11-615 no.85-57 c.2

-

Statistique Canada



Methodology Branch

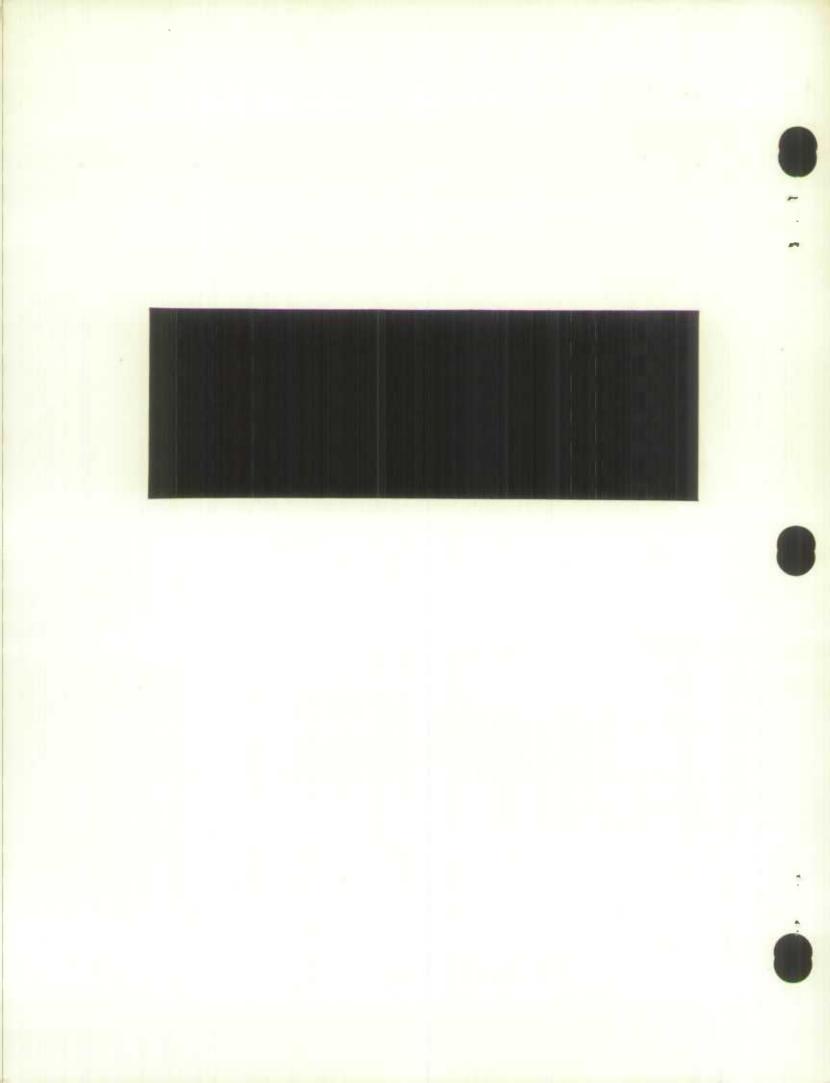
Direction de la méthodologie

Institutions & Agriculture Survey Methods Division

5

Canada

Division des méthodes d'enquêteinstitutions et agriculture



WORKING PAPER No. IASM-85-057E

CAHIER DE TRAVAIL No. DMEIA MÉTHODOLOGIE

METHODOLOGY BRANCH

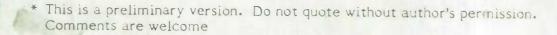
50412

Cir

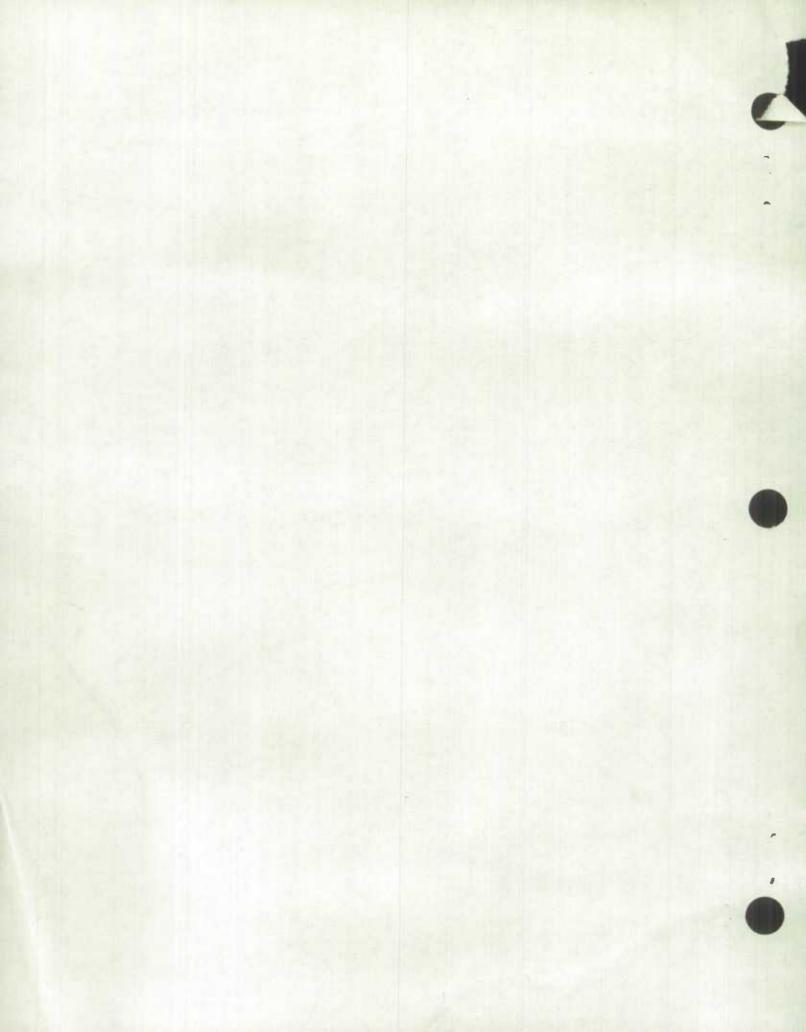
A CLUSTER ANALYSIS OF ACTIVITIES OF DAILY LIVING FROM THE CANADIAN HEALTH & DISABILITY SURVEY

by

D.A. Binder & G. Lazarus May, 1985



16



Abstract

The Canadian Health and Disability Survey, administered as a supplement to the Canadian Labour Force Survey in October 1983, collected data by means of a screening questionnaire and a follow-up questionnaire to those screened in. The data from the screening questionnaire, consisting of a set of activities of daily living, were used to group respondents according to identifiable characteristics. A description of the groups of respondents is provided and the potential for development of a disability severity scale is explored.

Résumé

L'Enquête sur la santé et les invalidités au Canada, qui constituait un supplément à l'Enquête sur la population active d'octobre 1983, est composée d'un questionnaire de sélection et d'un suivi auprès des répondants sélectionnés. Le questionnaire de sélection est un ensemble de questions relatives aux difficultés à éffectuer certaines activités de la vie quotidienne. Ces variables sont utilisées pour regrouper les répondants selon certaines caractéristiques. On donne une description des groupes formés et on analyse le potentiel du regroupement pour le développement d'une mesure de gravité de l'incapacité.

TABLE OF CONTENTS

1. BACKGROUND

- 1.1 Introduction
- 1.2 Questionnaire
 - 1.2.1 Screening section
 - 1.2.2 Follow-up section
- 1.3 Index of Severity

2. CLUSTERS

- 3. CHARACTERIZING THE CLUSTERS
 - 3.1 Introduction
 - 3.2 Preliminary Efforts
 - 3.3 Further Developments

4. EVALUATION

- 4.1 Methodology
- 4.2 Results
 - 4.2.1 Components
 - 4.2.2 Mean scores
 - 4.2.3 The scales
- 5. CLOSING REMARKS

APPENDIX

A DIAGRAM AND TABLES

B A TECHNICAL REPORT CONCERNING THE CLUSTER ANALYSIS

REFERENCES

I. BACKGROUND

1.1 INTRODUCTION

In response to a need for data on disabled persons in Canada, Statistics Canada undertook a program to create a disability database¹. The Canadian Health and Disability Surveys (CHDS) were administered as supplements to the Canadian Labour Force Survey (LFS) in October 1983 and June 1984. In both cases, separate questionnaires were administered to children and to adults. In the October survey, the adult questionnaire was administered to everyone in the LFS frame (which includes about 97% of the Canadian population aged 15 or more). In June, the adult survey was restricted to those aged 15 to 64 from the six provinces with the smaller sample sizes in October (i.e. Newfoundland, Prince Edward Island, Nova Scotia, New Brunswick, Manitoba and Saskatchewan). Children from all provinces were surveyed in both October and June.

This report concentrates on work which utilized only the data from the adults questionnaire in October 1983. This survey obtained 92,945 adult respondents from approximately 47,000 households.

1.2 QUESTIONNAIRE

The questionnaire included first a screening section and then a follow-up section which was administered only to those individuals who were selected by the screening section.

1.2.1 Screening Section

The screening section consisted of nineteen items - - - seventeen activities of daily living, an activity limitation item and an item about mental handicap. The activities of daily living (ADL's) are a set of activities which any person is required to perform during the course of his/her regular living pattern. The set used here was a modified version of those developed by the Organization for Economic and Co-operative Development (OECD) and has been utilized by several other countries².

1 Obstacles Report

² International Classification of Impairments, Disabilities, and Handicaps, World Health Organization. Geneva. Switzerland.

The ADL's are listed below along with the questionnaire item identification.

- Alo Walking 400 Metres
- All Walking up and down stairs
- A12 Carrying 5 kg. object for 10 metres
- A13 Moving from one room to another
- A14 Standing for long periods
- A15 When standing, bending down to pick up object
- A16 Dressing and undressing
- A17 Getting in and out of bed
- A18 Cutting own toenails
- A19 Using fingers to grasp or handle
- A20 Reaching
- A21 Cutting own food
- A22 Reading newsprint
- A23 Seeing clearly a face across the room
- A24 Hearing conversation with another person
- A25 Hearing conversation with two or more persons
- A26 Speaking and being understood

An example of the wording of these questions in the screening section of the questionnaire is as follows:

A20 Does have any trouble reaching?

The activity limitation item (A27) concerned limitation "in the kind or amount of activity he/she can do at home, at work or going to school because of a long-term physical condition or health problem".

The final item in the screen section (A28) concerned mental handicap.

It should be noted that the survey was concerned with long-term conditions or health problems - those that had lasted or were expected to last more than six months (excluding pregnancy).

An individual was screened in if he/she had trouble with at least one of the ADL's, the major activity limitation item or had a mental handicap. (Proxy responses were required for mentally handicapped individuals).

1.2.2 Follow-up Section of the Questionnaire

The follow-up section of the questionnaire was completed for individuals selected by the screening section. This section included an item which sought to determine if the respondent was completely unable to perform the ADL('s) he/she had trouble with. Other segments of the follow-up questionnaire pertained to: nature of the disability (related to trouble seeing or reading, trouble hearing, trouble speaking and being understood, and mobility); problems related to the ability to work or to the workplace; obstacles to education and availability of special educational facilities; problems related to local and long-distance travel; and problems in current residence and special facilities.

1.3 INDEX OF SEVERITY

One area of analytic interest is the development of an index of severity of disability. Our analysis is based on only the screen items since they have been used in a number of other surveys, whereas the follow up questions are specific to the CHDS.

One of our first steps in our efforts to establish this severity index was to cluster the screened-in respondents in such a way that those respondents in the same cluster tend to have a similar pattern of screening data.

In this paper we present the results of this cluster analysis, an interpretation of these results and an evaluation of the final clusters.

This analysis is described in Section 2 with technical details provided in Appendix B. The results of the cluster analysis and a discussion on the development of a severity scale are described in Section 3. An evaluation is given in Section 4. Closing remarks are provided in Section 5.

2. CLUSTERS

This section presents a description of the procedures used in the development of the clusters. The technical details may be found in Appendix B.

The cluster analysis was a procedure which grouped together those screened in respondents with similar but not necessarily identical 'profiles'. For our purposes, a respondent's profile consisted of the yes (has trouble) / no (does not have trouble) responses to the seventeen ADL's, and positive/negative responses to the major activity limitation item and the mental handicap item in the screening section of the questionnaire. Figure 2.1 illustrates the development of the clusters with the classification based on specific values for certain screening section items.

The chart may be read as follows. "Axx=1" indicates "yes - has trouble" with item xx in the screening section of the questionnaire while "Axx=0" indicates "no - does not have trouble" with item xx. Hence A16=1 implies that the respondent has trouble dressing and undressing himself/herself while A15=0 implies that the individual has no trouble when standing, bending down and picking up an object from the floor. (Refer to Section 1 for definitions of A10 to A28). The number in the top left hand corner of certain rectangles indicates the final cluster number (used as cluster identification in Table 2.1 and Table 2.2 of the next section). Since clusters were a subset of the total number, referring to the line number used in Table 2.2. It can be seen, for example, that L9 splits according to A22 into L10 (cluster 5 with profile identification consisting of A10=0, A25=1 and A22=1) and L11. The other number in the rectangle indicates the number of individuals for which "Axx=0" is true given all previous splits to that point. Hence there are 203 individuals in cluster number 5.

Figure 2.1 shows that six of the nineteen screening items are not used in the process of classifying respondents. These are All, Al3, Al8, A20, A23 and A24. This is because of their strong correlation with some combination of the other 13 items.

Table 2.1 gives the percentage of persons within each of the final clusters who responded "has trouble" to each of the nineteen items in the screening section of the questionnaire. The first column on the left gives the cluster identification number. The next column provides the number of individuals in the cluster. The remaining columns

give the unweighted percentages of individuals in the specified cluster who selected each of the screening items.

Table 2.2 presents these percentages for each of the "boxes" in Figure 2.1. There are also two columns called "Parent" and "Children". The "Parent" denotes the line number from which the new partition is derived. The "Children" give the line numbers of the next split in Figure 2.1, where applicable. For example, lines 3 and 4 are "Children" of line 2 because lines 3 and 4 are further splits of the partition in line 2. Conversely, line 2 is the "Parent" of lines 3 and 4.

The symbols "U" and "Z" are used to show how the groups are defined. The symbol "U" means that the group is defined through that variable being one i.e. 100% by definition. The symbol "Z" is used when the defining screening section item is zero i.e. 0% by definition. Occasionally, there is an asterisk next to a value of 0.0 or 100.0 in these tables. This means that no one or everyone, respectively, reported positively to that item.

Table 2.2 is useful in that it exhibits the effects of the splits on the other screening section items of the groups in an orderly fashion. Table 2.1 shows screening section item incidence for each cluster.

3. CHARACTERIZING THE CLUSTERS

3.1 INTRODUCTION

The previous section described the development of the clusters. This section explores the ways and means of identifying the clusters and attempts to rank the clusters according to severity of disability. In this section we introduce the concepts of "umbrella" group and "trouble orientation".

3.2 PRELIMINARY EFFORTS

We begin here with a crude measure of severity; namely average number of ADL's, E(NADL), where the ADL's include items A10 to A26.

Table 3.1 presents the total E(NADL) for each cluster as well as the E(NADL) of each activity orientation for each cluster. With reference to the list of ADL's given in Section 1.2, A24 and A25 refer to hearing troubles, A22 and A23 are related to vision troubles, A10, A11, A12 and A14 are oriented toward troubles with mobility and the remaining ADL's possess an agility trouble orientation. This demarcation between mobility and agility could be criticized on a number of grounds, but it proved useful for our purposes. Section 4 presents some ideas which are more objective in nature.

The clusters can now be characterized according to their orientation. As a point of departure for further study, a cluster will be considered to have a hearing trouble and vision trouble orientation (HV) when the average number of the hearing trouble ADL's exceeds one and the average number of vision trouble ADL's exceeds one. For example, for cluster 5, the average number of the hearing troubles is 1.596 and the average number of the vision trouble ADL's is 1.463. Cluster 5 is therefore a member of this HV "umbrella" group. A study of Table 3.1 reveals that cluster 2 belongs to the HV group as well.

A cluster will be considered to have a hearing trouble orientation (H) when the average number of the hearing trouble ADL's exceeds one but the average number of the vision trouble ADL's is less than one. Clusters 1, 3, 4, 6 and 7 satisfy these requirements.

A cluster will be considered to have a vision trouble orientation (V) when the average number of the vision trouble ADL's exceeds one but the average number of the hearing trouble ADL's is less than one. Clusters 9, 12, 13 and 21 satisfy these requirements.

Two clusters, 17 and 24, are singled out because the major troubles are either speaking and being understood (A26) or mental handicap (A28). For reference purposes, these clusters will comprise a special (S) "umbrella" group.

We now turn to the classification of the remaining eighteen clusters.

A cluster will be considered to have a mobility trouble and agility trouble orientation (MA) when the average number of the mobility ADL's exceeds two and the average number of the agility ADL's exceeds two.

- 6 -

For example, for cluster 15, the average number of the mobility trouble ADL's is 3.776 and the average number of the agility trouble ADL's is 2.941. Cluster 15 is therefore a member of this MA "umbrella" group. A study of Table 3.1 reveals that clusters 8, 10, 11 and 14 also belong to the MA group.

A cluster will be considered to have a mobility trouble orientation (M) when the average number of the mobility trouble ADL's exceeds two but the average number of the agility trouble ADL'S is less than two. Clusters 16, 18, 19 and 20 satisfy these requirements.

The agility "umbrella" group consists of cluster 22 only. It satisfies the requirements that the average number of the mobility trouble ADL's is less than two but the average number of the agility trouble ADL's exceeds two. Clusters 23, 25, 26, 27, 28 and 29 comprise an "umbrella" group which we shall call "neither" (N). These clusters do not show signs of significant troubles in total or in any single orientation.

3.3 FURTHER DEVELOPMENTS

Table 3.2 presents the clusters according to the umbrella group composition. The clusters are listed arbitrarily in order of cluster number within umbrella group. The ID column on the right-hand side of the table is new, presenting a slightly modified version of the partitions discussed in the previous section. The idea behind the adjustment is that clusters in the HV, H, V and S groups which satisfied mobility or agility requirements should be labelled as such. Hence, cluster 2 is HVMA to recognize that the cluster shows a strong orientation toward mobility and agility troubles as well as to the hearing and vision troubles identified earlier. Cluster 5, on the other hand, is labelled HVN because the mobility and agility requirements match the N orientation.

The remaining entries in the ID column of Table 3.2 can be interpreted in a similar fashion. When more than one cluster is identified by the identical letter/letter-combination they are ordered by decreasing E(NADL) within umbrella group; e.g. M1, M2, M3 and M4.

Thus, Table 3.2 presents a preliminary version of an incomplete ordering of the clusters. The clusters can be compared using E(NADL) within "umbrella" groups. In some cases, differences in severity may be attributed to a broader orientation of troubles while in others the scale is strictly within a single orientation.

4. EVALUATION

In the previous section, we proposed an incomplete ordering of the clusters according to the orientation of troubles with ADL's. In this section, we evaluate this proposal.

4.1 METHODOLOGY

The evaluation was based on a subset of screened in cases, but using more information per case with the addition of responses to questions of the form:

(B101) Is ... completely unable to walk 400 metres without resting?

This line of questionning was used for each of the ADL'S, A10-A26. The evaluation involved only the 11,412 individuals who were screened in and who responded to these questions.

These completely unable items were coded with "1" when the individual indicated that he/she was completely unable to perform the specified ADL, otherwise (able or item non-response), a "0" was coded.

The means were obtained for the nineteen screening items and seventeen follow-up items for each cluster. The means for the completely unable items were then multiplied by the ratio of the overall average number of ADL's to the overall average of completely unable items in order to scale them consistently and to avoid the scaling problems associated with principal components analysis.

Principal components were obtained using the nineteen screening section and seventeen follow-up item means as variables, using the "clusters" as observations and weighting according to cluster size. The clusters were then ordered according to each of the first four principal component loadings.

The final stage involved the pooling of cluster cases according to "umbrella" group membership and finding the means of the first four principal component loadings for each of the eight "umbrella" groups, where the weights were the numbers of members in the "umbrella" groups.

4.2 RESULTS

The evaluation of the incomplete ordering of clusters is presented in two stages. In the first stage, we examine the principal components and attempt to label them according to the loadings. We also explore the "umbrella" group construct in terms of the principal component means. In the second stage, we examine the ordering of the clusters according to the first four principal components.

4.2.1 Components

The first four principal component loadings for the nineteen screening section items and the seventeen follow-up items are presented in Table 4.1. They explained just over seven-eighths of the total variance and appeared to be most useful for our purposes.

The loadings of the first principal component are positive on all but four items (A24, A25 and B241 are hearing oriented, A28 is mental handicap). The negative loadings are close to zero. This first component appears to be an overall measure of strength. The first principal component explained nearly 66% of the total variance and is denoted as "OVERALL".

There are negative loadings on A10, A11, A12, A14 and A15 of the second component. The loading for A15 is nearly zero, however. Loadings are positive for ADL's with an agility-trouble orientation as well as for hearing-trouble and vision-trouble orientations. It appears then that this component polarizes mobility trouble against agility, hearing and vision troubles. The second component is labelled "M/AHV".

The third principal component has positive loadings for mobility and hearing oriented ADL's and negative loadings for agility and vision oriented ADL's. This third component is denoted "MH/AV".

The fourth principal component has positive loadings for mobility and vision oriented ADL's and negative loadings for agility oriented ADL's. This fourth component is designated "MV/A".

4.2.2 Mean Scores

Table 4.2 presents the average deviations of the principal component loadings from the overall mean loadings for each of the eight "umbrella" groups. We can now check to see if the incomplete ordering presented earlier is consistent with the results from the principal components analysis. We note the following observations from Table 4.2.

- i) The mobility/agility "umbrella" group has the highest deviation on the first principal component "overall", while the "umbrella" group "neither" has the lowest deviation. The deviation for the hearing/vision group is positive as is the mean for the vision group. The hearing group deviation is negative, however, evidence that hearing-oriented troubles are inversely-related to severity of disability. There may be an inclination to draw the same kind of conclusion with respect to agility-oriented troubles. It is observed that the mobility/agility and mobility groups have positive deviations while the agility "umbrella" group has a negative deviation. However, in this case, the result is somewhat ambiguous because the agility-oriented ADL's included speaking trouble (A26), a so-called 'special' trouble area and it is clear indeed that the special "umbrella" group has a negative deviation for the first principal component.
- ii) The second component set mobility-oriented troubles (-) against agility, hearing and vision-oriented troubles (+). Positive deviations are recorded for the hearing/vision, hearing, vision and agility "umbrella" groups while negative deviations are associated with the mobility/agility, mobility and neither groups, as expected. The deviation for the special groups is nearly zero.
- iii) The third component set mobility-oriented and hearing-oriented troubles (+) against agility-oriented and vision-oriented troubles (-). Again, the results are consistent.
- iv) The fourth principal component set mobility and vision-oriented torubles (+) against agility-oriented troubles (-). The results are again consistent with the umbrella-group construct.

4.2.3 The scales

Table 4.3 shows the ranks of the clusters according to the first four principal component loadings and E(NADL). Recall that the component loadings are for 11,412 cases and utilize follow-up information while the E(NADL) scale is based on 12,907 cases and uses screening section information only.

The cluster ranking according to principal components was done as follows. The component representing overall strength (PRIN1) ranked clusters from highest to lowest scores. The ranking of clusters on PRIN2 tended to put clusters with mobility-oriented troubles at the bottom end as opposed to clusters with agility, hearing or vision oriented troubles which were ranked higher up on this scale. The ranking of clusters on PRIN3 tended to put clusters with mobility or hearing troubles at or near the bottom of the scale while clusters with agility or vision-oriented troubles were ranked higher. Finally clusters with agility-oriented troubles were ranked higher. Given the bipolar nature of components 2, 3 and 4, it was necessary to make an arbitrary decision as to a trouble orientation scale. As cluster 8 had shown itself to be highly severe according to the E(NADL) scale, it was determined that cluster 8 should be similarly ranked along the other scales.

For most clusters, the rankings fluctuate over a wide range. This reflects the nature of the criteria upon which the scales were based. The first principal component, which provides an overall measure of strength, may be the most suitable candidate for ranking the clusters. Firstly, it incorporates the screening section information used in the development of the E(NADL) measure. As a result, the rank orderings provided by the PRIN1 and E(NADL) scales are quite similar. The additional follow-up information used in the construction of this component leads us to believe that PRIN1 is better than other scales such as E(NADL). It is worth noting that the ranking was done on all 29 clusters and depicted in Table 4.3 on an "umbrella" group basis. The "umbrella" group information was not incorporated into the principal components analysis, however.

5. CLOSING REMARKS

A clustering technique was employed to group screened-in individuals according to similar screening section profiles. The clusters were then ordered according to the information contained in the screening section of the questionnaire (the incomplete ordering based on E(NADL) and presented in Table 3.2) and finally according to information contained in the screening and follow-up sections of the questionnaire (the PRIN1 scale presented in Table 4.3). This last scale is deemed presently to be the most suitable of those considered here. However, it could be argued that no single index of severity exists and in fact the severity index should be defined as a 4-dimensional scale corresponding to our principal components.

APPENDIX A

DIAGRAM AND TABLES



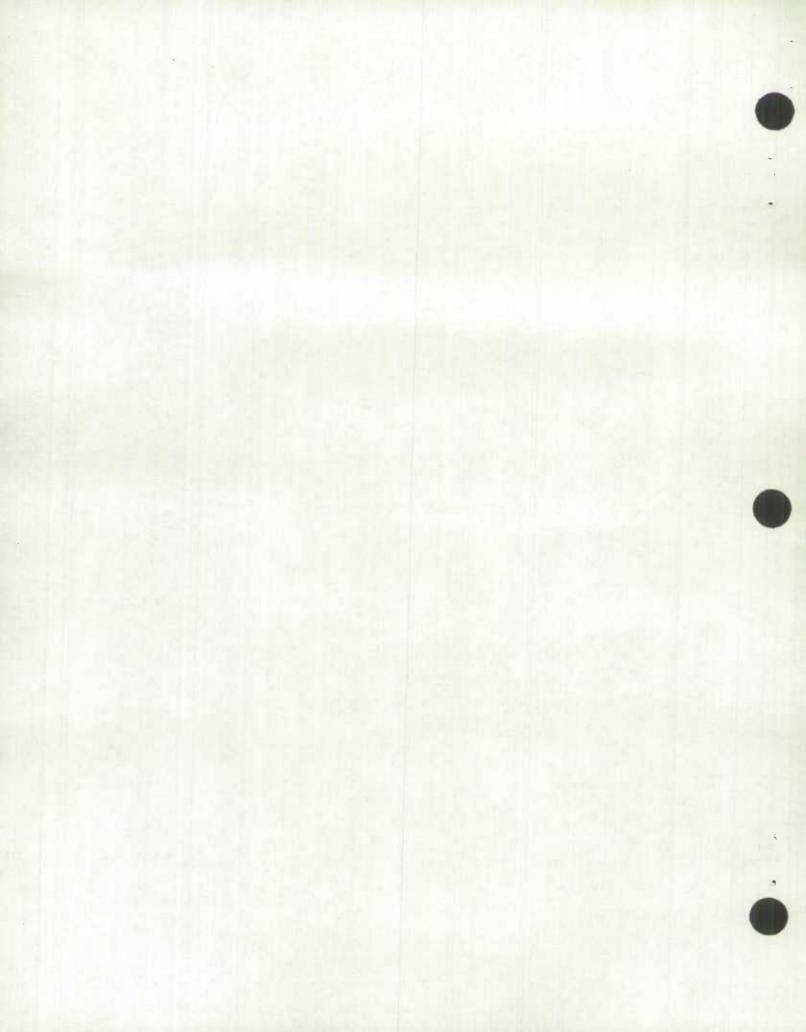
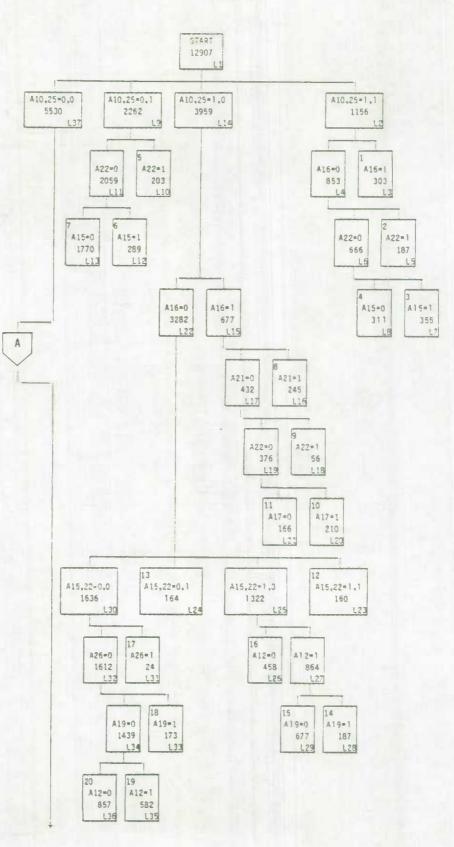
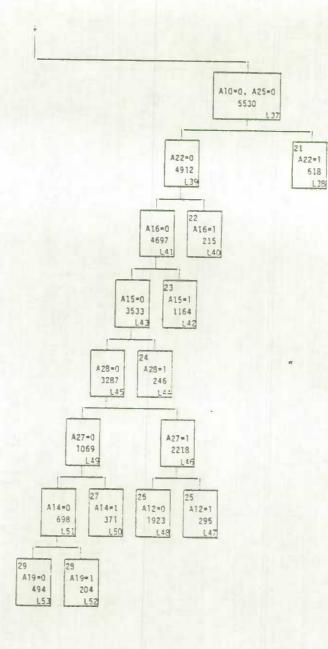


FIGURE 2.1





luster	Size	A10	All	A12	A13	A14	A15	A16	A17	A18	A19	A20	A21	A72	A23	A24	A25	A26	A27	A28
1	303	100.0	92.7	79.9	59.7	89.8	85.5	100.0	62.7	86.8	60.1	63.7	\$2.2	38.6	27.1	73.3	100.0	23.4	94.4	6.
2	187	100.0	77.0	63.1	16.0	77.0	55.6	0.0	11.2	46.5	31.0	35.3	11.8	100.0	50.0	71.7	100.0	9.6	85.0	1.
3	355	100.0	85.1	66.5	19.4	75.8	100.0	0.0	15.8	49.6	26.5	34.6	5.9	0.0	3.4	63.7	100.0	2.5	88.7	1
4	311	100.0	65.6	36.7-	6.4	55.9	0.0	0.0	4.5	21.5	17.7	16.4	1.9	0.0	1.6	57.9	100.0	2.6	73.3	1
5	203	0.0	18.7	18.2	3.4	25.6	24.6	4.9	6.9	21.7	20.7	17.7	8.4	100.0	46.3	59.6	100.0	12.8	55.7	7
6	289	0.0	36.3	23.2	4.8	49.5	100.0	11.8	16.6	28.4	21.1	24.9	3.5	0.0	1.4	50.9	100.0	4.2	71.3	1
7	1770	0.0	9.2	5.3	0.3	10.8	0.0	1.1	0.9	4.4	7.1	4.6	0.6	0.0	1.3	60.5	100.0	5.6	26.3	1
8	245	100.0	94.7	88.6	67.3	93.9	89.0	100.0	74.7	94.7	84.0	78.4	100.0	32.6	16.7	1.2	0.0	32.2	96.3	9
9	56	100.0	92.9	82.1	55.4	89.3	91.1	100.0	58.9	87.5	30.4	50.0	0.0	100.0	30.4	5.4	0.0	10.7	100.0	5
10	210	100.0	95.7	01.0	55.7	91.9	93.8	100.0	100.0	85.2	33.3	- 55.2	0.0	0.0	0.5	0.0	. 0.0	2.4	89.0	1
11	166	100.0	92.2	71.7	21.1	83.7	74.1	100.0	0.0	59.0	28.9	45.8	0.0	0.0	1.8	0.6	0.0	3.0	90.4	0
12	160	100.0	91.9	71.3	25.0	81.3	100.0	0.0	16.9	58.1	31.9	39.4	7.5	100.0	45.6	4,4	0.0	5.0	93.1	1
13	164	100.0	61.0	48.8	4,3	55.5	0.0	0.0	4.9	32.3	14.0	20.7	5.5	100.0	42.7	1.2	0.0	6.7	78.0	
14	187	100.0	91.3	100.0	23.6	81.4	100.0	0.0	16.8	40.2	100.0	34.4	1.5	0.0	1.0	0.9	0.0	1.3	89.4	1
15	677	100.0	93.6	100.0	29.9	84.0	100.0	0.0	19.3	56.1	0.0	66.3	16.6	0.0	2.1	2.1	0.0	5.9	92.0	1
16	458	100.0	74.9	0.0	10.9	65.7	100.0	0.0	12.9	32.8	16.4	20.7	0,7	0.0	0.0	0.4	0.0	2.0	82.3	0
17	24	100.0	66.7	58.3	12.5	37.5	0.0	0.0	0.0	37.5	20.8	16.7	20.8	0.0	0.0	0.0	0.0	100.0	91.7	33
10	173	100.0	74.0	55.5	7.5	59.5	0.0	0.0	10.4	29.5	100.0	29.5	12.1	0.0	0.0	0.0	0.0	0.0	82.1	1
19	582	100.0	79.6	100.0	11.5	60.8	0.0	0.0	2.9	14.6	0.0	19.4	1.0	0.0	0.5	0.2	0.0	0.0	73.5	1
20	857	100.0	59.0	0.0	2.7	45.6	0.0	0.0	2.2	10.4	0.0	8.0	0.0	0.0	0.7	0.4	0.0	0.0	66.7	0
21	618	0.0	14.7	12.6	1.9	19.4	13.9	5.5	4.7	22.2	11.5	9.7	7.1	100.0	41.1	2.6	0.0	8.7	55.3	9
55	215	0.0	26.5	40.9	7.0	41.4	59.1	100.0	32.1	47.4	35.8	41.9	19.5	0.0	1.4	1.4	0.0	7.0	76.3	- 4
	1164	0.0	29.0	26.1	2.1	43.3	100.0	0.0	13.0	19.0	13.5	18.1	1.9	0.0	0.8	0.7	0.0	1.2	66.6	0
24	246	0.0	2.4	2.4	0.0	2.0	0.0	0.0	0.4	7.7	3.3	0.8	2.0	0.0	0.0	0.0	0.0	27.2	62.2	100
25	295	0.0	35.6	100.0	2.4	32.9	0.0	0.0	3.1	8.5	18.0	23.7	1,4	0.0	0.3	0.0	0.0	1.4	100.0	0
	1923	0.0	13.5	0.0	0.3	16.8	0.0	0.0	1.8	4.2	9.1	7.3	1.2	0.0	0.7	0.5	0.0	1.9	100.0	0
27	371	0.0	17.0	13.7	0.3	100.0	0.0	0.0	2.4	6.2	5.4	2.4	0.3	0.0	0.3	0.3	0.0	0.3	0.0	0
28	204	0.0	10.3	6.9	0.0	0.0	0.0	0.0	0.1	7.8	100.0	11.6	8.3	0.0	0.0	0.5	0.0	0.5	0.0	0
29	494	0.0	38.7	26.3	0.6	0.0	0.0	0,0	2.2	10.9	0.0	18.0	1.6	0.0	6.5	5.7	0.0	8.5	0.0	0

TABLE 2.1: Screen Question (% YES to Screen Question)



		2.	

Line	"Parent"	Cluster	A10	A11	A12	A13	A14	A15	A16	A17	A18	A19	A20	A21	A22	A23	A24	٨25	A26	A27	A28	"Children"	Size
1	-		39.6	42.6	32.3	9.0	42.3	34.9	10.0	10.8	22.8	17.3	19.5	5.6	12.3	6.6	17.0	26 E	5.0	63.0	2.6		10.001
2	1		U	80.5	61.4	26.0	74.3	52.1	26.2	24.3	51.3	33.7	37.5	15.3	26.3	16.8	17.0	26.5 U	5.0	67.0	3.6	2, 9, 14, 37	12,907
3	2	1	U	92.7	79.9	59.7	89.8	85.5	U	62.7	86.8	60.1	63.7	42.2	38.6	27.1	73.3	U	9.2	85.5	2.5	3, 4	1,156
4	2	-	U	76.2	54.9	14.0	68.8	53.8	z	10.7	38.7	24.3	28.1	5.7	21.9	13.1	63.3	U	4.1	82.3	1.2	-	303
5	4	2	U	77.0	63.1	16.0	77.0	55.6	Z	11.2	46.5	31.0	35.3	11.0	U U	50.8	71.7	U	9.6	85.0	1.6	5, 6	853
6	4		U	76.0	52.6	13.4	66.5	53.3	Z	10.5	36.5	22.4	26.1	4.0	Z	2.6	61.0	U	2.6	81.5	1.1	7. 8	187
7	- 6	3	U	85.1	66.5	19.4	75.8	U	Z	15.8	49.6	26.5	34.6	5.9	Z	3.4	63.7	U	2.5	88.7	1.1	1+ 0	666 355
8	6	4	U	65.6	36.7	6.4	55.9	Z	Z	4.5	21.5	17.7	16.4	1.9	7	1.6	57.9	U	2.6	73.3	1.0	6n	355
										410	2103		10.4	1.7	5	1.0	31.9	0	6.0	13.3	1.0	-	311
9	1	-	Z	13.5	8.7	1.2	17.1	15.0	2.8	3.4	9.0	10.1	B.4	1.7	9.0	5.3	59.2	U	6,1	34.7	2.1	10, 11	2,262
10	9	5	7	18.7	18.2	3.4	25.6	24.6	4.9	6.9	21.7	20.7	17.7	8.4	U	46.3	59.6	Ű	12.8	55.7	7.9		203
11	9		Z	13.0	7.8	1.0	16.3	14.0	2.6	3.1	7.8	9.1	7.4	1.0	Z	1.3	59.1	U	5.4	32.6	-1.5	12. 13	2,059
12	11	6	Z	36.3	23.2	4.8	49.5	U	11.8	16.6	28.4	21.1	24.9	3.5	7	1.4	50.9	U	4.2	71.3	1.0		289
- 13	11	7	Z	9.2	5.3	0.3	10.8	Z	1.1	0.9	4.4	7.1	4.6	0.6	Z	1.3	60.5	U	5.6	26.3	1.6		1,770
14	1		U	79.2	58.1	19.4	68.2	52.3	17.1	18.3	37.0	21.6	30.3	8.6	11.6	5.7	0.8	2	4.2	81.4	1.9	15, 22	3,959
15	14		U	94.2	81.5	51.4	90.4	87.0	U	62.9	82.4	50.4	60.9	36.2	20.1	9.2	1.0	Z	14.0	92.9	4.7	16. 17	677
16	15	. 8	U	94.7	88.6	67.3	93.9	89.0	U	74.7	94.7	84.0	78.4	U	32.6	16.7	1.2	2	32.2	96.3	9.8		245
17	15	-	U	94.0	77.5	42.4	88.4	85.9	U	56.3	75.5	31.3	50.9	Z	13.0	4.9	0.9	Z	3.7	91.0	1.9	18, 19	432
18	17	9	U	92.9	82.1	55.4	89.3	91.1	U	58.9	87.5	30.4	50.0	Z	U	30.4	5.4	Z	10.7	100.0*	5.4		56
19	17	-	U	94.1	76.9	40.4	88.3	85.1	U	55.9	73.7	31.4	51.0	Z	2	1.1	0.3	Z	2.7	89.6	1.3	20, 21	376
20	19	10	U	95.7	81.0	55.7	91.9	93.8	U	U	85.2	33.3	55.2	7	7	0.5	0.0*	Z	2.4	89.0	1.9		210
21	19	11	U	92.2	71.7	21.1	83.7	74.1	U	Z	59.0	28.9	45.8	Z	Z	1.8	0.6	Z	3.0	90.4	0.6		166
22	14	-	U	76.1	53.3	12.8	63.6	45.2	2	9.1	27.6	15.7	23.9	3.0	9.9	5.0	0.8	Z	2.2	79.1	1.3	23, 24, 25, 30	3.282
23	22	15	U	91.9	71.3	25.0	81.3	U	2	16.9	58.1	31.9	39.4	7.5	U	45.6	4.4	Z	5.0	93.1	1.9	-	160
24	22	13	U	61.0	48.8	4.3	55.5	2	2	4.9	32.3	14.0	20.7	5.5	U	42.7	1.2	Z	6.7	78.0	4.3	-	164
25	22	*	U	85.9	65.4	20.1	76.3	U	2	15.8	39.9	19.8	34.2	3.3	Z	8,0	0.9	Z	2.1	87.3	1.0	26, 27	1,322
26	25	16	U	74.9	Z	10.9	65.7	U	2	12.9	32.8	16.4	20.7	0.7	Z	0.0*	0.4	Z	2.0	82.3	0.4		458
27	25	-	U	91.8	U	25.0	81.9	U	Z	17.4	43.6	21.6	41.3	4.7	2	1.3	1.2	Z	2.3	89.9	1,3	28, 29	864
28	27	14	U	91.3	U	23.6	81.4	U	Z	16.8	40.2	U	34.4	1.5	7	1.0	0.9	z	1.3	89.4	1.2		187
29	27	15	U	93.6	U	29.9	84.0	U	Z	19.3	56.1	Z	66.3	16.6	2	2.1	2.1	Z	5.9	92.0	1.6	-	677

Con't

Line	"Parent	Cluster	A10	A11	A12	A13	A14	A15	A16	A17	A18	A19	A20	A21	A22	A23	A24	A25	A26	A27	A28	"Children"	Size
30	22		U	68.0	42.3	6.5	52.4	7	7	3.3	14.3	10.9	14.5										
31	30	17	U	66.7	58.3	12.5	37.5	2	7	0.0*	37.5		14.5	2.0	L	0.6	0.2	1	1.5	71.1	1.3	31, 32	1,636
32	30		U	68.1	42.1	6.4	52.6	2				20.8	16.7	20.8	1	0.0*	0.0*	Z	U	91.7	33.3	-	24
33	30	18	U	74.0	55.5			2	2	3.3	14.0	10.7	14.5	1.7	2	0.6	0.2	7	Z	70.8	0.8	33, 34	1,612
34						7.5	59.5	2	1	10.4	29.5	U	29.5	12.1	Z	0.0*	0.0*	Z	Z	82.1	1.2	-	173
	32	10	U	67.3	40.4	6.3	51.8	2	2	2.5	12.1	Z	12.6	0.4	2	0.6	0.3	Z	2	69.5	0.8	35, 36	1,439
35	34	19	U	79.6	U	11.5	60.8	2	2	2.9	14.6	Z	19.4	1.0	2	0.5	0.2	Z	2	73.5	1.0	-	582
36	34	20	U	59.0	L	2.7	45.6	Z	Z	2.2	10.4	Z	8.0	0.0*	2	0.7	0.4	7	Z	66.7	0.6	-	857
37	1	-	z	20.5	17.5	1.2	27.3	24.9	4.5	5.7	12.2	13.8	12.5	3.0	11.2	5.7	1.2	7	4.2	66.0	5.8	38. 39	6 130
38	37	21	7	14.7	12.6	1.9	19.4	13.9	5.5	4.7	22.2	11.5	9.7	7.1	U	41,1	2.6	7	8.7	55.3	9.2	30 + 37	5,530
39	37		2	21.1	18.1	1.1	28.3	26.3	4.4	5.8	11.0	14.1	12.9	2.5	7	1.7	1.0	2	3.7	67.4	5.3	40 41	618
40	39.	22	7	26.5	40.9	7.0	41.4	59.1	U	32.1	47.4	35.8	41.9	19.5	1	1.4	1.4	2	7.0	76.3		40, 41	4,912
41	39		7	20.9	17.0	0.9	27.7	24.8	7	4.6	9.3	13.1	11.6	1,7	2	1.2	1.0	2			- 4.7	-	215
42	41	23	2	29.0	25.1	2.1	43,3	1	7	13.0	19.0	13.5	18.1	1.9	1			4	3.5	67.0	5.3	42, 43	4,697
43	41		7	18.3	14.0	0.5	22.5	7	7		6.1	13.5			1	0.8	0.7	6	1.2	65.6	0.4		1,164
44	43	24	7	2.4	2.4	0.0*	2.0	4	7	1.9			9.5	1.6	1	1.3	1,1	4	4.3	67.1	6.7	44, 45	3.533
45	43	-		19.5	14.9			6	6	0.4	7.7	3.3	0,.8	2.0	2	0.0*	0.0*	6	27.2	62.2	U		246
46	45		4			0.5	24.1	4	2	2.0	6.0	13.8	10.1	1.6	l	1.4	1.2	2	2.6	67.5	7	45. 49	3,287
47		-	1	16.5	13,3	0.5	18.9	1	4	2.0	4.7	10.3	9.5	1.2	2	0.6	0.5	2	1,8	U	2	47, 48	2,218
	45	25	4	35.6	U	2.4	32.9	l	L	3.1	8.5	18.0	23.7	1,4	Z	0.3	0.0*	2	1.4	U	2	-	295
48	46	26	1	13.5	Z	0.3	16.8	Z	2	1.8	4.2	9,1	7.3	1.2	Z	0.7	0.5	2	1.9	U	Z		1,923
49	45		z	25.7	18.2	0.4	34.7	z	Z	2.1	8.7	21.0	11.4	2.4	7	3.1	2.8	7	4.1	7	7	50, 51	1,069
50	49	27	z	17.0	13.7	0.3	U	7	2	2.4	6.2	5.4	2.4	0.3	7	0.3	0.3	7	0.3	7	7	504 51	371
51	49	-	2	30.4	20.6	0.4	Z	7	7	1.9	10.0	29.2	16.2	3.6	7	4,6	4.2	7	6.2	7	7	52. 53	698
52	51	28	7	10.3	6.9	0.0*	7	7	7	0.1	7.8	U	11.8	8.3	7	0.0*	0.5	7	0.5	7	7	JE 9 JJ	
53	51	29	Z	38.7	26.3	0.6	2	7	2	2.2	10.9	Z	18.0	1.6	7	6.5	5.7	2	8.5	2	2		204

Table 3.1

Cluster	Hearing	Vision	<u>Mobility</u>	Agility	Total
1	1.733	0.657	3.624	5.841	11.855
2	1.717	1.508	3.171	2.170	8.566
3	1.637	0.034	3.274	2.543	7.488
4	1.579	0.016	2.582	0.710	4.887
5	1.596	1.463	0.625	1.211	4.895
6	1.509	0.014	1.091	2.152	4.766
7	1.605	0.013	0.253	0.246	2.117
8	0.012	0.493	3.772	7.203	11.480
9	0.054	1.304	3.643	4.480	9.841
10	0.000	0.005	3.686	5.256	8.947
11	0.006	0.018	3.476	3.319	6.819
12	0.044	1.456	3.445	2.838	7.783
13	0.012	1.427	2.653	0.884	4.976
14	0.009	0.010	3.727	3.178	6.924
15	0.021	0.021	3.776	2.941	6.759
16	0.004	0.000	2.406	1.964	4.374
17	0.000	0.000	2.625	2.083	4.708
18	0.000	0.000	2.890	1.890	4.780
19	0.002	0.005	3.404	0.494	3.905
20	0.004	0.007	2.046	0.233	2.290
21	0.026	1.411	0.467	0.852	2.756
22	0.014	0.014	1.088	3.498	4.614
23	0.007	0.008	0.984	1.688	2.687
24	0.000	0.000	0.068	0.352	0.482
25	0.000	0.003	1.685	0.587	2.273
26	0.005	0.007	0.303	0.258	0.573
27	0.003	0.003	0.310	1.170	1.486
28	0.005	0.000	0.172	1.285	1.462
29	0.057	0.065	0.650	0.418	1.190





	Cluster	No. of OBS	E(NADL)	ID
		HV (Hearing/Visio	n)	
	2	187	8.566	HVMAI
	5	203	4.895	HVN1
	Total	390		
2.		<u>H</u> (Hearing)		
	1	303	11.855	HMA1
	3	355	7.488	НМА2
	4	311	4.829	HMI
	6	289	4.760	HAI
	7	1,770	2.120	HNI
1.1	Total	3,028		
3.		V (Vision)		
	9	56	. 9.841	VMAI
	12	160	7.783	VMA2
	13	164	4.976	VM1
	21	618	2.756	VNI
	Total	998		
		S (Special)		
	17	24	4.708	SMAI
	24	246	0.482	SNI
	Total	270		
		MA (Mobility/Agili	tv)	
	8	245	11.480	MAI
	10	210	8.947	MA1 MA2
	11	166	6.819	MA4
	14	187	6.924	MA4 MA3
	15	677	6.759	MA5
	Total	1,485	01777	

	Table 3.2	
Ordering o	f Clusters by "Umbrella" Group	DS

.

1

	1.00			
	Cluster	No. of OBS	E(NADL)	ID
1.1				
6.		M (Mobility		
	16	458	4.374	M2
	18	173	4.780	M1
	19	582	3.905	M3
	20	857	2.290	M4
	Total	2,070		
7.		A (A -: 1:++)		
/ •		<u>A</u> (Agility)		
	22	215	4.614	AI
	Total	215		
8.		N (Neither)		
	23	1,164	2,687	N1
	25	295	2.273	N2
	26	1,923	0.573	N6
	27	371	1.486	N3
	28	204	1.462	N4
	29	494	1.190	N5
	Total	4,451		

Table 3.2 (cont'd)Ordering of Clusters by "Umbrella" Groups

-

TABLE 4.1

Alo 0.21 27 0.24 Al1 0.16 16 0.08 Al2 0.16 12 0.03 Al3 0.08 0.04 02 Al4 0.13 10 0.03 Al5 0.15 02 07 Al6 0.10 0.15 10 Al7 0.07 0.07 06	0.25 0.07 0.02 05 01
A17 0.07 0.07 06 A18 0.12 0.07 04 A19 0.06 0.09 06 A20 0.09 0.02 04 A21 0.05 0.10 07 A22 0.04 0.26 13 A23 0.02 0.12 04 A24 01 0.24 0.36 A25 02 0.39 59 A26 0.01 0.06 02 A27 0.08 13 11 A28 00 0.02 03 B101 0.44 22 0.26 B111 0.24 0.07 05 B121 0.41 09 00 B131 0.06 0.10 06 B141 0.40 04 0.09 B151 0.27 0.13 11 B161 0.10 0.20 13 B171 0.08 0.13 09 B181 0.32 0.26 15 B191 0.05 0.08 06 B201 0.13 0.09 07 B211 0.09 0.18 12 B231 0.03 0.18 09 B241 00 0.09 0.13 B251 0.01 0.05 02	29 17 11 04 08 07 06 0.43 0.18 05 09 00 0.03 0.01 0.25 05 0.02 07 03 26 15 11 05 09 09 09 09 09 09 09 09 09 09 09 09 09 09 09 09 09 00 03 26 15 09 09 09 00 03 09 09 00 03 09 09 09 00 03 09 00 03 00 03 00 03 00

PRINCIPAL COMPONENT ANALYSIS



*

	TABLE 4.2	
PRINCIPAL COMPONENT	UMBRELLA GROUP	MEAN
	HEARING/VISION	
PRIN1	346	0.68
PRIN2	346	1.26
PRIN3	346	0.61
PRIN4	346	1.06
	HEARING	
PRIN1	2741	-0.33
PRIN2	2741	0.54
PRIN3	2741	0.81
PRIN4	2741	-0.25
	VISION	
PRINI	888	0.30
PRIN2	888	0.69
PRIN3	888	-0.76
PRIN4	888	1.27
	SPECIAL	
PRIN1	151	-1.02
PRIN2	151	-0.04
PRIN3	151	-0.47
PRIN4	151	-0.06
	MOBILITY/AGILITY	
PRINL	1311	3.31
PRIN2	1311	-0.33
PRIN3	1311	-0.21
PRIN4	1311	-0.33
	MOBILITY	
PRINI	1893	0.30
PRIN2	1893	-0.80
PRIN3	1893	0.18
PRIN4	1893	0.33
	AGILITY	
PRINI	195	-0.19
PRIN2	195	0.31
PRIN3	195	-0.80
PRIN4	195	-0.78
	NEIPIHER	
PRIN1	3887	-1.11
PRIN2	3887	-0.16
PRIN3	3887	-0.41
PRIN4	3887	-0.22

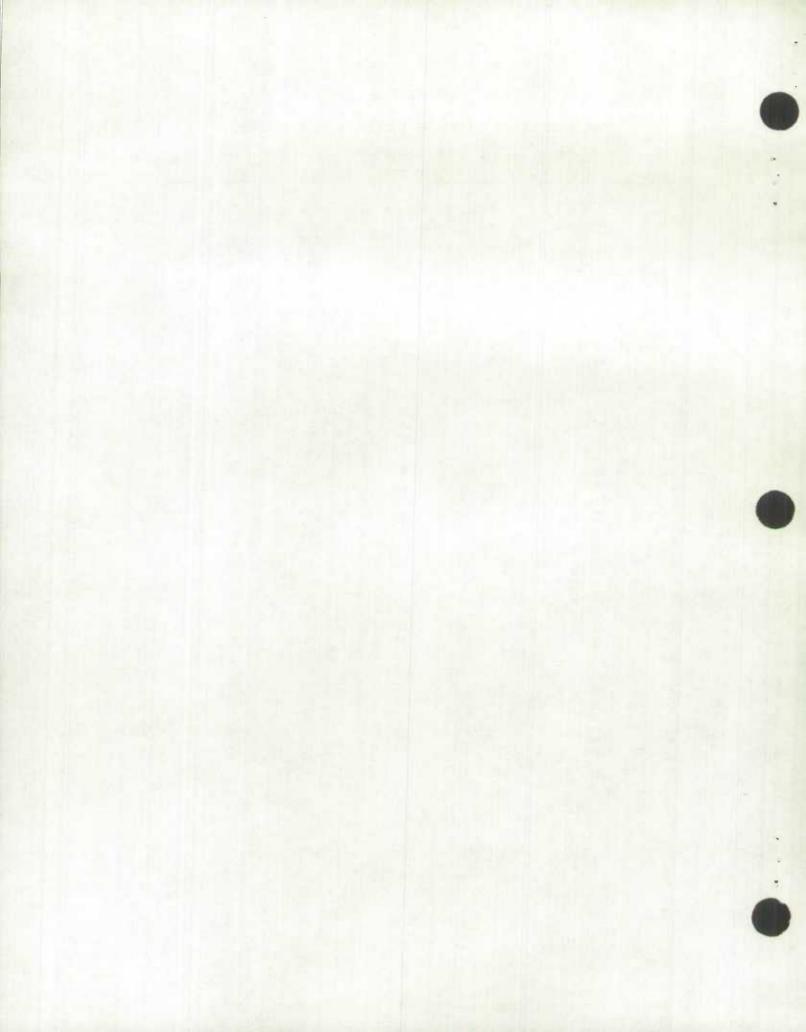
TABLE 4.2



TABLE 4.3

CLUSTER RANK ACCORDING TO ALTERNATIVE SCALES

Cluster	ID	PRIN1 (overall)	PRIN2 (M/AHV)	PRIN3 (MH/AV)	PRIN4 (MV/A)	E(NADL)
2	HVMA1	9	4	27	28	5
5	HVN1	22	2	22	25	12
1	HMAl	3	3	24	6	1
3	HMA2	10	14	28	10	7
4	HMl	16	15	29	20	13
6	HAl	20	8	25	3	15
7	HNl	29	7	26	9	24
9	VMA1	2	6	4	23	3
12	VMA2	4	10	7	27	6
13	VM1	13	11	11	29	11
21	VN1	23	5	2	26	20
8	MA1	1	1	1	1	2
10	MA2	5	20	13	4	4
14	MA3	6	24	16	7	8
11	MA4	7	23	17	8	9
15	MA5	8	28	20	18	10
18	M1	14	26	19	21	14
16	M2	15	25	18	17	18
19	M3	11	29	23	24	19
20	M4	18	27	21	22	22
22	Al	17	9	3	2	17
23	N1	21	17	6	5	21
25	N2	19	22	10	16	23
27	N3	24	19	15	12	25
28	N4	28	12	9	11	26
29	N5	25	16	12	15	27
26	N6	26	18	8	14	28
17	SMAL	12	21	14	19	16
24	SN1	27	13	5	13	29



APPENDIX B

A TECHNICAL NOTE CONCERNING THE CLUSTER ANALYSIS

I General

We present the methodology used to derive the final clusters. The analysis was performed on the 12,907 cases which were screened in, based on at least one positive response (Yes, has trouble) to A10 - A28.

The final clustering method used to develop the algorithm was completed in two phases. The first phase was primarily graphical and greater insight regarding the nature of the data was gained. A more objective clustering methodology was then used in the second phase. At the outset of the second phase, the analysis involved both unweighted and weighted data. The results were essentially the same. It was decided to continue without the sampling weights because of the added complexity which would be incurred by their inclusion.

2 First Phase Analyses

The first phase used mostly graphical techniques and PROC FASTCLUS in SAS. The procedure was as follows:

- a) Define an initial grouping based on primarily graphical methods.
- b) Use the results of this grouping and apply PROC FASTCLUS to it a number of times, revising the groupings and variables used on successive runs.
- c) Plot the final groupings to look for missed clusters and clusters to collapse.
- d) Iterate on steps (b) and (c).

For step (a), we started by transforming the variables A10 to A28 along the principal components based on the total covariance matrix (TCOV). By investigating plots of the first few principal components, lines were drawn on the plots to define plausible groups. For each group identified, the TCOV of the subset was recomputed and the procedure was repeated until no sufficiently large identifiable group was apparent. A total of 36 groups was thus identified.

In step (b), the operation involved was PROC FASTCLUS from SAS. Our implementation of this procedure computed the group means of certain variables for each of the groups being run. The procedure then assigned each individual to the group with the closest mean (sums of squared differences as a distance measure). After

passing through all the data, the group means were updated and the procedure repeated until convergence was achieved (i.e. no change in grouping from one iteration to the next). This procedure should work well under certain distributional assumptions (such as independence and constant variances of the input variables within each group). In order to mimic some of these assumptions, the input variables were the standardized principal components based on the within covariance matrix (WCOV resulting from PROC CANDISC).

After this was applied to the initial 36 groups, it turned out that one variable was constant within the resulting groups (A28 - mental handicap). In this case PROC CANDISC was repeated on each subset separately (A28=0 vs A28=1) using the final grouping from the previous run as the initial grouping for the new run. This procedure was repeated until no more subsets could be found with constant values within each group but non-constant values between groups.

Now 36 groups were defined. All but two of these groups could be defined based on only zero-one values of A10 to A28. In step (c), a series of plots using standardized principal components of the within covariance matrix, for two groups at a time, were constructed. Each two groups were chosen so as to be near each other, in the sense that their definition changed on only one ADL variable. Based on the plots, some groups were collapsed, resulting in looking at further plots after collapsing. In some cases, further splits were identified and subjected to PROC CANDISC as in step (b).

The iterative procedure of successive collapsing and splitting of groups was never completed because it was becoming too subjective and was also time consuming. Instead, we changed our approach to the clustering algorithm used in the second phase.

3 Second Phase Analysis

We learned two important facts in the first phase. First, we found that it is reasonable to define our clusters based on whether an ADL was zero or one. Secondly, we found that the correlations among the ADL's cannot be ignored in the clustering.

For the second phase, there were two main steps.

a) a divisive step, where we sequentially partitioned the 12907 individuals.

b) an agglomerative step, where we collapsed the partition.

In stage (a), we used PROC CANDISC of SAS. The 12,907 screened in respondents were partitioned into two subsets corresponding to one of the 19 screening items. This was done for each of the screening items and the partition which gave the most significant split on the remaining items was selected. The two groups resulting from the split were then studied separately for further splitting.

The most significant split was defined by the smallest value of Wilks' lambda, or equivalently, the largest F value.

After each split, we plotted the first seven standardized principal components based on WCOV for that split, to ascertain if we should consider splitting further. In cases where the plots did not show any potential for further splitting, we stopped unless the groups were still so large (greater than 500, say) that further splitting was justified because of too many hidden observations on the bivariate plots. In cases where further splitting was performed, we repeated the process of checking all possible partitions within the previously defined splits using PROC CANDISC and selecting the most significant.

In step (b) we collapsed as much as possible the groups in the reverse order of their creation. The criteria for collapsing were subjective, based on the following general considerations:

- Were groups sufficiently large to justify being separated?
- Was the F value reasonably large?
- Did the plots show a clear separation?
- Did the plots, before splitting the group, look very different than either of the plots after splitting the group?

In marginal cases, groups were collapsed in order to achive the smallest number of final groups and to avoid creating groups with very few individuals where possible. The final partition consisted of 29 clusters, two of which contained fewer than 100 respondents.

REFERENCES

- 4 -

Anderberg, Michael, R., (1973). CLUSTER ANALYSIS FOR APPLICATIONS. Academic Press, New York.

OBSTACLES, Report of the Special Committee on the Disabled and Handicapped, House of Commons, Ottawa, Canada, 1981.

Tatsuoka, Maurice M., MULTIVARIATE ANALYSIS: TECHNIQUES FOR EDUCATIONAL AND PSYCHOLOGICAL RESEARCH, New York: John Wiley and Sons Inc., 1971.

INTERNATIONAL CLASSIFICATION OF IMPAIRMENTS, DISABILITIES, AND HANDICAPS, World Health Organization, Geneva, Switzeland, 1980.







200 27







