graph yields a measure of the operator's sensitivity irrespective of decisional manipulations. This measure is termed d' (see Figure 3).

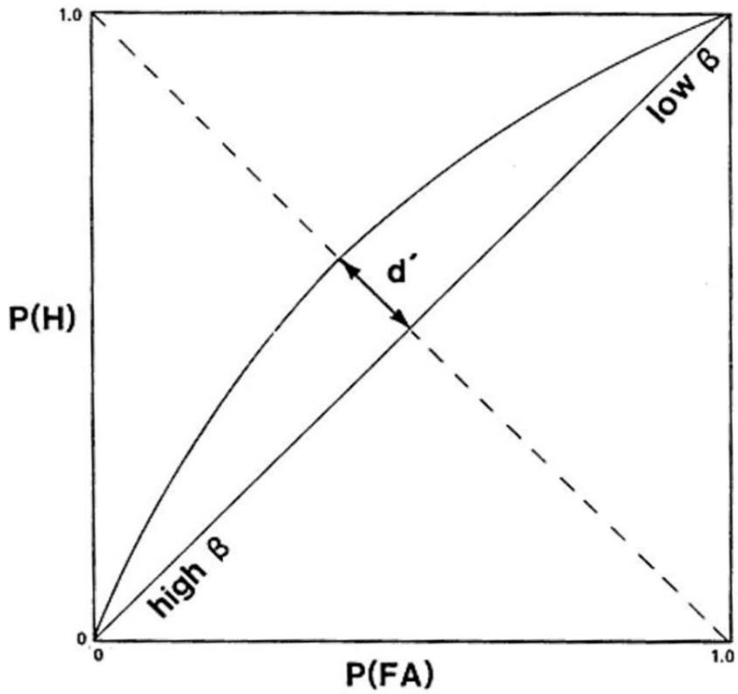

Figure 3. ROC Curve

d' also corresponds to the distance between the noise and signal-plus-noise curves (see Figure 3). As such, it is a measure of the difficulty of the task. As d' increases, the amount of overlap between the two curves decreases and the detection task becomes easier. Similarly, as d' decreases, the amount of overlap increases, making the detection task more difficult. Goodenough and Metz (1974) suggested that vigilance decrement is in reality a decrease in d'.

The theory is not without its critics, however. Broadbent (1971) questioned the verity of four TSD assumptions. First, TSD proposes that the distributions of the likelihood ratios are normal and have equal variance. This may not be true. Second, TSD assumes that an observer can arrange the perceived sensory evidence along a single dimension. Given the diversity of sensory inputs and the human's inability to aggregate evidence over time, this assumption is also questionable. Third, it is assumed that the observer knows the λ value for every possible experienced event. However, the human is an imperfect aggregator of evidence. Fourth, TSD assumes that the observer is always rational. This is again questionable. TSD, however, has served as a good model for research.

Another theory, proposed by Welford (1981), stated that the vigilance decrement was the result of a decrease in attention (arousal) over time. Four independent lines of research supported this theory. First, psychophysiological recordings of operators during vigilance tasks indicated that poorer detection performance was associated with lower physiological arousal (Blakeslee, 1979 & O'Hanlon, 1979). Second, stimulants such as nicotine and caffeine that raise an operator's physiological arousal decreased the magnitude of the vigilance decrement. Third, loud auditory stimulation, which also results in increased arousal, decreased the magnitude of the vigilance decrement. Finally, vigilance performance under the threat of electric shock was superior to that without the threat of shock (Kennedy & Coulter, 1975).

Welford's theory, however, could not explain the vigilance decrement that occurred when newly trained operators were on watch. The decrement they exhibited was much greater than that shown by experienced operators, and yet the novices seemed to be the more highly motivated and aroused of the two groups. To explain this phenomenon, Boker (1974) proposed another theory. It was based on the operator's expectancy of the signal occurrence. During training, operators are typically exposed to many signals in a short period of time. This high signal rate is meant to efficiently enhance the trainee's ability to recognize signals. The high signal rate, however:

- Causes the trainee to adopt a conservative decision rule (i.e., a stimulus must be strong to be considered a true signal), and
- Creates the expectancy of a high signal rate in the real world.

The operator thus misses weak signals especially as time-on-watch progresses. Further, the newly trained operator confronted with a low probability task, gradually becomes more familiar with the signal probability in the real world and attempts to match his or her response probability with it (Craig, 1978). It was found that the vigilance decrement of newly trained operators could be reduced if the probability of a signal was minimized during training.

The final vigilance theory to be described is the Vicious Circle Hypothesis. This hypothesis suggests that an observer's decisional rule (β) is determined by a subjective estimate of the probability of a signal (P(S)). But this is based on the probability of the observer saying yes and this in turn is based on the observer's β. Over time the observer will become more conservative. This causes him or her to miss weak signals and to decrease his or her subjective estimate of signal probability. This decreases the probability of saying yes, which makes the observer become even more conservative, hence the vicious circle.

The research to date has not clearly differentiated the strengths of the four theories presented above. This is partially because the research, especially that sponsored by the military, has been directed at more applied problems. The applied research suggests that vigilance decrement can be decreased by:

- Providing the operator with knowledge of results (i.e., the number of signals missed and the number of false alarms) (Poulton, 1973),
- Increasing the operator's motivation,
- Providing the operator with a more accurate (i.e., real-world) estimate of signal probability (Craig, 1980),
- Periodically exposing the operator to sample signals to maintain an accurate image of the signal characteristics in his or her memory,
- Introducing false signals, and

- Training the observers to use either fixed or completely random search patterns rather than alternating patterns (Fisk & Schneider, 1980).

The first four methods are difficult to implement in the real world. The fifth is akin to crying wolf and can give the operator an inaccurate estimate of system status. But, Mitta and Folds reported empirical data that high false alarm and hit rates as well as short detection latency enhance operator performance. Their data were collected in an automated traffic incident detection system. But vigilance performance even in high signal-rate environments (e.g., 30 signals/minute) Parasuraman, 1980) can be degraded over time. Also, the detection latency to real signals is greatly increased by the presence of non-relevant signals (Thackray, Bailey, & Touchstone, 1979). Finally, introducing false signals was originally meant to increase operator stimulation thus avoiding under load. However, Warm, Dember, and Hancock (1996) cite a number of experiments to support their conclusion that "under load is a myth: rather than being under stimulating, vigilance tasks are exacting, capacity-draining assignments that are associated with a considerable degree of mental demand and frustration" (p. 195).

Finally, a sobering finding is that vigilance of pairs can also degrade. For example, Wiener (1980) identified vigilance problems of both the pilot and the ATC to be at fault in midair collisions. There is clearly yet much to be learned about human vigilance and automation.

References

Baker, J. R. (1974). Immediate and delayed retention effects of interspersing questions in written instructional passages. *Journal of Educational Psychology, 66*, 96–98.

Blakeslee, P. (1979). Attention and vigilance: performance and skin conductance response changes. *Psychophysiology, 16 (5),* 413–319.

Broadbent, D. E. (1971). *Decision and stress.* New York, New York: Academic Press.

Craig, A. (1980). Effect of prior knowledge of signal probabilities on vigilance performance of a two-signal task. *Human Factors, 22 (3),* 361–371.

Craig, A. (1978). Is the vigilance decrement simply a response adjustment towards probability matching? *Human Factors, 20,* 441–446.

Fisk, A. D. & Schneider, W. (February 1980). Controlled and automatic processing during tasks requiring sustained attention. A new approach to vigilance (TR 8006). Arlington, Virginia: Office of Naval Research.

Goodenough, D. J. & Metz, C. E. (1974). Effects of listening interval on auditory detection performance. *Journal of Acoustical Society of America, 55(1),* 111–116.

Kennedy, R. S. & Coulter, X. B. (1975). Research note: The interactions among stress, vigilance, and task complexity. *Human Factors, 17,* 106–109.

Mitta, D. A. & Folds, D. J. Incident detection system design: the effects of high false alarm rate on operator performance, 1–7.

O’Hanlon, J. F. (April 1979). *Research on the psychophysiological basis of human vigilance* (TR 787-F). Arlington, Virginia: Office of Naval Research.

Parasuraman, R. (February 1980). *Memory load and event rate control sensitivity decrements in sustained attention* (TR-26). Arlington, Virginia: Office of Naval Research.

Poulton, E. C. (1973). The effect of fatigue upon inspection work. *Applied Ergonomics, 4,* 73–83.

Swets, J. A. (Ed.), (1964). Signal detection and recognition by human observers. New York: Wiley.

Swets, J. A., Tanner, W. P., & Birdsall, T. G. (1961). Decision processes in perception. *Psychological Review, 68,* 201–340.

Thackray, R. I., Bailey, J. P., & Touchstone, R. M. (1979). The effect of increases monitoring load on vigilance performance using a simulated radar display. *Ergonomics, 22 (5),* 529–539.

Warm, J. S., Dember, W. N., & Hancock, P. A. (1996). Vigilance and workload in automated systems. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: theory and applications,* (pp. 183–200). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Welford, A. T. (1981). Signal, noise, performance, and age. *Human Factors, 23(1),* 97–109.

Wiener, E. L. (1980). Midair collision: The accidents, the systems, and the realpolitik. *Human Factors, 22*, 521–533.