gerards
methodology

APPLYING THEORY OF SIGNAL DETECTION IN MARKETING: PRODUCT DEVELOPMENT AND EVALUATION

Robert C. Angus and Terry C. Daniel

Theory of signal detection (TSD) is a methodology developed by electrical engineers in the early 50's. It has recently been extended by psychologists to analyze perception of stimuli and observer decision criterion [1]. These elements are important to those developing new products, especially products which are multi-dimensional in terms of the stimuli to which consumers react. Accordingly, TSD should provide a useful method of evaluating the consumer preferences or their ability to discern product differences and how this ability is affected by the interaction of various product components. A brief description of an experiment involving different ice cream mixes will illustrate an application of TSD procedures to evaluate consumer reactions.

The objective was to determine the extent to which a taste-testing panel could discern richness of ice cream products and to investigate how fat level, flavor, and overrun (density of mix) influenced their judgment. The three factors (fat, flavor and overrun) were held at high, medium and low levels, resulting in 27 different ice cream mixes. Members of the panel were asked to rate each mix independently

---

Robert C. Angus is Professor of Agricultural Economics and Terry C. Daniel is Assistant Professor of Psychology, University of Arizona.

Contributed paper, AAEA, Edmonton, Canada, 1973

as rich or not rich.

This single-stimulus procedure is known to be susceptible to the confounding influence of response bias [2, p. 18]. An oberver may exhibit a high judgment accuracy for rich mixes, but have little ability to distinguish rich from not-rich. In this case, the observer has a high level of correct responses to rich mixes and a large number of incorrect responses to not-rich mixes. An observer's performance is determined by the relation between his correct and incorrect responses. TSD is helpful in this situation. It offers a means of distinguishing an observer's response bias, his decision criteria, from his true ability to discriminate between signal and non-signal.

The Decision Model

The classical signal detection approach requires an observer to judge whether a signal is present or absent. The signal stimulus was usually defined in physical terms. A correct decision, a hit, occurred if the observer identified the signal when it was present. Identification of a signal when it was not present was called a false alarm.

The ability of an observer to detect a signal from noise is described as the observer's receiver operating characteristic (ROC). The ROC is a bivariate function relating an observer's hit rate and false alarm rate. The ROC provides an estimate of the distance between the hypothetical noise distribution and the signal-plus-noise distribution. It is formulated by obtaining hit rates and false alarm rates for a number of decision criteria levels.

The ice cream mix experiment requires a contemporary application of TSD. In this case, there is no clear cut definition of richness as applied to a specific ice cream mix. Thus, we have no *a priori* physical definition for richness, and no reason to call a response a hit or a false alarm. For this application, the theoretical decision model can be illustrated by plotting the probability, given a specific mix, over a dimension of perceived richness, Figure 1. The dashed line represents a criterion for a judgment of "rich" selected by an observer. Perceived richness values to the left of this line are judged rich, and values to the right, not-rich. The dispersion inherent in each distribution is due to variation of the perceptual process of an individual from one observation to the next. The overlap of the distributions represents opportunities for confusion. In general, Figure 1 represents an easy discrimination because the overlapping areas are small.

Method of Analysis

Members of the taste panel reported their decisions as to the richness on a ten point scale. A scale value of nine meant that the subject was absolutely certain the mix was rich. A zero indicated absolute certainty that the mix was not rich. Absolutely certain decisions implied a subject used his most rigorous decision criterion. A rating of four implied the subject guessed and that the mix could be judged rich only under less rigorous criteria.

Individual TSD distributions or ROC functions may be visualized as a plot of cumulative probabilities obtained by cumulating the responses in each certainty category, Table 1. In our experiment, the mix which contained medium levels of fat, overrun and flavor (MMM) was arbitrarily selected as a standard of comparison [3]. Confidence level frequencies and cumulative probabilities have been computed for this mix. Each

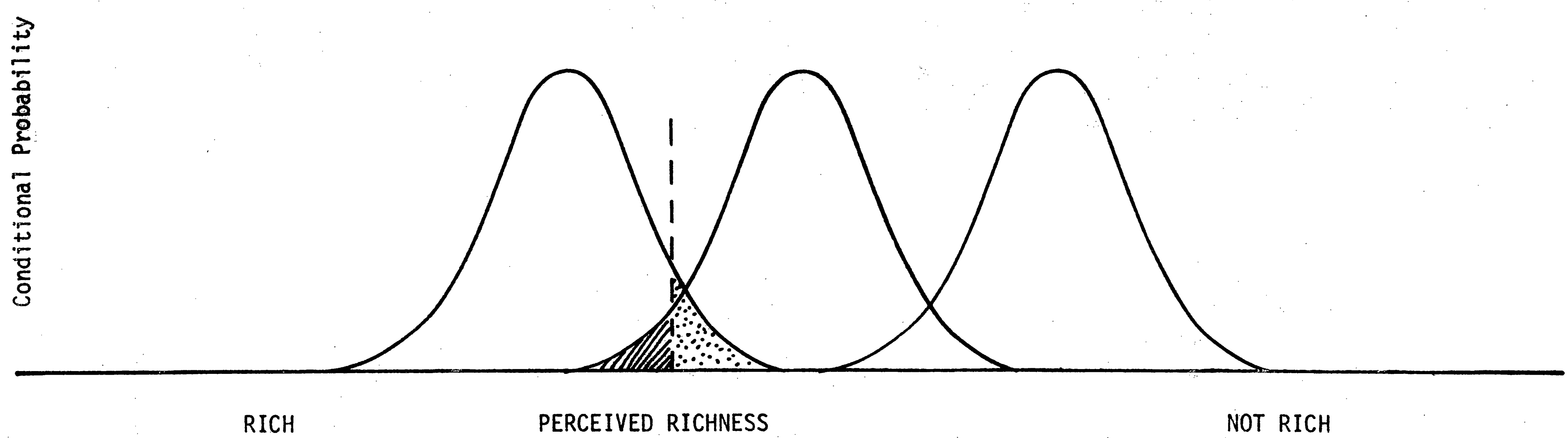

Figure 1. Theory of Signal Detection Model.

Table 1 - Data Illustrating TSD Computations

Certainty Level	Medium, Medium, Medium Mix Frequency	Cumulative Proportion	Z Transformation	Low, Low, Low Mix Frequency	Cumulative Proportion	Z Transformation
0 Certainty Not Rich	1	1.000	2.502	5	1.000	2.502
1	3	.979	2.038	8	.896	1.257
2	4	.917	1.382	6	.729	.608
3	3	.833	.965	7	.604	.262
4	4	.771	.739	4	.458	- .104
5	4	.688	.486	6	.375	- .316
6	7	.604	.262	8	.250	- .672
7	11	.458	- .104	4	.083	-1.382
8	5	.229	- .739	0	.000	-2.502
9 Certainty Rich	6	.125	-1.149	0	.000	-2.502
	48			48		
			.7102			-.3166

$D_m = .7102 - (-.3166) = 1.0268$

cumulative probability has been transformed into a Z score by use of the cumulative normal distribution. Confidence level frequencies and cumulative frequencies were computed for mixes to be compared, in this illustration the mix with the lowest fat, overrun and flavor (LLL). Z scores for the standard mix, MMM, if plotted against themselves, form the 45^{o} diagonal, Figure 2. A plot of the Z transforms of MMM and LLL mixes from Table 1 result in the line labeled ROC LLL. The distance of the ROC for LLL from MMM indicates the degree of discrimination. The further the ROC from the diagonal, the greater the degree of discrimination.

Several methods of measuring the distance of the ROC from the diagonal have been used. The first two, d' and d_s, are labeled in Figure 2. They may be computed in this application as follows:

$$d' = \text{mean } Z_{mmm} - (\sigma Z_{mmm}/\sigma Z_{111}) \text{ mean } Z_{111}$$

$$d_s = 2d'/[1 + (\sigma Z_{mmm}/\sigma Z_{111})]$$

The third distance metric, d_m, is simply the mean of the Z scores for the standard of comparison (MMM) minus the mean Z score for the mix which is compared (LLL), Table 2. Note that the difference in Z scores for the zero certainty score (d_o) is zero. This is because the Zs are transforms of the cumulative probability which must sum to one. Standard deviations and standard errors of d_i may also be computed. In any case, the larger the d_m value in absolute terms, the greater the observer's ability to distinguish between an ice cream mix and the standard, MMM.

Each of the measures, d', d_s, and d_m, has its particular advantages. d_m was used here because it is least affected by the slope of the ROC. It is easy to compute and considers every obtained point on the ROC.

Now that the d_m metric has been explained, it is possible to

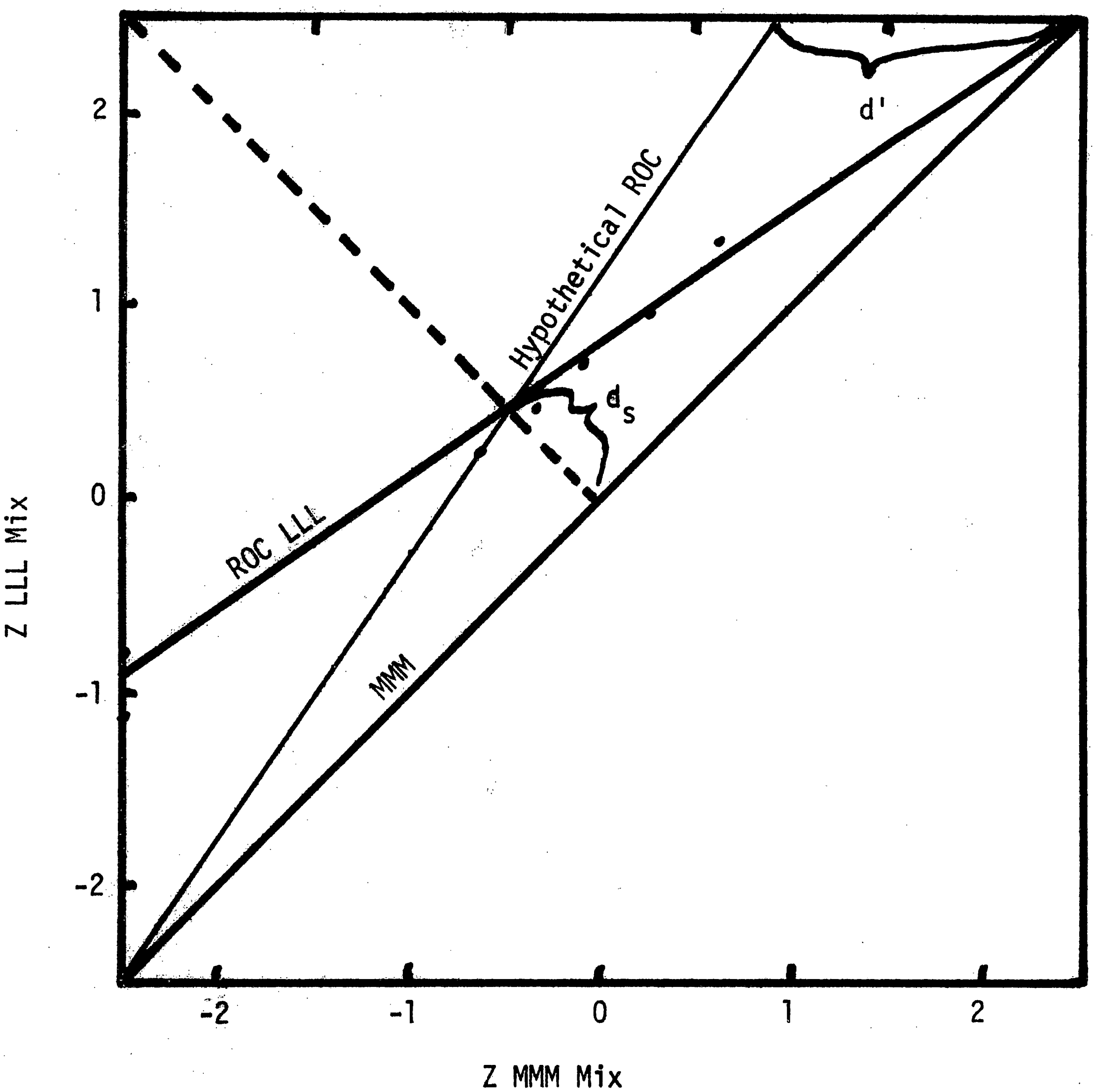

Figure 2. Normalized ROC Function.

Table 2. Computations for the d_m Distance Metric

Certainty Level		Difference MMM		LLL		d_i		
0	Certain Not Rich	$Z_{mmm}0$	-	$Z_{111}0$	=	d_0	=	0
1		$Z_{mmm}1$	-	$Z_{111}1$	=	d_1		
2		$Z_{mmm}2$	-	$Z_{111}2$	=	d_2		
3		$Z_{mmm}3$	-	$Z_{111}3$	=	d_3		
4		$Z_{mmm}4$	-	$Z_{111}4$	=	d_4		
5		$Z_{mmm}5$	-	$Z_{111}5$	=	d_5		
6		$Z_{mmm}6$	-	$Z_{111}6$	=	d_6		
7		$Z_{mmm}7$	-	$Z_{111}7$	=	d_7		
8		$Z_{mmm}8$	-	$Z_{111}8$	=	d_8		
9	Certain Rich	$Z_{mmm}9$	-	$Z_{111}9$	=	d_9		

$$d_m = \sum_{i=1}^{9} Z_{mmm}/N_c - 1 \; - \; \sum_{i-1}^{9} Z_{111}/N_c - 1 \; = \; \sum_{i=1}^{9} d_i/N_c - 1$$

illustrate how the TSD technique eliminates bias due to the judgment criterion. The mean confidence score for each mix was a function of perceived richness and judgment criterion or bias. The d_m is a function of perceived richness only. Both perception and judgment criterion effects are compounded in a single rating response. Therefore, an analysis of variance of simple mean ratings fails to separate perception from judgment criterion.

This argument can be illustrated by the following hypothetical situation. Consider two taste panels, each of which are atypical of the population in terms of judgment criterion. For example, a home economist group used in a pretest rated the high fat, low overrun and low flavor mix (HLL) as being more than fairly certain rich, a rating of eight. A second test group rated the same mix as fairly certain not rich, a rating of two. It appears from the ratings that the home economist group perceived the mix as being richer than did the other group. However, it is possible that the difference was due entirely to differences in judgment criteria and that both panels perceived the mix as equally rich. This appears to be the case because the home economist group rarely rated any mix below fairly certain rich, that is almost all mixes were rated as eights or nines. Conversely, the second group rarely rated any of the mixes above just guessing not rich, a rating of four. Since the signal detection index expresses differences in ratings of the HLL mix from ratings of the MMM mix by the same group, this judgment bias is eliminated and the d_m values were found to be the same for the two groups. This illustrates the importance of transforming the ratings to signal detection metrics, d_m, rather than accepting ratings at face value.

The example above illustrates how differences in criteria, by changing the origin of rating responses, can make comparisons of raw rating scores ambiguous. A similar problem arises if one panel uses an expanded rating scale. The Z transformation in the TSD analysis standardizes the units for the d_m metric. Other techniques may be used to accomplish standardization of rating scores, but these require substantial expansion of the number of judgments per stimulus per subject.

The results of our experiment with the 27 ice cream mixes are summarized in Figure 3. The numbers on the graphs at each mix represent current ingredient costs in dollars per gallon of finished product. Our research groups had hypothesized that the richest mix would be 14 percent fat, 78 percent overrun and triple flavor at 93 cents [4]. Note that our panel perceived 14 percent fat, 102 percent overrun and double flavor mix as richest. The cost of ingredients for this mix was 77 cents. Note that in some instances a few cents difference in cost can produce substantial increases in perceived richness of the product, for example MMM at 75 cents versus HHM at 77 cents. On the other hand increasing the cost of mix rather substantially in some cases may produce no change, or even a decrement, in perceived richness. The use of TSD methodology in this experiment has insured that these results are not restricted to the particular judgment criteria of our panel, but would be expected to hold for other panels having more stringent or less stringent criteria for judging product richness.

While our results were somewhat surprising, the conclusion is not at all unique. Over 60 years ago, a group of experimenters studied ice cream mixes with methods and analyses common for that period [5]. The following is a quote from Vermont Agricultural Experiment Station Bulletin

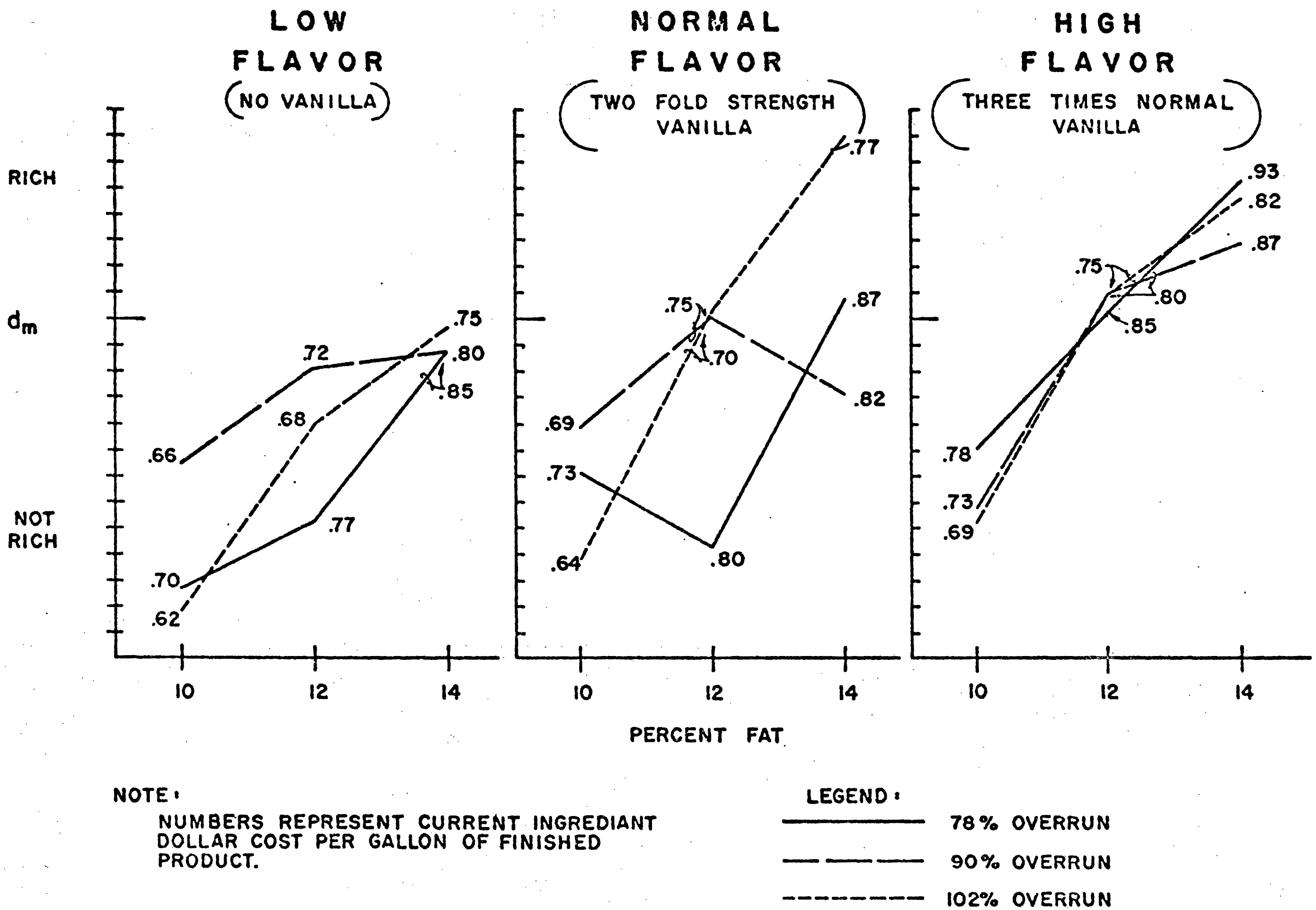

Figure 3. Relative Richness (d_m value) and Cost of Mix.

155 dated 1910:

"It may seem strange to some that an experiment station should approve the incorporation of air into ice cream. They may reason that the station advocates the dilution of the product, the selling of 'wind' as ice cream, a course quite as open to objection from the ethical, if not the legal standpoint as is the dilution of milk with water. It is a fact that an ice cream the volume which is approximately a third air is more satisfactory to the consumer than one containing no air. It has a more velvety feel to the tongue, and conveys a sensation of richness without causing the unpleasant effects of an excessively rich cream."

REFERENCES

[1] Green, D. M. and J. A. Swets, Signal Detection Theory and Psychophysics, John Wiley, New York, 1966.

[2] Wheeler, L., T. Daniel, G. Seeley and W. Swindell, "Detectability of Degraded Visual Signals: A Basis for Evaluating Image-Retrieval Programs," Technical Report 73, Optical Sciences Center, University of Arizona, 1971.

[3] Daniel, T. C., L. Wheeler, R. S. Boster and P. R. Best, Jr., "An Application of Signal Detection Analysis to Forest Management Alternatives," Man-Environment Systems. In Press.

[4] Swartz, N. A., Development and Marketing New Dairy-Based Foods, M.S. Thesis, University of Arizona, 1973.

[5] Vermont Agricultural Experiment Station, Bulletin 155, University of Vermont, 1910, pp. 32-33.