A MULTIVARIATE MODEL FOR RESPONSE RELIABILITY IN SURVEYS

Miles Davis and John McCoy
Social Security Administration

Introduction

Nonsampling errors associated with response consistency can present serious problems in the analysis and interpretation of sample survey data. When surveys represent special populations with multiple problem characteristics, such as the poor, ethnic minorities, the aged, and those with health disorders and disabilities, such errors may have serious consequences for the results obtained. Therefore, an understanding of such sources of error is highly desirable.

Problems of survey reliability are complex and therefore call for appropriate multivariate analytic procedures sensitive to the interactive configurations of the data. In the investigation reported here, the research problem is approached with joint attention given to respondents and questions. The guiding theme concerns identifying patterns of consistency and inconsistency that are dependent on both questions and respondents. The technique selected for this task involves the spectral decomposition of a contingency table (Good). Its application, similar in several respects to the methods of principal components or to the mathematical first steps of factor analysis, is further discussed below.

Origins and characteristics of the data

Data were collected in the interview and reinterview phases of the Survey of Low-Income Aged and Disabled (SLIAD). The larger survey was designed as a before-after investigation of noninstitutionalized persons interviewed first in 1973 and recontacted for follow-up interview in 1974. Four national probability samples were represented: (1) low-income persons aged 65 and older, (2) disabled persons aged 18 and older, both screened from the Current Population Survey, (3) Old Age Assistance recipients, and (4) recipients of Aid to the Blind and Aid to the Permanently and Totally Disabled.

Reinterviews were conducted immediately after the 1974 follow-up survey. A total of 1,432 cases were selected from each of the four samples and further stratified as (1) rural nonproxy, (2) non-rural nonproxy, and (3) proxy. Stratification by proxy was done to see if responses obtained from persons other than the designated sample person would be less reliable. Differences between rural and urban reliability patterns were also of interest. However, only responses obtained from 434 rural nonproxy respondents are analyzed in this paper.

The reinterview differed from other census reinterview investigations in two major respects. First, response reconciliation, a procedure that provides reinterviewers with knowledge about a respondent's prior responses, was not practiced. Second, rather than determine what changes, if any, might have occurred in household composition since the prior interview, a detailed questioning procedure was followed. This was intended to maintain the independence of the two survey procedures and to reduce possible effects introduced by interviewers.

The concept of reliability applied in the analysis means simply that a response pattern is deemed reliable if it is repeated. Reliable response includes literally everything that happened to a data element from its verbal elicitation by interviewers to its representation as a magnetic mark on a tape. Response consistency is represented as a dichotomous variable. A response was consistent only if its designated codes were duplicated in the reinterview. Responses that were not on the main diagonal of a square table were defined to be inconsistent. All consistent responses were coded 1 and inconsistent responses were coded 0. This binary notation allows compact storage in the computer. Degrees of reliability arising from varying distances from the main diagonal in nondichotomous square tables are not considered.

A partial display of data appears in Table 1. Rows represent respondents and columns represent questions. For visual clarity, 1 is printed * (star) and 0 is printed ▯ (box). During statistical analysis * is scored as +1 and ▯ is scored as -1. Row and columns sums of these scores are shown in Table 1, bordering the data matrix. The two-way analysis of variance of the data is:

	Sum of squares	df	Mean square	F
Questions	2,405.505	60	40.092	63.077
Respondents	813.633	433	1.879	2.956
Interaction	16,512.790	25,980	.636	
Total, corr.	19,731.928	26,473	.745	
Mean	6,742.072	1	6,742.072	
Total	26,474.	26,474	1.	

Questions are a very important source of variation, much more so than respondents, although both have significant F statistics. The magnitude of the interaction mean square, .636, indicates complex interdependence between questions and respondents. The nature of this interaction is examined in subsequent analysis.

Singular decomposition of a rectangular matrix

In his monograph entitled The Estimation of Probabilities, I.J. Good (1965, pp.61-63) describes succinctly the singular decomposition of a contingency table. As noted earlier, the procedure produces results similar to those of principal components in multivariate analysis or to the mathematical first steps in factor analysis, but it is also applicable to non-square matrices. Good states that his ideas were drawn from Smithies (1958), who describes the singular decomposition of the kernels of integral equations, and from Halmos (1958, p.156), who gives the spectral decomposition of a self-adjoint linear transformation of a finite-dimensional vector space. Whittle (1952) discusses principal components and factor analysis.

Principal components are often computed from correlation or covariance matrices, which are square, symmetric and non-negative definite. Good's singular decomposition can be applied to non-symmetric or non-square matrices. Let A be a matrix of real elements, with s rows and t columns. If there exist vectors x and y of unit length, with s and t elements, respectively, and a number k such that

$$Ay = kx \quad \text{and} \quad A'x = ky,$$

then k is a <u>singular value</u> or <u>eigenvalue</u> of A and x and y are <u>singular vectors</u> or <u>eigenvectors</u>. Observe that

$$AA'x = k^2x \quad \text{and} \quad A'Ay = k^2y,$$

so that k^2 is an eigenvalue of the Gram matrices AA' and A'A; and x is an eigenvector of AA' and y is an eigenvector of A'A. Call x a <u>left eigenvector</u> and y a <u>right eigenvector</u>. Both matrices AA' and A'A are square, symmetric and non-negative definite, with the same non-negative eigenvalues, min(s,t) in number, except that the larger of the two has $|s-t|$ extra zero eigenvalues.

The singular decomposition is performed by finding the eigenvalues and eigenvectors of the smaller of A'A and AA'. The method of Jacobi is used as described in Ralston and Wilf (1960). Details are given below. Eigenvalues k_r^2 are computed and sorted into descending order and indexed by r in that order. Since all of them are non-negative, their positive square roots are defined to be the eigenvalues k_r of A. The eigenvectors of A'A are x_r and those of AA' are y_r. Once one set of eigenvectors is found, the other is computed by:

$$x_r = Ay_r/k_r \quad \text{or} \quad y_r = A'x_r/k_r$$

<u>Jacobi's method</u>

Jacobi's method operates iteratively on a square, symmetric, non-negative definite matrix S. In the n^{th} iteration, the matrices $E^{(n)}$ and $V^{(n)}$, both of the same size as S, are operated on to form $E^{(n+1)}$ and $V^{(n+1)}$. Initially, $E^{(o)} = S$ and $V^{(o)} = I$, the identity. The operation is:

1) The off-diagonal element of $E^{(n)}$ that is largest in absolute value is found. Call it $E^{(n)}_{ij} = E^{(n)}_{ji}$.
2) An orthogonal matrix $T^{(n)}$ is constructed, such that the element $E^{(n+1)}_{ij} = 0$, where $T^{(n)\prime}E^{(n)}T^{(n)} = E^{(n+1)}$. (' means transpose.)
3) $V^{(n+1)} = V^{(n)}T^{(n)}$.
4) When the absolutely largest off-diagonal element is small enough, stop.

Then $E^{(n)}$ is a matrix with the eigenvalues of S on the diagonal and zeroes elsewhere, to a sufficiently good approximation. $V^{(n)}$ is an orthogonal matrix whose columns are the eigenvectors of S, again approximately.

$T^{(n)}$ is computed thus: In the identity matrix I, replace the diagonal elements (i,i) and j,j) with c_n and the off-diagonal elements (i, j) and (j,i) by s_n and $-s_n$, respectively. c_n and s_n are the cosine and sine of an angle of rotation. They are computed thus: $c_n = 1/\sqrt{1+t_n^2}$ and $s = t_nc_n$,

where $t_n = r_n - (\text{signum } r_n)\sqrt{1+r_n^2}$ if $r_n \neq 0$ and $t_n=1$ if $r_n=0$, and $r_n = (E^{(n)}_{ii} - E^{(n)}_{jj})/(2E^{(n)}_{ij})$, for $E^{(n)}_{ij} \neq 0$.

If $E^{(n)}_{ij} = 0$, the process should have stopped, since $E^{(n)}_{ij}$ is the absolutely largest off-diagonal element.

<u>An example with small contrived data to help us understand the huge real data</u>

Suppose we had data from a questionnaire with s = 5 respondents and t = 3 questions. Reading * as +1 and □ as -1, and calling the resulting numerical matrix A, we form the Gram matrices

*	*	*	3
*	□	□	-1
*	*	□	1
*	□	*	1
□	*	□	-1
3	1	-1	3

$$A'A = \begin{matrix} 5 & -1 & 1 \\ -1 & 5 & -1 \\ 1 & -1 & 5 \end{matrix} \qquad \text{and } AA' = \begin{matrix} 3 & -1 & 1 & 1 & -1 \\ -1 & 3 & 1 & 1 & -1 \\ 1 & 1 & 3 & -1 & 1 \\ 1 & 1 & -1 & 3 & -3 \\ -1 & -1 & 1 & -3 & 3 \end{matrix}$$

We apply the Jacobi method to A'A, the smaller matrix:

$$E^{(o)} = A'A = \begin{matrix} 5 & -1 & 1 \\ -1 & 5 & -1 \\ 1 & -1 & 5 \end{matrix} \qquad V^{(o)} = \begin{matrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix}$$

Reduce the element $E^{(o)}_{12} = E^{(o)}_{21} = -1$ to 0:

$r_1 = (E^{(o)}_{11} - E^{(o)}_{22})/(2E^{(o)}_{12}) = (5-5)/(2(-1)) = 0$; $t_1 = 1$;

$c_1 = 1/\sqrt{2}$ and $s_1 = 1/\sqrt{2}$: $E^{(1)} = T^{(1)\prime}E^{(0)}T^{(1)} =$

$$\begin{bmatrix} .70711 & -.70711 & 0 \\ .70711 & .70711 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 5 & -1 & 1 \\ -1 & 5 & -1 \\ 1 & -1 & 5 \end{bmatrix}$$

$$\begin{bmatrix} .70711 & .70711 & 0 \\ -.70711 & .70711 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 0 & 1.41421 \\ 0 & 4 & 0 \\ 1.41421 & 0 & 5 \end{bmatrix}$$

$V^{(1)} = V^{(0)}T^{(1)} =$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} .70711 & .70711 & 0 \\ -.70711 & .70711 & 0 \\ 0 & 0 & 1 \end{bmatrix} =$$

$$\begin{bmatrix} .70711 & .70711 & 0 \\ -.70711 & .70711 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Reduce the element $E_{13}^{(1)} = E_{31}^{(1)} = 1.41421$ to 0:

$r_2 = (E_{11}^{(1)} - E_{33}^{(1)})/(2E_{13}^{(1)}) = (6-5)/(2(1.41421)) =$ $.35355$; $t_2 = -.70711$; $c_2 = .81650$; $s_2 = -.57735$

$E^{(2)} = T^{(2)\prime}\, E^{(1)} T^{(2)} =$

$$\begin{bmatrix} .81650 & 0 & .57735 \\ 0 & 1 & 0 \\ -.57735 & 0 & .81650 \end{bmatrix} \begin{bmatrix} 6 & 0 & 1.41421 \\ 0 & 4 & 0 \\ 1.41421 & 0 & 5 \end{bmatrix} \cdot$$

$$\cdot \begin{bmatrix} .81650 & 0 & -.57735 \\ 0 & 1 & 0 \\ .57735 & 0 & .81650 \end{bmatrix} = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

$V^{(2)} = V^{(1)} T^{(2)} =$

$$\begin{bmatrix} .70711 & .70711 & 0 \\ -.70711 & .70711 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} .81650 & 0 & -.57735 \\ 0 & 1 & 0 \\ .57735 & 0 & .81650 \end{bmatrix} =$$

$$\begin{bmatrix} .57735 & .70711 & -.40825 \\ -.57735 & .70711 & .40825 \\ .57735 & 0 & .81650 \end{bmatrix}$$

The eigenvalues of A'A are 7, 4 and 4, so the eigenvalues of A are $\sqrt{7} = 2.64575$, $\sqrt{4} = 2$ and $\sqrt{4} = 2$.

The question eigenvectors of A are the columns of $V^{(2)}$. The respondent eigenvectors of A are the columns of $AV^{(2)}\,(E^{(2)})^{-1/2}$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & -1 \\ 1 & 1 & -1 \\ -1 & 1 & -1 \\ -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} .57735 & .70711 & -.40825 \\ -.57735 & .70711 & .40825 \\ .57735 & 0 & .81650 \end{bmatrix} \cdot$$

$$\cdot \begin{bmatrix} .37796 & 0 & 0 \\ 0 & .5 & 0 \\ 0 & 0 & .5 \end{bmatrix} = \begin{bmatrix} .21822 & .70711 & .40825 \\ .21822 & 0 & -.81650 \\ -.21822 & .70711 & -.40825 \\ .65465 & 0 & 0 \\ -.65465 & 0 & 0 \end{bmatrix}$$

Finally, we sort the rows of A into descending order of the respondent eigenvector values corresponding to the largest eigenvalue, $\sqrt{7}$. Similarly, we sort the columns on the largest eigenvector. The data are displayed as in Table 1:

	1	3	2	sum	1	2	3
4	*	*	□	1	.65465	0	0
1	*	*	*	3	.21822	.70711	.40825
2	*	□	□	-1	.21822	0	-.81650
3	*	□	*	1	-.21822	.70711	-.40825
5	□	□	*	-1	-.65465	0	0
sum	3	-1	1	3			

Analyzing the real data

The data matrix (shown partially in Table 1), with 434 respondents (rows) and 61 questions (columns) was analyzed, using the methods described above, by means of APL functions written by the first author, on computers at the Parklawn Computation Center, Rockville, Maryland, through a remote terminal at the Social Security Administration. The Jacobi method required just over 5100 iterations and supplied the eigenvalues k_r of the data matrix. The eigenvalues are listed in Table 2. The matrix of question eigenvectors associated with the first two eigenvalues is shown in Table 3. The respondent eigenvectors corresponding to the two largest eigenvalues are listed in Table 1 to the right of the data matrix. The eigenvectors corresponding to the first eigenvalue are used to rank respondents and questions in descending order of the elements of the eigenvector. The first eigenvectors may be thought of as measures of reliability for questions and respondents.

Questions are plotted in Figure 1, using their first and second eigenvector elements as horizontal and vertical axes. Each point is labelled with its question number. A list of question numbers and a very brief description of the content of each question is in Table 4. Figure 1 shows interesting clusters of similar questions. The most reliable questions, 33, 35, 37 and 41, form a tight cluster about (.200,-.035). They are race, sex, confinement to wheelchair or bed, and inability to speak English, respectively, all clear and relatively permanent characteristics. Other interviewer questions, in the range 34-47, form a slightly less reliable cluster to the left of the most reliable one. Questions 28 and 29, about receipt of SSI and Social Security benefits, also fall in this cluster. An obvious cluster contains questions 48-51 near the point (.050, .410) all of which concern stairs or steps. Questions 58-61 near (.050, .240) are about land usage, railroad tracks and abandoned buildings. Questions 52-57, forming a loose cluster near (.150, .065), involve description of the block in which the respondent lives, as do questions 58-61. The Haber Functional Limitation questions, 18-26, form a very loose cluster centered near (.080,-.040), mixed with questions involving distances, 4-13, and occupation, 14-17. The least reliable question is 31, at (-.068,-.052), about total annual income of the nuclear family.

The second eigenvector, unlike the first, has no obvious interpretation, but it does serve to separate questions into interesting clusters. It is also the orthonormal contrast accounting for the second largest part of the total sum of squares of data elements. The total sum of squares, 26,474, is equal to the sum of squares of the eigenvalues. The square of the first eigenvalue is 9638.134, or 36.41 percent of the total. The square of the second is 1724.254, or 6.50 percent of the total. The higher eigenvalues become smaller rather gradually, without a sharp gap. One can learn more by looking at successive eigenvectors, but with diminishing returns for the effort.

Figure 2 shows respondents plotted as questions are in Figure 1. More reliable respondents are plotted toward the right. The most reliable is 398 at (.068,.039). The cluster centered at (.055,.050) contains the bulk of reliable respondents. Respondent 66, at the bottom of the graph, was inconsistent on stairs and land use questions, giving a very negative second eigenvector.

Discussion

These graphs of the eigenvectors provide a rich source of information for understanding the multivariate configurations manifested by question/respondent consistency. Spectral decomposition is currently being applied to the remaining nonproxy nonrural and proxy matrices. Comparative analysis will provide further information about the relative impact of rural-urban location and respondent/proxy sources of data. Further investigation is needed both for interpreting respondent eigenvectors and particularly for distinguishing consistency patterns associated with respondent characteristics and interviewers.

Finally, response consistency need not be coded as a dichotomous or binary variable. Such a measure could range over a finite interval, say 0 to 1, or -1 to +1, at some cost in computer storage to be sure, but with no difficulty in theory or computation. The singular decomposition procedure could be applied to any real matrix with its interpretation dependent on the meaning of the data. Moreover, a similar balanced interpretation of rows and columns would be possible.

References

Good, I. J. 1965. The Estimation of Probabilities. Cambridge, Mass.: M. I. T. Press.

Halmos, P. R. 1958. Finite-Dimensional Vector Spaces. Princeton, N. J.: Van Nostrand.

Ralston, A. and H. S. Wilf. 1960. Mathematical Methods for Digital Computers. New York: Wiley.

Smithies, F. 1958. Integral Equations. Cambridge: Cambridge University Press.

Whittle, P. 1952. "On principal components and least squares methods of factor analysis," Skand. Aktuarietidskr. 35, 223-239.

Table 1: Data matrix ordered by first eigenvectors

```
                             Questions
Respon-                                                                  Respondent
dents                                                                    eigenvector
     3334343424324   155 5513534 511142  23 14211 2126225245456513   Row
     3571408384997795156137022666438626287045559234710029380911841   sum      1       2
398  ************************□***********************************□    57   .0682   .0391
 17  *****************************************□*****□*□**□*******□    51   .0666   .0541
332  *****************************************□□****□**□********□□    49   .0660   .0508
375  *********************************□***********□*□****□****□*□□    47   .0653   .0461
227  ********************************□************□*□****□***□***□    49   .0651   .0329
228  ******************□******************□********□************□□    51   .0648   .0497
385  *************************□***********□******□*************□□□    49   .0646   .0372
 12  ****************************□****□*************□****□******□□    49   .0645   .0547
 87  ***********************□*********□*************□*□*******□**□    49   .0634   .0448
 38  ************************************□****□**□□***□*********□*    49   .0627   .0459

195  *********□*******□□*□□**□***□**□*□□*□**□***□□□□*****□******□□    23   .0336   .0447
 36  *****************□□*□□**□□**□***□***□□**□***□***□*□□*□□□□□□□*    17   .0333  ¯.1020
 66  *******□*********□□*□□**□***□**********□□*□□*□**□□□□□□□□□□□□□    11   .0333  ¯.0971
184  ********□*******□*******□*□□***□□*□□□*□□□*□**□□**□***□□□□**□□    15   .0332  ¯.0229
 29  *□**************□**□**□****□□*□□□**□***□□□****□**□□**□□□□**□□    17   .0332  ¯.0200

277  ********□*****□**□□*□□**□***□□□****□*□**□□*□*□**□*□□*□□□□□□□*     9   .0235  ¯.0997
334  *****************□□*□□**□□**□□□□*□**□**□□□□□*□□*□□□□*****□□□*     9   .0232  ¯.0124
312  *****□*****□□*□*□**□□*□**□□□***□**□□□□□******□****□*□*******□    19   .0214   .0668
340  **************□□*□□□□□*□□*□*□****□□□**□□****□*□*□□*□*□□□□□□*□     5   .0186  ¯.0746
165  **□□*□□□*□□**□*************□**□□□**□***□*□**□*□***□*□□□□□□**□    11   .0183  ¯.0128
141  *******□****□□**□*□*□*□□□□□□□□*□*□□□**□□***□□□□□*□□*□******□□     1   .0084   .0705
144  *************□□□□□□□□□□*□**□□□****□□**□*□**□□□**□□□□*□□□□□□**    ¯1   .0078  ¯.0968
264  **********□□*□□**□□*□□**□*□□□□□□*□*□*□***□□□**□□□□□□*□□□□□□**    ¯5   .0048  ¯.0687
247  □□**********□□□□*□□*□□□□□□□□□□***□*□□*□□*□*□*□□*□□*□□*□□□□□□□□   ¯15  ¯.0063  ¯.0812
 81  **□□*□□□*□□□□**□*□□*□□**□*□*□*□*□□****□□□□□□□□*□*□□*□□□□□****   ¯11  ¯.0139  ¯.0285
                                                                -
     444443333333333333232222222222222111111111111111           ¯1
Col. 2211199877766434319077885754211019878745343211009978755544413
sum  2286842486080460640668482046688804626882824462082282662060244   13360
```

Table 2
Eigenvalues of the data

Eigenvalue number	Eigenvalue
1	98.174
2	41.488
3	38.750
4	30.608
5	28.234
6	26.811
7	24.112
8	23.464
9	22.386
10	21.734
11	21.451
12	21.019
13	20.549
14	20.177
15	19.805
16	19.573
17	19.389
18	18.513
19	18.290
20	17.875
21	17.364
22	16.999
23	16.894
24	16.614
25	16.066
26	15.703
27	15.321
28	15.106
29	14.527
30	14.176
31	13.933
32	13.154
33	12.799
34	12.774
35	11.976
36	11.895
37	11.468
38	11.374
39	11.024
40	10.455
41	10.171
42	9.976
43	9.659
44	9.408
45	9.233
46	8.911
47	8.729
48	8.230
49	7.977
50	7.887
51	7.713
52	7.450
53	7.140
54	6.314
55	5.759
56	5.443
57	4.814
58	4.546
59	3.958
60	3.715
61	3.439

Table 3
Question eigenvectors

Question number	Eigenvector 1	Eigenvector 2
1	.151	-.040
2	.097	-.018
3	.064	.003
4	.083	-.055
5	.171	-.017
6	.129	-.042
7	.175	-.029
8	.093	-.036
9	.172	.012
10	.145	-.028
11	.167	-.020
12	.069	.025
13	.113	-.005
14	-.002	-.045
15	.081	-.094
16	.106	-.058
17	.060	-.054
18	.112	-.041
19	.071	.004
20	.049	-.039
21	.054	-.102
22	.047	.000
23	.040	-.069
24	.060	.018
25	.073	-.024
26	.101	-.092
27	.091	-.029
28	.184	-.031
29	.179	-.031
30	.091	-.083
31	-.068	-.051
32	.143	-.029
33	.202	-.035
34	.199	-.044
35	.202	-.028
36	.136	-.006
37	.201	-.031
38	.190	-.033
39	.181	-.036
40	.191	-.046
41	.201	-.036
42	.105	-.002
43	.186	-.014
44	.183	-.006
45	.073	.212
46	.130	-.008
47	.177	-.017
48	.037	.410
49	.035	.405
50	.036	.406
51	.031	.360
52	.138	.078
53	.150	.059
54	.125	.061
55	.166	.073
56	.157	.040
57	.149	.067
58	.026	.247
59	.046	.236
60	.051	.242
61	.028	.220

Table 4
Content of questions

Question number	Content
1	Parents present in childhood
2	Head of family in childhood
3	Childhood head of family occupation
4	Distance to grocery store
5	Unit of distance to grocery store
6	Distance to drug store
7	Unit of distance to drug store
8	Distance to restaurant
9	Unit of distance to restaurant
10	Distance to hospital
11	Unit of distance to hospital
12	Distance to friend
13	Unit of distance to friend
14	Work history
15	Industrial code
16	Private or public employment
17	Occupational code
18*	Walking
19*	Using stairs
20*	Standing
21*	Sitting
22*	Stooping
23*	Lifting
24*	Carrying weights
25*	Reaching
26*	Using fingers
27	Home ownership, single or joint
28	SSI benefits this year
29	Social Security benefits this year
30	Welfare in past 12 months
31	Annual income of nuclear family
32	Age
33	Race
34	Ethnic descent
35	Sex
36	Education
37#	Confined to wheelchair or bed
38#	Blind or near blind
39#	Very hard of hearing
40#	Unable to speak clearly
41#	Unable to speak English
42#	Type of proxy response
43#	Number of floors in residence
44#	Floor of residence
45#	Street level approach
46#	Residence in city or farm
47#	Living quarters
48#	Stairs to reach residence
49#	Interior or exterior stairs
50#	Number of stairs
51#	Steps without handrail
52@	Pedestrian sidewalks
53@	Detached single family dwellings
54@	Mobile homes
55@	Attached or row houses
56@	Apartment buildings
57@	Abandoned automobiles
58@	Abandoned buildings
59@	Railroad tracks
60@	Industrial land usage
61@	Commercial land usage

* Haber Functional Limitation Scale
\# Interviewer observation
@ Interviewer block description

Figure 1

Eigenvector values for each question
Eigenvector 2 plotted against eigenvector 1
Questions are identified by number

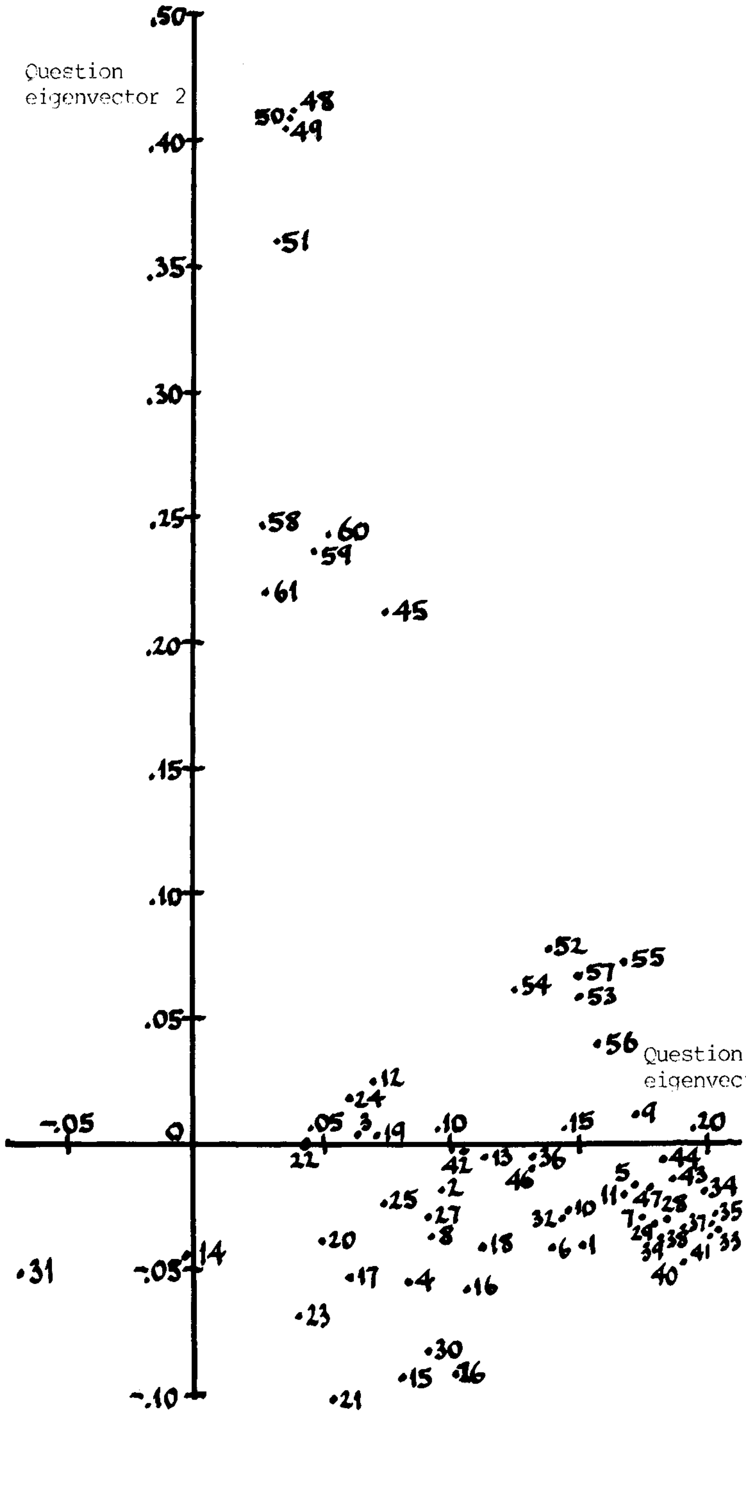

Figure 2

Eigenvector values for each respondent
Eigenvector 2 plotted against eigenvector 1
Some respondents are identified by number