# REDUNDANCY IN TASK ASSIGNMENTS AND GROUP PERFORMANCE*

ROBERT B. ZAJONC AND WILLIAM H. SMOKE

UNIVERSITY OF MICHIGAN

The problem of combining abilities of group members to maximize the performance of the group as a whole is examined in terms of redundancy in task assignments. In particular, ways of distributing a given number of items of information among a given number of individuals to obtain the maximum probability of each item being recalled by at least one individual are studied. It is shown that there exists an optimal distribution scheme which is independent of the amount of material originally given, the size of the group, and individual differences in ability.

The problem of assessing determinants of group performance has been investigated from two different points of view. For the most part, studies in this area have been concerned with the effects of group variables, such as the presence of others [1, 2], cohesiveness [e.g., 8], leadership style [e.g., 5], and the like, on the performance of groups and of individuals working in groups. Recently, some attempts have been made to analyze the group product by means of a combinatorial analysis of individual abilities. Lorge and Solomon [4] performed such an analysis in the area of group problem solving, and Hays and Bush [3] have used it in group learning. In principle, this latter approach is analogous to that of Moore and Shannon in what they called the "crummy relay problem" [6], which refers to constructing reliable circuits out of unreliable relays. von Neumann [9] has shown that by using a number of components of limited unreliability a reliable machine may be constructed. Moore and Shannon have demonstrated that a reliable circuit may be designed by using arbitrarily unreliable relays. The increase in circuit reliability is obtained essentially by increasing the redundancy among relays. It would seem that the study of group performance in terms of redundancy among the abilities of individual members would hold considerable promise.

Consider, for example, a group of $N$ individuals. Let $H$ items of information be given to these individuals. The object is to recover this information from the group as a whole after some interval of time. For the present purposes it is irrelevant which individual remembers a particular item, although the item should be remembered by somebody in the group. There is evidence in

the area of individual recall to the effect that the proportion of items recalled is inversely related to the number of items originally assigned [10]. Consequently, the probability that a given item is remembered by a given individual is some inverse function of the number of items he was asked to learn. On the other hand, the probability that at least one individual of those assigned the given item remembers it increases with the number of individuals assigned the item. The first consideration implies minimizing the number of items per individual, the second maximizing it. The problem then is to discover the optimal distribution of items among individuals. The group as a whole is considered to remember an item when that item is remembered by at least one individual.

## Case I

The following conditions are imposed on Case I.

(a) The probability $p(i, j)$ that item $i$ is remembered by individual $j$ is equal to the constant $p$, $0 \leq p \leq 1$, or to zero according to whether or not item $i$ is assigned to $j$.

(b) Each individual is assigned the same number, $h$, of items. Thus $\sum_i p(i, j) = hp$.

(c) Each item is assigned to an equal number, $n$, of individuals. Thus the probability, $P$, that a given item is recalled by at least one individual is given by

(1)                     $$P = 1 - \prod_n (1 - p) = 1 - (1 - p)^n,$$

the same for all items.

Under the conditions of Case I the problem is reduced to finding the assignment of items which generates the greatest value of $P$, i.e., finding values of $p$ and $n$ that will maximize $P$. If it is assumed that one is dealing not with number of items (a discrete measure) but with amount of material or amount of information (continuous measures), $P$ may be regarded as a differentiable function of $p$. The necessary condition for $P$ to be maximum as a function of $p$ is that $dP/dp = 0$. Since $h$ is a differentiable function of $p$, and since by (b) and (c) $Nh = Hn$, it follows that $n$ is also a differentiable function of $p$. Thus, from (1)

(2)          $$\frac{dP}{dp} = 0 = -(1 - p)^n \left[ \frac{dn}{dp} \log (1 - p) - \frac{n}{1 - p} \right].$$

Given $p \neq 1$, $dP/dp = 0$ only if the bracketed expression is zero, or

(3)                $$\frac{dn}{dp} \log (1 - p) - \frac{n}{1 - p} = 0.$$

Since $n = Nh/H$,

$$(4) \qquad \frac{dh}{dp} \log (1 - p) - \frac{h}{1 - p} = 0.$$

Thus far the function relating $p$ to $h$ has not been specified. However, the condition in (4) holds for *any* set of values $p$ and $h$ which satisfy the relation $p = 1 - e^{k/h}$, for an arbitrary constant $k$. Thus

$$P = 1 - [1 - (1 - e^{k/h})]^{Nh/H} = 1 - e^{Nk/H}$$

is a constant since all the terms in the expression are constants. Thus, if the probability of a given item being recalled by a given individual were given by $p = 1 - e^{k/h}$, all assignments would be equally good. Under these conditions what is lost in $p$ by assigning more items to each individual is gained by increasing $n$, and consequently the probability that an item is remembered by the group as a whole is independent of the assignment of items to the group members.

If, however, the relation between $p$ and $h$ is not given by $p = 1 - e^{k/h}$ but by some other function $p = f(h)$, then in general not all assignments will be equally good. In fact, on the basis of empirical data available in this area [10] it would appear that the function is of the form $p = e^{-k^2h^2}$, for an empirical parameter $k$ which depends on such factors as time, nature of the material, its organization, meaningfulness, or the like. This function fits data gathered by Oberly [7] with $k = .10$. Assuming $p = e^{-k^2h^2}$ and finding $dh/dp$,

$$(5) \qquad (1 - p) \log (1 - p) - 2p \log p = 0,$$

which is satisfied approximately for $p = .84$. Figure 1 shows the relation between $P$ and $p$ for selected values of $k$ with $N/H = .01$.

Solving for $h$, $h = \sqrt{\log .84/-k^2} = .42/k$. Thus, the best assignment results when each individual is assigned $.42/k$ items. In terms of $p$, the maximum value of $P$ obtains when each individual is given the number of items which would result in his forgetting about 16% of the material. Since the relationship between $h$, which denotes individual loads, and $n$, which reflects the amount of task-assignment redundancy, is known, the optimal amount of redundancy may be obtained. Thus, for Case I, $n = (.42/k)(N/H)$, the optimal amount of task-assignment redundancy.

It is rather interesting to note that the result obtained is entirely independent of the size of the group, $N$, and the number of items, $H$. Thus for any given number of individuals and any given number of items, $.42/k$ items per person represents the optimal assignment. Of course, the optimal amount of redundancy and consequently the maximum value of $P$ vary with the ratio of individuals to items. The larger this ratio the higher the maximum possible value of $P$. Figure 2 represents the relationship between the maximum values of $P$ and the $N/H$ ratio for some selected values of the parameter $k$.
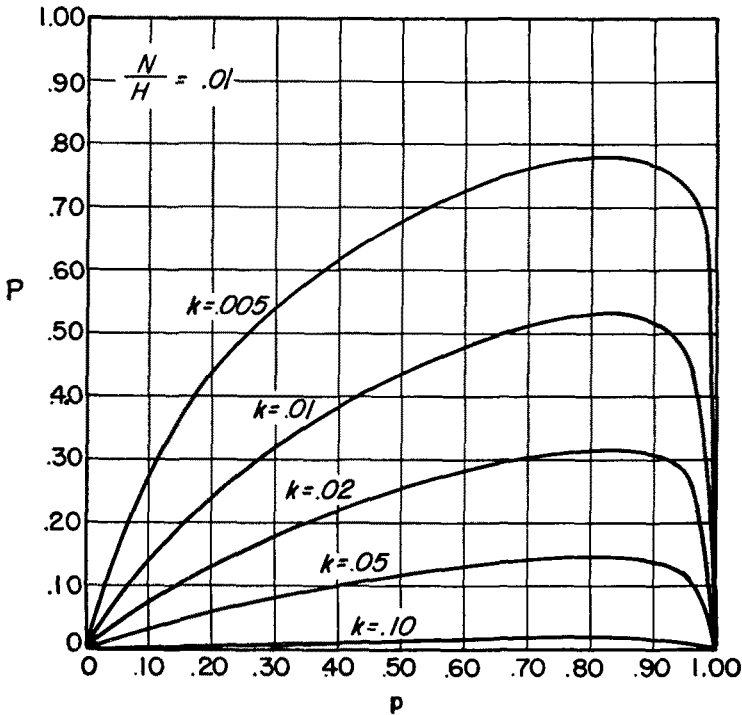
The Relationship Between $P$ and $p$ for Different Values of the Constant $k$.

The above solution was obtained by assuming $h$ to be a continuous variable. Given a group of $N$ individuals and a collection of $H$ discrete items, it will not be possible in general to assign the items to individuals in more than a small number of ways. Hence the assignments do not vary continuously. As a matter of fact, for given values of $N$ and $H$ only some of all the possible assignments satisfy the conditions (b) and (c). It can be demonstrated that if $D$ is the greatest common divisor of $N$ and $H$, the number of assignments of $H$ items to $N$ group members satisfying (b) and (c) is equal to $D$, given that two assignments are not considered distinct when they assign the same number of items to each individual.

*Case II*

In Case II items of equal difficulties and individuals with equal recall capacities were considered. Now the case where there exist individual differences in recall will be examined.

Again a set of conditions is imposed.

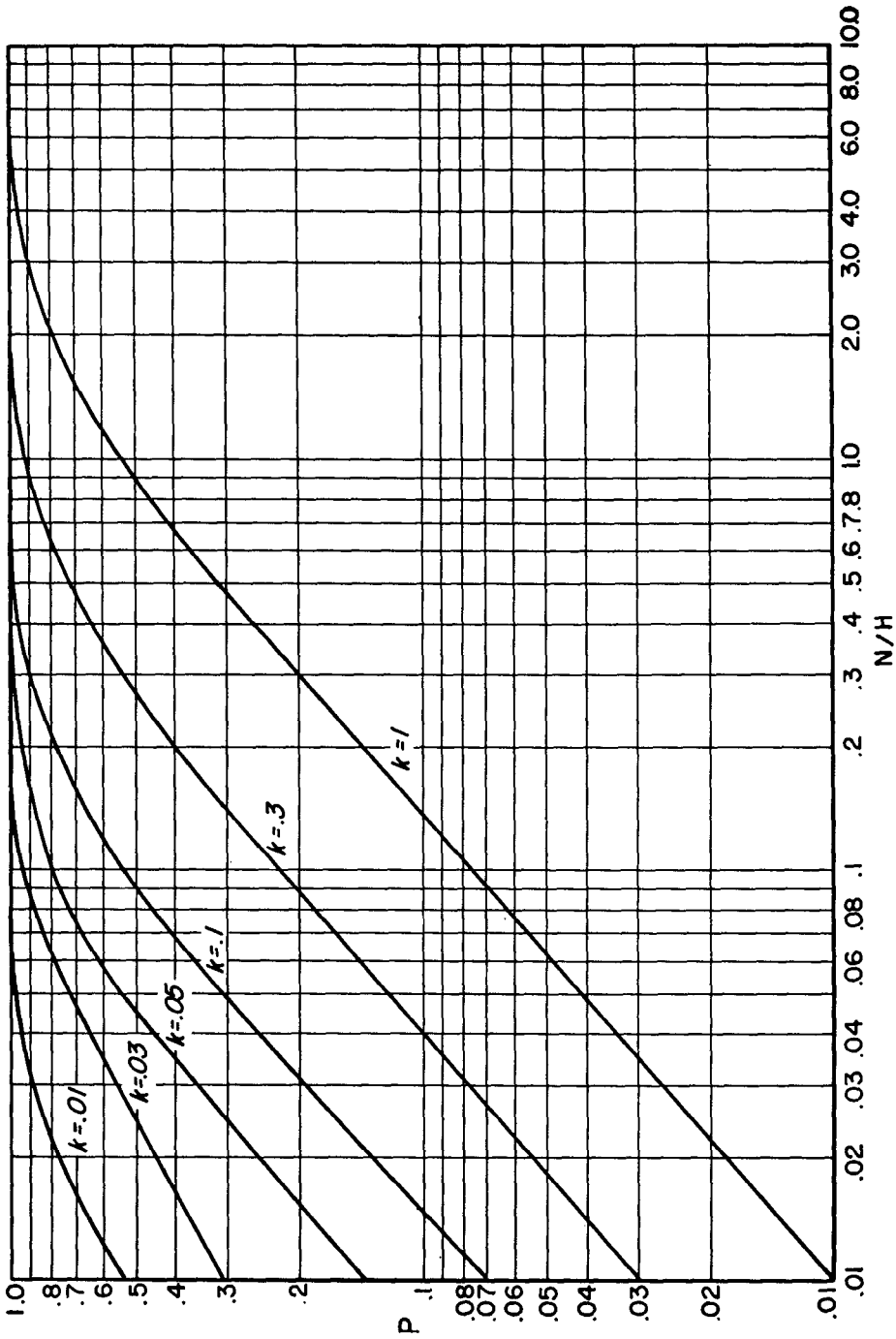(a) The probability, $p(i, j)$, that item $i$ is remembered by individual $j$, is

FIGURE 2

Maximum Values of $P$ as a Function of $N/H$.

equal to the constant $p$, $0 \leq p \leq 1$, or to zero according to whether or not item $i$ is assigned to $j$.

(b) The number, $h_j$, of items assigned to individual $j$ is such that $p = e^{-k_j{}^2 h_j{}^2}$, where $k_j$ is an empirical parameter obtained with respect to the individual $j$.

(c) The items are so distributed that for each item $i$ the probability that $i$ is remembered by at least one individual is equal to the constant $P' = 1 - \prod_j [1 - p(i, j)]$, the same for all items.

Thus, the above conditions imply that different individuals will be assigned different numbers of items, depending on their individual abilities to remember them. The conditions also imply that each item $i$ is assigned to the same number of individuals $n$, or that the redundancy is equal for all items. Hence

$$(6) \qquad P' = 1 - \prod_j [1 - p(i, j)] = 1 - (1 - p)^n.$$

Again, the necessary condition that $P'$ be a maximum as a function of $p$ is that $dP'/dp = 0$. Thus

$$(7) \qquad \frac{dP'}{dp} = 0 = -(1 - p)^n \left[ \frac{dn}{dp} \log (1 - p) - \frac{n}{1 - p} \right].$$

If $p \neq 1$, then

$$(8) \qquad (1 - p) \log (1 - p) \frac{dn}{dp} - n = 0.$$

In terms of the above conditions the number of assignments is equal to $nH$ and to $\sum_j h_j$. Hence $nH = \sum_j h_j$. Thus

$$\frac{dn}{dp} = \frac{1}{H} \sum_j \frac{dh_j}{dp}.$$

From restriction (b),

$$\frac{dp}{dp} = 1 = -2k_j^2 h_j \frac{dh_j}{dp} e^{-k_j{}^2 h_j{}^2} = -2k_j^2 h_j \frac{dh_j}{dp} p = (2p \log p) \frac{1}{h_j} \frac{dh_j}{dp}.$$

Thus

$$h_j = (2p \log p) \frac{dh_j}{dp}$$

and

$$\sum_j h_j = (2p \log p) \sum_j \frac{dh_j}{dp}.$$

Therefore

$$\frac{dn}{dp} = \frac{1}{H} \sum_i \frac{dh_i}{dp} = \frac{1}{2p \log p} \left(\frac{1}{H}\right) \sum_i h_i = \frac{n}{2p \log p}$$

and

$$(1 - p) \log (1 - p) \frac{dn}{dp} - n = 1 - p) \log (1 - p) \frac{n}{2p \log p} - n,$$

or

(9) $$(1 - p) \log (1 - p) - 2p \log p = 0.$$

Note that individual differences do not influence the solution, as (9) is satisfied for $p = .84$. However, the number of items, $h_i$, to be assigned to the different individuals will depend on their recall abilities which are reflected in the constants $k_i$. If the assignment of items to individuals satisfies (a) and (b), then for two individuals $j_1$ and $j_2$

$$\exp (-k_{j_1}^2 h_{i_1}^2) = p = \exp (-k_{j_2}^2 h_{i_2}^2)$$

or

$$k_{j_1} h_{i_1} = k_{j_2} h_{i_2} .$$

Thus

$$h_{i_2} = \frac{k_{j_1}}{k_{j_2}} h_{i_1}$$

and in general

$$h_{i_a} = \frac{k_{j_1}}{k_{j_a}} h_{i_1}$$

for each member $j_a$ of the group. Since $\sum_i h_i = nH$,

$$nH = \sum_j \frac{k_{j_1}}{k_{j_a}} h_{i_a} = h_{i_1} k_{j_1} \sum_j \frac{1}{k_j} .$$

Hence

$$h_{i_1} = \left(\frac{nH}{k_{j_1}}\right) \left[\frac{1}{\sum_j \frac{1}{k_j}}\right],$$

and in general

(10) $$h_{i_a} = \left(\frac{nH}{k_{j_a}}\right) \left[\frac{1}{\sum_j \frac{1}{k_j}}\right]$$

for each individual $j$. In this case individual differences are exploited by assigning fewer items to less able members and more items to the capable individuals.

While the value $p = .84$ is optimal under the restrictions specified above, it remains to be determined whether a different solution is obtained by relaxing restriction (a) such that $p(i, j)$, is no longer required to be a constant across individuals.

## Conclusions

The solutions presented provide a standard against which empirical results may be compared. Empirical tests must conform to the predictions, otherwise the restrictions imposed on the solutions could not have been met. Thus by careful experimental controls it can be discovered what variables determine the departures from the prediction.

For instance, the conditions imposed above require that $p(i, j)$, the probability of the individual $j$ recalling the item $i$, be constant. This, of course, necessitates a complete independence of the recall probabilities in terms of the items, as well as in terms of individuals. The probability $p(i_1, j_1)$ must be independent of $p(i_1, j_2)$ and of $p(i_2, j_1)$. Therefore, the cases examined are valid not for groups of interacting members but for collections of individuals working independently of one another.

This requirement, however, is not at all a shortcoming. On the contrary, it allows one to study the effects of group interaction on individual and group performance in recall. The existence of group interaction would probably lower the value of the parameter $k$, and this effect can be evaluated readily by empirical tests. The difference between the values of the parameter $k$ for individuals working together and for individuals working alone would provide information on the effect of group interaction on individual performance, and the difference between the corresponding values of $P$, information concerning the effect of group interaction on group performance.

It is noted that the model presented is not restricted to recall, and that it may, with slight modifications, be applied to other behaviors, such as learning, problem solving, or decision making.

## REFERENCES

[1] Allport, F. H. The influence of the group upon association and thought. *J. exp. Psychol.*, 1920, **3**, 152–182.
[2] Dashiell, J. F. An experimental analysis of some group effects. *J. abnorm. soc. Psychol.*, 1930, **25**, 190–199.
[3] Hays, D. G. and Bush, R. R. A study of group action. *Amer. sociol. Rev.*, 1954, **19**, 693–701.
[4] Lorge, I. and Solomon, H. Two models of group behavior in the solution of Eureka-type problems. *Psychometrika*, 1955, **20**, 139–148.
[5] Maier, N. R. F. The quality of group decisions as influenced by the discussion leader. *Hum. Relat.*, 1950, **3**, 155–174.

[6] Moore, E. F. and Shannon, C. E. Reliable circuits using less reliable relays. Part I. *J. Franklin Inst.*, 1956, **262**, 191–208.

[7] Oberly, H. S. A comparison of the spans of "attention" and memory. *Amer. J. Psychol.*, 1928, **40**, 295–302.

[8] Schachter, S., Ellertson, N., and Gregory, D. An experimental study of cohesiveness and productivity. *Hum. Relat.*, 1951, **4**, 229–238.

[9] von Neumann, J. *Probabilistic logic.* Pasadena: California Inst. Technol., 1952.

[10] Woodworth, R. S. *Experimental psychology.* New York: Holt, 1938.