ASD-TN-61-55

October 1961
Reprinted
April 1970

A Hierarchical Grouping Procedure Applied to a Problem of Grouping Profiles

By
Joe H. Ward, Jr.
Marion E. Hook

PERSONNEL LABORATORY
AERONAUTICAL SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
LACKLAND AIR FORCE BASE, TEXAS

Fred E. Holdrege, Col USAF
Chief

A. Carp
Technical Director

Personnel Laboratory
Aeronautical Systems Division (AFSC)

NOTICES

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the United States Government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.

Qualified requestors may obtain copies of this report from the Armed Services Technical Information Agency, Documents Service Center, Arlington 12, Virginia. Department of Defense contractors may obtain unclassified documents from ASTIA on request by stating their official need and citing the Defense contract involved.

ASTIA release to OTS not authorized.

Copies of ASD Technical Reports and Technical Notes should not be returned to the Aeronautical Systems Division unless such return is required by security considerations, contractual obligations, or notice on a specific document.

ASD-TN-61-55
October 1961

A HIERARCHICAL GROUPING PROCEDURE APPLIED TO A PROBLEM OF GROUPING PROFILES

by
Joe H. Ward, Jr.
Marion E. Hook

Project 7734; Task 773403

Personnel Laboratory
AERONAUTICAL SYSTEMS DIVISION
AIR FORCE SYSTEMS COMMAND
UNITED STATES AIR FORCE
Lackland Air Force Base, Texas

FOREWORD

A new approach to clustering is being used by Personnel Laboratory for the solution of a variety of problems. This approach is based on the concept that items should be grouped at each stage of a hierarchy so as to maximize payoff or minimize the cost of grouping in terms of some relevant criterion. Dr. Ward describes this concept mathematically in *Hierarchical Grouping to Maximize Payoff,* (WADD-TN-61-29). A unique feature of the technique is that the criterion function to be optimized is selected by the investigator and can be varied from problem to problem. Applications of the technique now programmed by Personnel Laboratory involve the clustering of (1) people to maximize their similarity in terms of measured characteristics, (2) jobs to minimize cross-training time when personnel are reassigned to new jobs within clusters according to established rotation policies, and (3) job descriptions to minimize errors in describing a large number of jobs with a limited number of descriptions. This paper describes an application to a problem of grouping test profiles so as to maximize the homogeneity of profiles within the same clusters, taking account of all profile variables and all groups at the same time.

ABSTRACT

This report describes an application of a hierarchical grouping procedure to a problem of grouping test profiles so as to maximize the homogeneity of profiles within the same clusters, taking account of all profile variables and all clusters at the same time. The data are 25 test profiles to which Sawrey, Keller, & Conger applied a different grouping technique. The results of the two grouping techniques are compared.

Desirable characteristics of the hierarchical grouping technique are demonstrated. Any measure of profile similarity can be used. It is unnecessary to specify in advance the number of groups to be formed or to select nucleus groups. The cost of grouping can be evaluated in terms of any criterion expressed as a functional relation, or objective function. The resulting hierarchical structure of *k* profiles is that which, at each stage (k, k-1, . . ., 1), least impairs the objective-function value. Thus the hierarchical grouping technique shows not only the order in which profiles must be grouped so as to yield the optimal value of the objective function when the number of profiles is systematically reduced, but also the costs of the grouping at each stage of the hierarchy.

A HIERARCHICAL GROUPING PROCEDURE APPLIED TO A PROBLEM OF GROUPING PROFILES[1]

Investigators often desire to group large numbers of persons, jobs, or objects into smaller numbers of mutually exclusive classes in which the members have similar characteristics. When the grouping is done in a manner that establishes a taxonomy of mutually exclusive clusters wherein each larger unit is a unique combination of the next-subordinate units, the clusters are called "hierarchical groups." Such groupings have proved particularly useful for classification purposes. For example, plants and animals may be hierarchically grouped with respect to genetic characteristics; library holdings grouped in terms of their contents to facilitate storage and retrieval of information; persons (or jobs) grouped in terms of specified characteristics for purposes of personnel administration. While grouping ordinarily results in some loss of information, it may generate new information as well as increase the efficiency with which large masses of data can be considered. Therefore the technique is applied in many situations.

Until recently, grouping has rarely been done by mathematical techniques. The procedures used required arbitrary or subjective decisions; optimally homogeneous groups were not formed, and the loss resulting from the grouping was not quantified. Much attention has been given to two technical problems, namely, the measurement of the relative similarity of profiles and the formation of groups of profiles. Many different ways of handling the first problem have been discussed in the past decade (Rao, 1952; Cronbach & Gleser, 1953; Helmstadter, 1957; Thorndike, Hagen, et *al.*, 1957; Sawrey, Keller, & Conger, 1960). Two approaches to the second problem were noted by Sawrey, Keller, & Conger. To the best of our knowledge, however, no one has applied a hierarchical grouping procedure to a problem of classifying profiles. As we shall show, this approach is advantageous when computer facilities are available.

Ward (1961) has mathematically described a procedure for forming hierarchical groups of mutually exclusive sets in a manner that satisfies any stated criterion of the investigator. The purpose of the present report is to describe an application of the technique to a problem of grouping profiles. The data for 25 profiles to which the general computer program has been applied are those given by Sawrey, Keller, & Conger (1960, p. 661). These authors compare the results of their technique of grouping profiles with the results of a Q-technique factor analysis. Hence the interested reader can examine the results of three different treatments of the same profile data.

METHOD

HIERARCHICAL GROUPING PROCEDURE

Objective function. The hierarchical grouping procedure is based on the premise that the most accurate information is available when each individual constitutes a group. Consequently, as the number of groups is systematically reduced, k, $k-1$, . . ., 1, the clustering of increasingly dissimilar individuals will yield less precise information. The extent of the inaccuracy associated with grouping can be quantified by a value-reflecting number, the objective function. This objective function may be any functional relation that reflects the investigator's criterion.

At each stage in the profile-grouping problem described here, the goal is to form a group such that the sum of the squared within-group deviations about the group mean of each profile variable is minimized for all profile variables in all groups at the same time. A satisfactory objective function for this purpose can be formulated and described mathematically, as follows.

---

[1] Manuscript released by the authors for publication as an ASD Technical Note in October 1961.

Given n individuals each of whom has been observed on p characteristics, we define the following vectors:

y = a vector, of dimension np (n times p), which has as elements the observed values of the p characteristics

and a set of np predictor vectors of the form

$x^{(r,s)}$ = a vector, of dimension np, in which the elements are *1s* if the corresponding elements of y came from the rth person on the sth characteristic; and *0s* otherwise

$[r = 1, \ldots, n;\ s = 1, \ldots, p]$

We can now express y as a linear combination,

$$y = \alpha_{(1,1)}x^{(1,1)} + \alpha_{(1,2)}x^{(1,2)} + \ldots + \alpha_{(r,s)}x^{(r,s)} + \ldots + \alpha_{(n,p)}x^{(n,p)} + e$$

where

e = a residual vector = the null vector, since y can be expressed without error as a linear combination of the vectors $x^{(r,s)}$

$[r = 1, \ldots, n;\ s = 1, \ldots, p]$

The purpose of the grouping operation, then, is to unite groups in a manner that reduces the number of predictor vectors $[x^{(r,s)}]$ by p at each iteration while keeping the elements of the error vector near zero. The objective function can be given by

$$z[i, j, k\text{-}1] = \sum_{i=1}^{np} e_i^2$$

where

e = the error vector that results from the reduction of k groups to k-1 groups by uniting group i with group j

Accordingly, if we define

y_{rsg} = the observation of the sth characteristic for the rth person in the gth group

$[r = 1, \ldots, n_g;\ s = 1, \ldots, p;\ g = 1, \ldots, k\text{-}1]$,

the value of the objective function, expressed as a sum of squared deviations (SSE), is given by

$$\begin{aligned} z[i,j,k\text{-}1] &= \sum_{s=1}^{p}\left\{\sum_{g=1}^{k-1}\sum_{r=1}^{n_g}(y_{rsg})^2 - \sum_{g=1}^{k-1}\left[\frac{1}{n_g}\Big(\sum_{r=1}^{n_g} y_{rsg}\Big)^2\right]\right\} \\ &= \sum_{s=1}^{p}\sum_{g=1}^{k-1}\sum_{r=1}^{n_g}\left[(y_{rsg})^2\right] - \sum_{s=1}^{p}\sum_{g=1}^{k-1}\left[\frac{1}{n_g}\Big(\sum_{r=1}^{n_g} y_{rsg}\Big)^2\right] \end{aligned}$$

Hierarchical grouping. Since the computer program yields a complete hierarchical structuring, it is unnecessary to specify in advance the number of groups to be formed or to select the nuclei of potential groups.[2] Given a total of k sets, or groups, this program insures their reduction to k-1 sets with the least possible impairment of the optimal level of the objective function at each stage in the grouping operation. The SSE reflecting the optimal level of the objective function is 0.00. This occurs in this application when there are 25 groups each containing one profile. The

---

[2] The mathematical description of the procedure given by Ward (1961) includes formulas for determining both the number of possible ways of forming groups and the number of distinguishable unions possible.

reduction to 24 groups (1 group having two profiles; 23 groups each having one profile) is made by considering all possible 25(24)/2 pairings of the 25 profiles and selecting that pairing for which the objective-function value is the smallest. Thus the first iteration results in pairing the two most homogeneous profiles in the array. Moreover the cost of this grouping is available in terms of the SSE associated with 24 groups.

Repetition of this process permits systematic reduction of the number of groups, 24, 23, . . ., 1. At the outset of the second and subsequent iterations, each group previously formed is treated as one unit, regardless of the number of profiles in the group. At the beginning of the second iteration there are 24 groups (of which 23 are the original groups each containing one profile); at the beginning of the third iteration there are 23 groups; and so on. Hence the grouping procedure routinely involves consideration of the effects of: (*a*) pairing each two of the remaining one-profile groups; (*b*) pairing each remaining one-profile group with each previously formed profile cluster; and (c) pairing each two previously formed profile clusters. The pairing selected is always that which reduces by one the number of groups while minimizing the objective function.

When the complete hierarchical solution has been obtained, the SSE values may be compared to ascertain the relative homogeneity of the groups formed at different stages in the process. A sharp increase in the objective function indicates that much of the classification system's accuracy has been lost by reducing the number of groups by one at this stage. Such information on the relative "costs" of different numbers of groups provides valuable guidance whenever it is necessary to decide upon the specific number of profile categories to be used for classification purposes.

PROFILE DATA

For convenience, the matrix of d^2s used to illustrate the hierarchical grouping procedure is reproduced here from Sawrey, Keller, & Conger (1960, p. 661). Table 1 (p. 4) summarizes the distance functions obtained for each pairing of 25 Conger-Wilson Designs Test profiles which were obtained from Air Force enlisted men. The d^2 for each pair has been computed by squaring the difference in scores on each profile element and summing these squares. The six profile elements are "designed to show the subject's relative preferences for a variety of formal art elements (e.g., warm vs. cool color, strong vs. weak contrast, symmetry vs. asymmetry, etc.)" (p. 662). Scores on each element have a possible range from 0 to 10.

The objective function used in the hierarchical grouping technique was computed from the d^2 matrix. This is readily done, for it can be demonstrated that if we identify any two persons in the *g*th group as r and r',

$$\frac{1}{2}\sum_{g=1}^{k}\frac{1}{n_g}\left[\sum_{r=1}^{n_g}\sum_{r'=1}^{n_g}\sum_{s=1}^{p}(y_{rsg}-y_{r'sg})^2\right]=\sum_{s=1}^{p}\sum_{g=1}^{k-1}\sum_{r=1}^{n_g}\left[(y_{rsg})^2\right]-\sum_{s=1}^{p}\sum_{g=1}^{k-1}\left[\frac{1}{n_g}\left(\sum_{r=1}^{n_g}y_{rsg}\right)^2\right]$$

where (as before)

r = person $[r=1, \ldots, n_g]$
g = group $[g=1, \ldots, k]$
s = characteristic $[s=1, \ldots, p]$
y_{rsg} = observation of the *s*th characteristic for *r*th person in the *g*th group.

RESULTS

The complete hierarchical structuring of the 25 profiles that minimizes the objective function at each stage is given in Table 2. This grouping was completed within approximately 15 minutes

(Text continues on page 6)

TABLE 1. Matrix of d^2s Between Persons[a]

	Identifying Number																								
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	-																								
2	131																								
3	77	42																							
4	36	131	91																						
5	29	128	82	5																					
6	35	108	86	5	4																				
7	37	88	82	55	50	46																			
8	18	71	39	36	29	29	19																		
9	96	59	45	84	85	83	37	42																	
10	53	34	34	35	34	24	38	17	37																
11	58	57	47	34	31	25	29	16	26	7															
12	72	61	3	98	85	95	81	40	48	47	56														
13	112	35	73	60	67	45	77	62	64	15	28	98													
14	2	147	99	40	35	39	35	24	106	63	66	94	122												
15	92	93	91	66	65	59	25	40	20	39	16	96	62	94											
16	134	3	39	130	131	111	105	78	70	37	66	60	36	152	112										
17	43	42	52	35	36	24	18	15	35	6	11	63	23	47	31	49									
18	57	112	118	71	64	56	8	31	57	52	35	119	89	51	23	135	30								
19	47	70	64	55	48	44	4	15	23	26	15	65	61	49	13	87	14	10							
20	51	54	66	25	28	14	32	23	51	6	11	81	15	55	37	59	4	40	26						
21	107	34	82	67	70	48	52	57	51	18	25	103	7	113	43	43	16	60	40	14					
22	103	30	62	63	60	42	70	53	63	12	23	83	5	115	59	35	22	80	52	16	8				
23	107	64	56	87	84	82	38	49	3	40	25	59	63	117	15	79	38	54	22	52	46	58			
24	41	78	84	51	46	38	2	19	41	30	23	87	61	39	23	95	12	6	4	22	38	54	40		
25	40	51	37	40	31	29	15	10	24	11	8	40	42	48	24	62	9	29	7	17	31	31	23	13	

[a] From Sawrey, Keller, & Conger (1960, p. 661).

Corrected data for Table 2 (p.5), ASD-TN-61-55

(From Ward, J.H., Jr., & Hook, M.E. Application of an hierarchical grouping procedure to a problem of grouping profiles. <u>Educational and Psychological Measurement</u>, 1963, Vol. XXIII, No. 1, 69-81.)

TABLE 2

Structure of 25 Profiles Resulting from Hierarchical Grouping Procedure
Number of Groups

Profile No.	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14	14
4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6
7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7	7
24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24	24
19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19	19
18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18	18
8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8
11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11	11
25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25	25
10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10	10
17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17	17
20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20	20
9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9	9
23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23	23
15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15	15
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16	16
13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13	13
22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22	22
21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21	21
3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3
12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12	12
Objective Function (SSE)	605.9	439.0	323.5	237.6	186.9	144.4	104.0	65.0	53.2	42.0	34.7	29.5	25.3	21.3	18.0	15.3	12.8	10.5	8.5	6.5	5.0	3.5	2.0	1.0	0.0
Δ 1		166.9	115.5	85.9	50.7	42.5	40.4	39.0	11.8	11.2	7.3	5.2	4.2	4.0	3.3	2.7	2.5	2.3	2.0	2.0	1.5	1.5	1.5	1.0	
Δ 2		51.4	29.6	35.2	8.2	2.1	1.4	27.2	0.6	3.9	2.1	1.0	0.2	0.7	0.6	0.2	0.2	0.3	0.0	0.5	0.0	0.0	0.5		

on an IBM 650 Computer. With the exception of the last three rows, Table 2 is a copy of the computer's printout of the final results of the hierarchical grouping program. If the table is read from right to left, the columns show the order in which profiles are clustered so as to minimize the error term. Column 24 (number of groups = 24), for example, shows that Profiles 1 and 14 are the first grouped, hence the most homogeneous; Column 23 shows the next two profiles to be grouped with least impairment of the objective function are Profiles 7 and 24.

The last three rows of Table 2 reflect the cost of each grouping. These values are readily calculated from information in a computer printout (not shown here). Values of the objective function (SSE row of Table 2) indicate the over-all accuracy of the classification system at each level in the hierarchy. For example, when all 25 profiles are in one group, the SSE value is 605.0; at the seven-group stage, the SSE value of 102.0 reflects considerable within-group homogeneity; but the SSE value of 63.5 for eight groups indicates that these groups have considerably greater homogeneity than the seven groups. The rows identified as $\Delta 1$ and $\Delta 2$ facilitate comparisons. The $\Delta 1$ entry (38.5) between Columns 7 and 8, for example, shows the change in the error that results from merging the cluster containing Profiles 2 and 16 with that containing Profiles 13, 22, and 21. The $\Delta 2$ entries, e.g., 26.5 below the $\Delta 1$ entry of 38.5, show the differences between the $\Delta 1$ entries and reflect the acceleration of error.

DISCUSSION

In this application of the general procedure for hierarchical grouping, the objective at each stage has been to form a grouping that minimizes the sum of the squared within-group deviations about the group mean of each profile variable for all profile variables for all groups at the same time. In discussing their grouping procedure, Sawrey, Keller, & Conger speak of forming profile groups "in which the group members are similar to each other and, at the same time, dissimilar from the members of all other groups" (p. 660). Since their method involves selecting nucleus groups, it appears that, for a specific number of groups, they sought a solution that would maximize the between-group sum of squares and minimize the within-group sum of squares for all profile elements. The hierarchical grouping procedure described here minimizes the within-group and maximizes the between-group sums of squares at each level in the hierarchical structure. However, this does not necessarily result in the minimal within-group and maximal between-group sums of squares for a specific number of groups.

It is interesting to examine the groups that result from the two different techniques of grouping. Let us first compare the seven original "dissimilar nucleus groups" established by the Sawrey-Keller-Conger (SKC) method (Table 3) with the initial seven groups produced by the hierarchical technique (Column 17, Table 2). Five of the seven groups in each array are the same at this stage. In addition, two of the three profiles in SKC Group VI have been grouped in the hierarchical process; the third profile is added on the next iteration. The major difference in the two arrays is that the hierarchical technique groups Profiles 17 and 20 before it forms SKC Group V (which appears after two more iterations). Otherwise, the two grouping procedures yield practically the same results at this stage.

With the Sawrey-Keller-Conger method of selecting arbitrary d^2 limits as criteria, the first additions to the original disparate groups are Profile 21 (Group V) and Profile 18 (Group VII). The same profiles are added to the same groups in the hierarchical grouping operation (Columns 12 and 11, Table 2). When $d^2 < 18.17$, the Sawrey-Keller-Conger technique clusters Profiles 10 and 20 with Group V. At the same time, Profiles 17 and 25 are left ungrouped because each now can be placed in either one of two groups. This treatment of Profiles 17 and 25 obviously affects the centroids of the groups to which these profiles might have been added and the subsequent classification of the three remaining profiles. In contrast, the hierarchical grouping technique soon

TABLE 3. Results of Grouping 25 Profiles by Sawrey-Keller-Conger Method[a]

Original 7 Dissimilar Nucleus Groups		Order in Which Profiles are Added to Dissimilar Nucleus Groups[b]				
SKC Group	Profile Number	$d^2 < 6.06$	$d^2 < 8.08$	$d^2 < 12.12$	$d^2 < 18.17$	$d^2 < 24.23$
I	1, 14					
II	16, 2					
III	3, 12					
IV	9, 23					
V	13, 22		21		10, 20	
VI	6, 4, 5					
VII	24, 7, 19		18			

Note. — "Intermediate" profiles not included in final groups: 8, 11, 15, 17, 25.

[a] Contents of this table extracted from Sawyer, Keller, & Conger (1960, pp. 664-667).

[b] With $d^2 < 18.17$ limit on maximum distance from group centroid, Profile 17, which could be added to both Group V and Group VII, is designated an "intermediate" profile and not added to either cluster. At this same limit, Profile 25, which could go into either Group IV or Group VII, is designated an "intermediate" profile. With $d^2 < 24.23$ limit, Profiles 8, 11, and 15 are classified as "intermediate" profiles.

clusters Profile 17 with 20 and Profile 11 with 25. This results in two "isolates" (Profiles 8 and 15) at the 11-group stage, which is comparable to the final array with five "intermediates" (Profiles 8, 11, 15, 17, and 25) produced by the Sawrey-Keller-Conger technique.

Much additional information is available in Table 2. At the 18-group stage, where Profiles 13 and 22 (SKC Group V) as well as Profiles 10, 17, and 20 are each a one-profile group, the SSE value is 10.5. The pairing of Profiles 17 and 20 at the next stage results in a small SSE increment (2.0). Note, however, the sharp rise in the error term when the number of groups is reduced from 12 groups (SSE = 28.5) to 9 groups (SSE = 51.5). Detailed information on specific profiles is easily extracted. Uniting Profiles 17 and 20, for example, produces a smaller increment in error than grouping Profile 4 with 5 and 6 (SKC Group VI) or Profile 13 with 22 (SKC Group V); Profile 10 is more homogeneous with the 17-20 cluster than it (or Profile 21) is with the 13-22 cluster. Furthermore with complete hierarchical structuring Profiles 11 and 25 are paired before Profiles 21 and 18 are added to their respective clusters. As illustrated here, valuable insights can be derived from the complete hierarchical structure of the profiles that results from systematically reducing the number of groups, one by one, from *k* to 1 in a manner that least impairs the objective function.

This grouping technique has several desirable features. Once the objective function has been established, the computer program for hierarchical grouping requires few arbitrary decisions. It eliminates the need to define nucleus groups or to set limits to specify the order in which profiles will be added to nucleus groups. This grouping technique is systematic and replicable. The computations do not require an undue amount of machine time. Given a matrix containing measures of the similarity of profiles, the complete hierarchical grouping of 100 profiles can be accomplished in about one hour if the investigator has access to equipment such as an IBM 650 Computer. Since the matrix can contain any measure of profile similarity and since the objective function can reflect any criterion specified by the investigator, the general computer program is applicable to many profile-grouping problems.

The evaluation of the effects of grouping on an objective function is a valuable feature of the hierarchical grouping technique. When the investigator can examine both the content of clusters and the error associated with grouping for each stage in the hierarchical structure, he has complete information for all levels of profile homogeneity represented in the sample. This information provides useful guidance for handling many problems, e.g., How many groups should be used for specified classification purposes? Which groups should be compared to evaluate the effects of different treatments? In some situations, use of relatively large rather than small numbers of groups may not improve the accuracy of the classification system enough to justify associated increases in administrative costs or time delays. Under other circumstances, a sudden rise in the costs of grouping may indicate that it is not appropriate to reduce the number of groups beyond a certain level. These are but a few of many possible examples. The specific use made of the information on grouping costs is dictated by the problem under study.

SUMMARY

The application of a hierarchical grouping procedure to a problem of grouping profiles is described. The matrix for 25 profiles (based on an art preference test) to which the computer program was applied was taken from a published report which describes a different grouping technique and a Q-technique factor analysis of the data. The profile clusters obtained by that grouping technique are compared with those obtained by the hierarchical grouping technique.

The hierarchical grouping technique has a number of desirable characteristics. Any measure of profile similarity may be used in the matrix describing the profiles. There is no need to specify in advance the number of groups to be formed, to select nucleus groups, or to set arbitrary limits for use in adding profiles to groups.[3] Grouping can be based on any criterion expressed as an objective function. The objective function can be quantified in terms of SSEs. The resulting hierarchical structure of the *k* profiles is that which, at each state (*k*, *k*-1, . . . , 1), least impairs the objective function. Hence the hierarchical grouping technique shows not only the order in which profiles are grouped so as to yield an optimal value of the objective function when the number of profiles is systematically reduced from *k* to 1, but also the costs of each grouping.

REFERENCES

Bottenberg, R.A. & Christal, R.E. *An iterative technique for clustering criteria which retains optimum predictive efficiency.* Lackland Air Force Base, Texas: Personnel Laboratory, Wright Air Development Division, March 1961. (WADD-TN-61-30; ASTIA Document No. AD-261 615)

Cronbach, L.J. & Gleser, Goldine C. Assessing similarity between profiles. *Psychol. Bull., 1953*, **50**, 456-473.

Helmstadter, G.C. An empirical comparison of methods for estimating profile similarity. *Educ. Psychol. Measmt*, 1957, **17**, 71-82.

Rao, C.R. *Advanced statistical methods in biometric research.* New York: Wiley, 1952.

Sawrey, W.L., Keller, L., & Conger, J.J. An objective method of grouping profiles by distance functions and its relation to factor analysis. *Educ. psychol. Measmt*, 1960, **20**, 651-674.

Thorndike, R.L., Hagen, Elizabeth P., Orr, D.B., et al. *An empirical approach to the determination of Air Force job families.* Lackland Air Force Base, Texas: Air Force Personnel and Training Research Center, August 1957. (*Technical Report* AFPTRC-TR-57-5, ASTIA Document No. AD-134 239)

Ward, J.H., Jr. *Hierarchical grouping to maximize payoff.* Lackland Air Force Base, Texas: Personnel Laboratory, Wright Air Development Division, March 1961. (WADD-TN-61-29; ASTIA Document No. AD-261 750)

---

[3] Dr. William L. Sawrey was kind enough to review our manuscript and permit use of the Sawrey-Keller-Conger data. In a personal letter commenting on the conclusions, he rightly points out that neither the computerized procedure nor his provides the actual decision on how many groups to use.