

On the Optimal Number of Alternatives at a Choice Point¹

AMOS TVERSKY

University of Michigan, Ann Arbor, Michigan

Given a fixed total number of alternatives for a multiple-choice type test, the use of three alternatives at each choice point will maximize discriminability, power and information of a test. A proof is presented and applications to test construction, task design, and information processing are briefly discussed.

Consider any test task or questionnaire which can be represented as a sequence of choice points at each of which one out of a set of alternatives is chosen. Multiple-choice tests, mazes or personality check lists are examples. Given a fixed total number of alternatives for the whole test, we wish to find the optimal number of choice points and the optimal number of alternatives at each choice point.

Let k be the total number of alternatives, and let x_i denote the number of alternatives at the i th choice point. Thus:

$$\sum_{i=1}^r x_i = k \quad (1)$$

where r is the number of choice points.

Three criteria for optimality which one may attempt to maximize will be considered.

(a) Discrimination capacity: the number n of possible distinct response patterns of a given test

$$n = \prod_{i=1}^r x_i \quad (2)$$

For example, consider a multiple-choice diagnostic test or an attitude questionnaire based on Likert-type items. Let every sequence of responses be regarded as a different "personality type" or "attitude profile." Thus, by maximizing n , we maximize the number of distinct types or profiles among which the test enables us to discriminate.

¹ This research was supported by the Air Force Office of Scientific Research under contract AF-AFOSR-196-63, and by Public Health Service Research grant MH-04236.

(b) Power: defined as 1 minus the probability of attaining perfect performance by chance alone. Assuming equal probability of guessing for all alternatives:

$$\text{Power} = 1 - \left(\prod_{i=1}^r x_i \right)^{-1} \tag{3}$$

Clearly, we are interested in constructing tests with maximal power.

(c) Uncertainty associated with the set $[A]$ of all possible response patterns to a given task.

$$H[A] = - \sum_{i=1}^n p_i \log_2 p_i = \log_2 \prod_{i=1}^r x_i \tag{4}$$

Since the three criteria proposed are strictly monotonically related to each other, maximizing (2) maximizes (3) and (4) as well. If the same number of alternatives is used at each choice point, i.e., $x_i = x_j$, for any i and j then

$$n = \prod_{i=1}^r x_i = x^r \tag{5}$$

and since $x \cdot r = k$

$$n = x^{k/x}.$$

Let us denote by $f_k(x)$ the discrimination function $x^{k/x}$ of such a test. $f_k(x)$ can be maximized by setting $(d/dx)f_k(x) = 0$.

$$\begin{aligned} \frac{d}{dx} f_k(x) &= \frac{d}{dx} (x^{k/x}) = \frac{d}{dx} (e^{\ln x^{k/x}}) \\ &= e^{\ln x^{k/x}} \frac{d}{dx} (\ln x^{k/x}) = x^{k/x} \left(\frac{k}{x^2} - \ln x \frac{k}{x^2} \right) \\ &= x^{k/x} \frac{k}{x^2} (1 - \ln x) \end{aligned} \tag{6}$$

Since both x and k are positive $(d/dx)f_k(x) = 0$ and only if

$$1 - \ln x = 0, \quad \ln x = 1, \quad \text{or} \quad x = e.$$

Thus $f(x)$ has a unique maximum at $x = e = 2.718$.

A family of discrimination functions for some different values of k is given in Fig. 1. Note that though the value of $f_k(x)$ depends rather heavily on k , the location of its maximum is completely independent of k .

(Note, incidentally, that the well-known function $f(p_i) = -p_i \log_2 p_i$, is a logarithmic transformation of the discrimination function $f_k(x)$ where $k = 1$. The above result may be used to solve for its maximum.

Let $\log_2 y = -p_i \log_2 p_i$ and let $x = 1/p_i$. Hence:

$$y = \left(\frac{1}{p_i}\right)^{p_i} = x^{1/x} = f_1(x) \tag{7}$$

By (6), $f_1(x)$ is maximum at $x = e$. Indeed $-p_i \log_2 p_i$ is maximum at $p_i = 1/e = 0.368$.)

Since x stands for the number of alternatives, we want to find the integer which is closest to the maximum point. Since $f_k(x)$ is single-peaked it should be either 2 or 3.

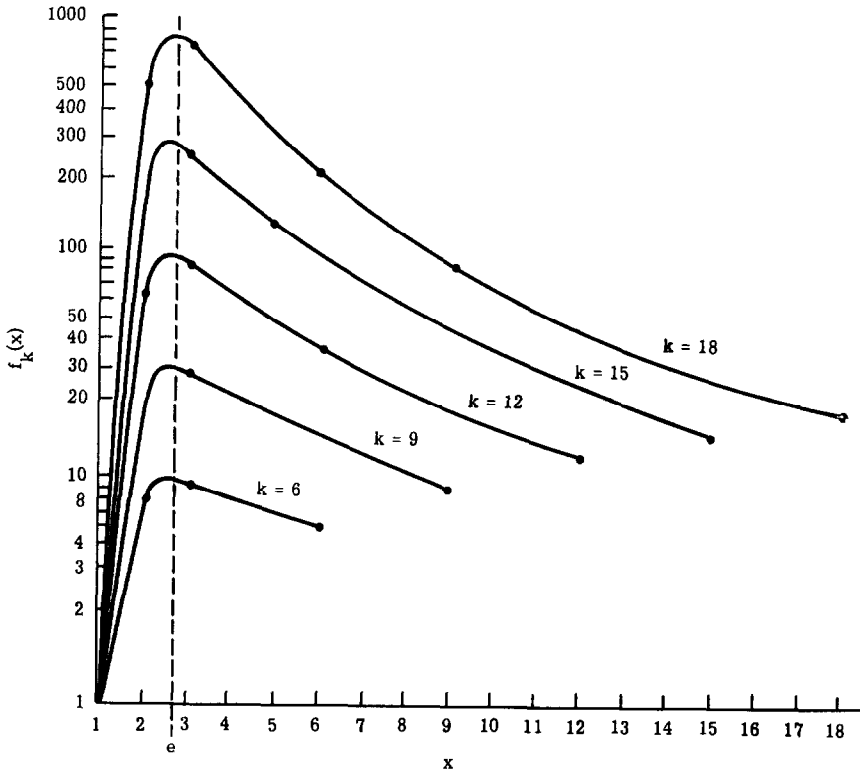


FIG. 1. A family of discrimination function $f_k(x)$ for some different values of k . The dots on the curves denote the points at which both x and k/x are integers.

To show that 3 gives the desired solution we have to show that $f_k(3) > f_k(2)$ for any $k > 0$, i.e., $3^{k/3} > 2^{k/2}$. Raising both sides of the above equation to the $6/k$ power yields the desired result.

Hence the use of three alternatives at each choice point will maximize discrimination capacity, power, and uncertainty of the test.

The final result for the case in which both x and k/r are integers is given by the following theorem.

THEOREM. *Consider any sequence of positive integers $x_1, x_2 \dots x_r$ such that $\sum_{i=1}^r x_i = k$. A sequence for which $\prod_{i=1}^r x_i$ is maximum will include a maximal number of 3's without including any 1.*

That is, the best solution will include either (i) $k/3$ choice points with 3 alternatives each, or (ii) $(k - 4)/3$ choice points with 3 alternatives each and a single choice point with 4 alternatives, or (iii) $(k - 2)/3$ choice points with 3 alternatives each and a single choice point with 2 alternatives, depending on whether the remainder of k when divided by 3 is zero, one, or two, i.e., whether $k \bmod (3)$ equals 0, 1, or 2.

The proof of the theorem is as follows: For every sequence of positive integers, denoted by S , with a fixed sum, k , which is not in any one of the forms (i), (ii), or (iii) it is possible to construct another sequence denoted by S' with the following properties:

- (1) $\sum S' = \sum S = k$.
- (2) $\prod S' > \prod S$.
- (3) S' is in one of the forms (i), (ii), or (iii).

The construction of such a sequence is done by successive replacements of elements in the original sequence which do not change its sum. Let S be any sequence of positive integers. If it contains 1, delete the 1 by adding it to some other element. The new sequence obtained will have a greater product since

$$x + 1 > x \cdot 1 \quad \text{for any } x \geq 1. \tag{8}$$

If the sequence contains 2's or 3's, leave them unchanged. If the sequence contains 4, replace it by a pair of 2's, leaving both the sum and the product unchanged.

Next, any even x can be expressed as $2t$ for some positive integer t . Hence replace any even $x > 4$ by its corresponding t -tuple of 2's. The product of the new sequence will exceed that of the old one since

$$2^t > 2t \quad \text{for any } t > 2. \tag{9}$$

Similarly, any odd $x > 3$ can be expressed as $2t + 3$ for some positive integer t . Thus we may replace any odd $x > 3$ by a single 3 and its corresponding t -tuple of 2's. This replacement will increase the product of the sequence because

$$2^t \cdot 3 > 2t + 3. \tag{10}$$

Repeated application of these replacements will yield a sequence consisting of 2's and 3's only.

Finally, replace any triple of 2's by a pair of 3's, thus increasing the product since

$$3^2 > 2^3. \quad (11)$$

The final sequence thus obtained will be in one and only one of the desired forms (i), (ii), or (iii). The maximality of $\Pi S'$ for this sequence follows from the inequalities (8) through (11).

The form of the desired sequence S' is unique except for order and except in case (ii) in which either a single 4 or a pair of 2's may be used.

Alternatively, the above theorem may be stated as follows: Consider all sequences of positive integers x_1, x_2, \dots, x_r with a fixed product k , i.e., $\prod_{i=1}^r x_i = k$. The sequence for which $\sum_{i=1}^r x_i$ is minimum will include a maximal number of 3's without including any 1.

DISCUSSION

The criteria proposed will be maximized by constructing tasks whose alternatives are of the form given by the above theorem. Certainly, additional criteria may be considered in deciding upon the number of alternatives to be used. There are, however, instances in which the above result may be directly applicable.

Whenever the amount of time spent on the test is proportional to its total number of alternatives, the use of three alternatives at each choice point will maximize the amount of information obtained per time unit. This seems to be true of multiple-choice tests consisting of questions like: "which of the following passages best describes X 's position?" in which the amount of time spent on the question is negligible compared to the time spent on choosing among the alternatives. The result is applicable, however, even in instances in which the proportionality assumption does not hold. All that is needed is that the relative gain in information will exceed the relative loss in time.

An estimate of the relative gain (or loss) of information, power, and discriminability as a function of some different values of x for a given k can be obtained from the graph of $f_k(x)$ in Fig. 1. Whenever additional criteria are explicitly introduced, the above estimate may be taken into account in constructing an optimally-designed task.

There exists some empirical evidence (Pressey, 1962), based on the study of auto-instructional items, which indicates that three-alternative test items are indeed optimal. Since neither time nor the total number of alternatives was controlled, the results are only suggestive.

Finally, the above result may shed some light on the study of information coding and processing. In a paper entitled "Information transmission with elementary auditory displays," Sumbly, Chambliss, and Pollack (1958) have used a set of auditory signals as an alphabet. They employed four stimulus variables, with two, three, and

five alternatives per variable. In summarizing the results Garner (1962) says: "Their results showed that three alternatives per variable gave the best performance, agreeing with the suggestion in the Pollack and Ficks (1954) results that three levels per dimension are better than two. Certainly, at this stage it seems that the maximum information transmission will be obtained with humans when no more than three alternatives are used with a single variable, but with many variables involved," (pp. 122-123). In other words, the data show that perceptual discrimination, measured by the amount of information transmitted, was maximized when each one of the dimensions has three levels.

One may hypothesize that the discriminability of the stimuli, or the memory load associated with them, is directly related to the total number of levels summed over dimensions. The fact that amplitude, for example, is a relevant dimension does not contribute to the *S*'s memory load; rather, it is the number of levels of amplitude which hinders discrimination. The total number of levels, however, was shown to be minimal whenever three-level factors are employed. If the above hypothesis is true, it follows that the use of three-level factors will minimize confusion and decrease memory load. This problem seems worthy of experimental investigation. Taken together with Garner's conclusions concerning human capacities to process multi-dimensional information, these results suggest that the use of three levels per dimension may be the most efficient way to code and process information.

ACKNOWLEDGMENTS

I wish to thank J. Give'on for his helpful suggestions and Dr. R. M. Dawes for calling my attention to the relationships between the finding of Pollack and Ficks and the result reported here.

REFERENCES

- GARNER, W. R. *Uncertainty and structure as psychological concepts*. New York: Wiley, 1962.
- POLLACK, I., AND FICKS, L. Information of elementary multidimensional auditory displays. *J. acoust. Soc. Amer.*, 1954, **26**, 155-158.
- PRESSFY, S. L. Basic unresolved teaching-machine problems, *Theory and Practice*, 1962, **1**, No. 1.
- SUMBY, W. H., CHAMBLISS, D., AND POLLACK, I. Information transmission with elementary auditory displays. *J. acoust. Soc. Amer.*, 1958, **30**, 423-429.

RECEIVED: June 1, 1963