FACTOR ANALYSIS OF SOLUTE PARAMETERS IN ETHER–STATIONARY PHASE INTERACTIONS

RUSSELL B. SELZER and DARRYL G. HOWERY*

Department of Chemistry, Brooklyn College of the City University of New York, Brooklyn, N.Y. 11210 (U.S.A.)

(Received May 29th, 1975)

SUMMARY

A data matrix of retention indices for 18 ethers on 25 stationary phases is factor analyzed. Six factors reproduce 96% of the data points to better than ± five retention index units. Over 30 vectors based on physically and chemically significant parameters of the ethers appear upon rotation to be factors. Among the successfully rotated parameters are molecular weight, total atom number, boiling point (°K), unity, unsaturation uniqueness, and the difference in the carbon number for the two alkyl groups. The relative importance of the physically significant vectors is estimated by reproducing the matrix using all possible combinations of vectors. In a six-factor space, the best set of physical vectors gave an average row average error of 5.3. The square of the boiling point (°K) is shown to be most nearly equivalent to the dominant eigenvector. Using combinations of six vectors consisting of rows and columns of data from the original matrix, the best reproductions for solute-associated and solvent-associated vectors had average row average errors of 3.3 and 2.9, respectively.

INTRODUCTION

Factor analysis of retention data[1–8] based on the method of Weiner and co-workers[9] has proved to be a highly promising approach for acquiring insights into the solute–solvent interactions in chromatography[10]. Parameters of the solvents and solutes can be tested individually for possible identification with the fundamental factors of the retention data space; knowledge of the functional form of the interactions is not required.

Factor analysis can contribute to an understanding of interactions in complex problems in two main ways. First, if one has a functional relation for an interaction, then factor analysis can be applied to test the proposed model. This approach was followed by Weiner *et al.*[9] and Weiner and Malinowski[11] in their highly original application of factor analysis to solvent effects on chemical shifts. Models for retention

---

* To whom correspondence should be addressed.

mechanisms in gas–liquid chromatography (GLC) were recently verified by Weiner et al.[7] using factor analysis. Unfortunately, the state of theories of interactions is such that we seldom have even a hint as to the mathematical form of the interaction terms. Thus, a second way of using factor analysis, the empirical approach, offers special promise for solving problems in many fields. In the empirical approach, even though we may be completely ignorant of the quantitative form of some or probably all of the energies of interaction, factor analysis allows one to find at least a partial solution to the problem by empirically testing in the rotational scheme of Malinowski various physical and chemical vectors associated with the row and column elements of the data matrix. Such a working-backwards approach can be utilized for problems that till now have not been amenable to analysis. The empirical approach has been applied by Howery, Weiner and co-workers to GLC[1–7,12] and to several other areas[8,13–15].

The vectors are chosen by a combination of chemical insight and empiricism; one acquires from an analysis of the results of the rotations estimates for the best functional dependence of the interaction terms. Then, by testing all possible combinations of the better vectors using the reproduction scheme of Malinowski, one can determine which combinations best reproduce the data matrix. If reproductions within or near the experimental error of the data points are found for both the row and the column elements, then one has found empirical solutions that are good approximations of the mathematically correct solution. At present, only factor analysis allows the possibility of such insight into multidimensional problems.

We present here a detailed factor analysis of the retention indices of ethers on a variety of stationary-phase solvents. In an effort to extend the scope of the empirical approach, new methods for estimating the relative importance and equivalence of physically and chemically significant parameters are explored.

PROCEDURE

From the compilation of McReynolds[16], a matrix of retention indices at 120 °C for 18 ethers on 25 stationary phases was chosen. The solutes and solvents are listed in Table I. Factor analyses were performed on IBM 360/145 and IBM 360/168 digital computers using a FORTRAN IV program developed by Weiner et al.[9] and expanded extensively by us. The program is available upon request.

RESULTS AND DISCUSSION

Using the approach described previously[6], six abstract eigenvectors seem to be required to reproduce the data matrix within experimental error. Five eigenvectors reproduced the matrix with an average row average error (a.r.a.e.) of 2.8 retention index (RI) units, the largest error being 14, and 34 points having errors exceeding five RI units. In a six-factor space the a.r.a.e. is 2.1. with the largest error being 8.2 and only 16 points having errors greater than 5. Since the average experimental error in McReynolds' data is in the range 3–5 RI units, we feel that the ether–solvent interaction space is adequately represented by six factors. Rotations of several physically significant vectors also gave considerably better fits for six factors than for five factors, further confirming our choice.

TABLE I

UNIQUENESS TESTS FOR SOLUTES AND STATIONARY PHASES

Number/Letter	Name	Uniqueness test value	Exner value
A. Solutes			
1	Dimethyl ether	0.84	0.50
2	Propyl methyl ether	0.14	1.17
3	Butyl methyl ether	0.17	1.15
4	tert.-Butyl methyl ether	0.59	0.81
5	Diethyl ether	0.23	1.11
6	Butyl ethyl ether	0.28	1.07
7	tert.-Butyl ethyl ether	0.19	1.13
8	Dipropyl ether	0.35	1.01
9	Isopropyl propyl ether	0.28	1.07
10	Diisopropyl ether	0.24	1.10
11	tert.-Butyl isopropyl ether	0.29	1.06
12	Dipentyl ether	0.42	0.96
13	Diisopentyl ether	0.41	0.97
14	Ethyl vinyl ether	0.37	1.00
15	Butyl vinyl ether	0.27	1.08
16	Isobutyl vinyl ether	0.28	1.07
17	2-Ethyl-1-hexyl vinyl ether	0.39	0.98
18	Allyl ethyl ether	0.25	1.10
B. Solvents			
A	Apiezon L	0.13	1.09
B	Carbowax 600	0.32	0.97
C	Carbowax 20M	0.34	0.95
D	Butyl tetrachlorophthalate	0.07	1.13
E	Di-(2-ethyl) sebacate	0.08	1.12
F	Diglycerol	0.90	0.38
G	Diisodecyl phthalate	0.19	1.05
H	Dioctyl sebacate	0.12	1.10
I	Dow Corning 550 fluid	0.12	1.10
J	Dow Corning FS 1265 fluid	0.53	0.80
K	Flexol 8N8	0.36	0.93
L	Hyprose SP80	0.60	0.74
M	Igepal CO 890	0.16	1.07
N	Isooctyldecyl adipiate	0.07	1.13
O	Pluronic acid	0.16	1.08
P	Polyphenyl ether-5 rings	0.24	1.02
Q	Quadrol	0.29	0.99
R	SE-30	0.14	1.09
S	SE-31	0.13	1.09
T	SE-52	0.14	1.09
U	Sucrose acetate isobutyrate	0.10	1.11
V	TMP tripelargonate	0.06	1.14
W	Tricresyl phosphate	0.07	1.13
X	Ucon LB-1715	0.07	1.13
Y	Zonyl E7	0.60	0.74

Tests for physically and chemically significant parameters

We assume that six solute–solvent interactions significantly influence the mea-
sured RIs. Our main problem is to find physically and chemically significant param-

eters of the ethers that, when used as test vectors in the rotational least-squares method of Weiner *et al.*[9], will align well in the six-factor space with one of the abstract eigenvectors. We are especially interested in testing a large number of chemical vectors based on structural differences in these closely related solutes.

Following the suggestion of Kindsvater *et al.*[8], we have used the value of the Exner function[17] to estimate the relative success of rotations. We have found this function to have mixed utility. In most cases, our qualitative evaluation of a rotation based on chemical insight agrees with calculated Exner values, *i.e.*, better rotations have lower Exner values. There are many cases, however, in which Exner values are quite contrary to our qualitative evaluations. Hence, we propose that the value of the Exner function should never be used as the sole criterion for evaluating rotations. We depend primarily on a qualitative point-by-point evaluation of the inputed and calculated vectors; the average error and Exner value are utilized as secondary information. For test vectors having similar characteristics, such as a series of uniqueness tests, the Exner value is most useful for comparing rotations.

The reciprocal, square and natural logarithm of each vector are tested routinely. Tests of the logarithms of vectors are usually of little value since the logarithms of most vectors cover too small a range and often approach a unity vector.

Physically significant factors

Tests for such physical properties of the ethers as molecular weight, molar refraction, boiling point (°K), and boiling point (°K) squared (SBK) gave quite good agreement between inputed known values and corresponding values predicted by the best-fit rotated vector. Tests for such vectors as boiling point (°C) and freezing point (°C) produced generally good fits though several points were predicted with greater than 10% error. A summary of all the better rotations involving physical vectors is presented in Table II, in which are given descriptions of the vectors and the results of the rotations. Properties that were predicted with large errors (and, therefore, do not seem to have the attributes of fundamental factors) include refractive index and density. Many physical properties, such as vapor pressure, viscosity, enthalpy of combustion, enthalpies of phase transitions, standard entropies and standard thermodynamic functions of formation, could not be validly tested for lack of data. (In a six-factor space, at least six points that span well the total vector space are required on a valid test vector.)

Chemically significant parameters

In an effort to identify structural parameters that might significantly influence the retention indices of ethers, a large number of structure-based vectors were formulated and tested. Among the relatively straightforward vectors that tested well enough to be possible fundamental factors are: unity (which should test well if the ethers contain a common factor), unsaturation uniqueness[5], carbon number, hydrogen number and total atom number. Other vectors that showed somewhat poorer agreement (correct qualitative pattern with a few poorly predicted points) but that still may be factors include: methyl ether uniqueness (for CH_3OR-type ethers), highest degree of substituent character for the carbon atoms in the ether bond (primary, secondary or tertiary character), and the difference in the number of carbon atoms in the two alkyl groups. A summary of the better rotations involving chemical vectors

TABLE II

BETTER ROTATIONS OF PHYSICALLY SIGNIFICANT VECTORS

Abbreviation	Vector name	Example values*	Points**	Qualitative evaluation***·§	Exner value§
MW	Molecular weight	102.2, 116.2	18	v	0.104
MR	Molar refraction	31.58, 36.22	18	v	0.086
CN	Carbon number	6, 7	13	v	0.069
TN	Total atom number	21, 24	13	v	0.088
CO	Carbon + oxygen number	7, 8	13	v	0.088
HN	Hydrogen number	14, 16	18	v	0.073
FC	Freezing point (°C)	−94.0, −75.51	12	f	0.958
FK	Freezing point (°K)	179.2, 192.5	12	g	0.597
BC	Boiling point (°C)	73.10, 86.37	16	g	0.104
BK	Boiling point (°K)	346.3, 353.5	16	g	0.251
UN	Unity	1, 1	18	v	§§
DU	Dimethyl ether uniqueness	0, 0	18	g	0.500
TU	tert.-Butyl methyl ether uniqueness	0, 0	18	g	0.810
MU	Methyl uniqueness	0, 0	11	g	0.575
UU	Unsaturation uniqueness	0, 0	16	g	0.259
VU	Vinyl uniqueness	0, 0	17	v	0.186
CD	Chain difference§§§	2, 1	10	f	0.336
CR	Chain ratio†	2, 1.33	9	g	0.549
HA	Highest adjacent††	3, 3	13	f	0.649
LA	Lowest adjacent††	1, 2	13	f	0.646
SA	Sum adjacent††	4, 5	13	g	0.592
RMW	Reciprocal MW	0.0098, 0.0086	18	g	0.300
RMR	Reciprocal MR	0.0317, 0.0276	18	g	0.269
RCN	Reciprocal CN	0.167, 0.143	18	g	0.206
RTN	Reciprocal TN	0.0476, 0.0417	18	g	0.193
RCO	Reciprocal CO	0.143, 0.125	18	g	0.193
RHN	Reciprocal HN	0.0714, 0.0625	18	g	0.264
RFC	Reciprocal FC	−0.0106, −0.0123	12	f	0.288
SCN	Square CN	36, 49	18	g	0.101
STN	Square TN	441, 576	18	v	0.088
SCO	Square CO	49, 64	18	v	0.088
SBK	Square BK	119890, 129240	15	v	0.026
LCN	Logarithm CN	1.79, 1.95	18	g	0.113

* Examples given are for tert.-butyl ethyl ether (solute 7) and tert.-butyl isopropyl ether (solute 11).
** Number of points on input vector.
*** v = Very good agreement between inputed and calculated vector; g = good agreement; f = fair (several points poorly predicted or pattern only predicted).
§ Data apply to rotations using six factors.
§§ Value undefined.
§§§ Total number of carbon atoms on largest radical − total number of carbon atoms on smallest radical.
† Total number of carbon atoms on largest radical/total number of carbon atoms on smallest radical.
†† Based on the degree of the two carbon atoms involved in the ether linkage, e.g., primary = 1, tertiary = 3.

is given in Table II. Typical vectors that do not appear to be factors include: ratio of the number of carbon atoms in the two chains, total number of carbon atoms off the main chains, ratio of the carbon number of the longer main chain to the carbon number of the shorter main chain, ratio of main-chain carbon number to branch-chain carbon number, isopropyl uniqueness, ethyl uniqueness, dialkyl uniqueness (for

ROR-type ethers), ratio of carbon number to hydrogen number, and number of carbon atoms between the nearest branch chain and the ether bond.

Details of the rotations for methyl uniqueness, boiling point (°K) squared and unsaturation uniqueness are presented in Table III. When the value for dimethyl ether is free-floated (left blank) in the methyl uniqueness test, a value satisfyingly close to 2.00 (1.72) is predicted for that solute.

TABLE III

DETAILS OF SELECTED ROTATIONS

MU -- Methyl uniqueness; SBK = square boiling point (°K); UU = unsaturation uniqueness.

Solute[a]	MU		SBK × 10^{-4}		UU	
	Test	Predicted	Test	Predicted	Test	Predicted
1	(2)	1.72	6.26	6.23	0	0.01
2	1	0.84	9.74	9.87	0	0.04
3	1	0.83	(11.80)	11.90	0	0.19
4	1,	1.00	10.80	10.70	0	−0.05
5	0	0.39	9.47	9.55	0	0.05
6	0	0.04	13.60	13.50	(0)	0.04
7	0	0.08	12.00	12.00	0	0.12
8	(0)	−0.28	(13.30)	13.50	(0)	−0.03
9	0	−0.30	12.70	12.60	0	−0.04
10	0	−0.10	11.60	11.60	0	−0.09
11	0	0.08	12.90	12.90	0	−0.08
12	0	0.00	21.50	21.40	0	0.07
13	0	0.11	19.80	19.90	0	−0.06
14		1.89	9.50	9.49	1	1.07
15		1.61	13.50	13.50	1	0.99
16		1.63	(12.70)	12.60	1	1.03
17		1.23	20.50	20.50	1	1.01
18		1.43	11.50	11.40	1	0.74

[a] See Table I for designations. Points in parentheses are known values free floated on test vector.

That so many successful or moderately successful rotations were observed is partially a consequence of the large variety of vectors tested and partially a consequence of the fact that many vectors may not be factors from a chemical viewpoint yet may be perfectly valid vectors mathematically. SBK, for example, is difficult to justify from our current knowledge of interactions, yet this vector, as we shall show later, seems to be equivalent to the first and most important eigenvector. Whether such vectors are true factors or are only correct in a mathematical sense poses interesting problems for both the theoretician and the experimentalist. Without quite detailed models for molecular interactions and a theory that allows one to transform from one set of interaction parameters to another, speculation is premature. Factor analysis seems to offer a valuable bridge between theory and experiment.

Uniqueness tests

The results of uniqueness tests, a standard vector in which the element of interest is given a value of "1" and all other elements in the vector are assigned "0", for each solute and solvent are shown in Table I. For solutes that are relatively

similar we do not expect to find much uniqueness in the individual solutes. Only the solutes dimethyl ether (the solute having the smallest molecular weight and, therefore, having the most poorly anchored points on test vectors) and *tert.*-butyl methyl ether exhibit significantly large uniqueness values (0.89 and 0.59, respectively). Aside from the vinyl ethers, correlations between groups of ethers were lacking. On a given vinyl ether uniqueness test, the other vinyl ethers tended to have predicted values almost as high as the solute being tested, consistent with the successful rotation for the unsaturation uniqueness vector mentioned above.

When the solute isobutyl methyl ether was included in the data matrix, atypical results were obtained. With four, five and even six factors the point corresponding to this ether and the solvent di-(2-ethyl) sebacate was predicted with errors exceeding 20 RI units. Only with seven or more factors did the error become small. In a five-factor space, the ether had a quite low uniqueness value; with six or more factors, however, the uniqueness exceeded 0.80. In the light of this unusual behavior, isobutyl methyl ether was not incorporated in the study reported here.

Reproductions using real vectors

One of the central objectives of theoretical chemistry is the identification of the complete set of interactions operative in a given problem; in terms of factor analysis, we would like to find a set of vectors that span the interaction space. Using the reproduction scheme of factor analysis, we can routinely assess the ability of a given set of vectors to reproduce the original data matrix. A set of test vectors is used to reproduce the data matrix in the same way as a set of abstract eigenvectors is employed in the reproduction scheme to determine the number of factors. The program has been modified to run all possible combinations of a chosen set of vectors, *e.g.*, the 33 physically and chemically significant vectors in Table II or the rows and the columns of data from the original matrix, in the reproduction scheme for any chosen number of factors. We now shall demonstrate procedures that allow one to estimate the relative importance of each test vector in each factor space, to estimate possible associations between the set of real vectors and the set of abstract eigenvectors, and finally to pinpoint sets of relatively equivalent real vectors. Though the approach may consume considerable computational time, the additional information acquired considerably extends the scope of factor analysis.

Using physically and chemically significant vectors

All combinations for one through six factors, *i.e.*, all combinations of 1, 2, 3, 4, 5 and 6 test vectors listed in Table II, were run in the reproduction scheme. A summary of the better combinations for factors 1–6 is given at the bottom of Table IV. The cut-off values for the a.r.a.e. are arbitrary. A relatively low cut-off would emphasize the more important vectors: we attempted to choose cut-offs that would bring out the key vectors and at the same time demonstrate the equivalence or lack of equivalence among vectors. Each value in the main body of Table IV gives the percentage of times the specified row element appeared in combinations having an a.r.a.e. less than the chosen cut-off values for the factor space designated by the column number.

For a given number of factors, the more important vectors are probably those which have the higher percentages. For example, in a six-factor space the vector

TABLE IV

PERCENTAGE OF BETTER REPRODUCTIONS CONTAINING SPECIFIC PHYSICAL VECTORS

Physical vector*	Number of factors used in reproduction					
	1	2	3	4	5	6
MW	0	3	5	1	17	8
MR	0	3	5	0	28	15
TN	0	0	8	10	22	21
CO	0	0	8	10	22	21
CN	0	0	10	12	42	36
HN	0	0	8	8	24	15
FC	0	15	6	5	0	0
FK	0	3	3	1	9	
BC	0	24	12	15	42	14
BK	0	30	21	8	10	13
UN	0	9	12	9	22	7
DU	0	0	0	7	0	17
TU	0	0	0	18	47	46
MU	0	3	9	35	49	2
UU	0	0	3	8	13	2
VU	0	0	5	9	21	15
CD	0	0	5	16	10	29
CR	0	0	0	11	0	49
HA	0	0	7	4	0	3
LA	0	0	7	18	6	4
SA	0	0	14	13	7	16
RMW	0	9	11	16	6	4
RMR	0	6	9	20	0	5
RTN	0	6	7	10	2	27
RCO	0	6	7	10	2	27
RCN	0	3	10	1	0	9
RHN	0	6	14	12	0	2
RFC	0	0	1	0	3	2
STN	0	9	5	3	0	18
SCO	0	9	5	4	0	23
SCN	0	9	5	4	0	29
SBK	100	46	75	99	100	100
LCN	0	0	8	6	0	13
Total number of combinations	33	528	5456	40920	237336	1107568
Cut-off (a.r.a.e.)	50.0	40.0	25.0	15.0	9.3	6.5
Number of reproductions less than cut-off	1	33	262	399	100	2010
Best combination: a.r.a.e.	47.8	18.5	12.5	9.8	8.3	5.3
Vectors involved**	SBK	SBK, RHN	SBK, LA, RMW	SBK, SA, BC, CD	SBK, VU, TU, BC, (TN, CO)	CN, TN, CD, SBK, CR, (STN, SCO, SCN)
Largest error	375	86	98	45	48	34

* See Table II for designations.
** Vectors in parentheses for five and six factors gave equivalent reproductions.

boiling point (°K) squared (SBK) occurs in essentially all the better reproductions, and the vectors chain ratio, *tert.*-butyl uniqueness and carbon number seem also to be relatively important. If for each vector the Exner value in Table II is compared to the percentage in Table IV, a fair positive correlation is noted between percentage and the reciprocal of the Exner value. For example, SBK has the highest percentage and the lowest Exner value. The chain ratio vector is a noticeable exception, having large values for both properties.

SBK appears to be an important vector in each of the factor spaces, having the highest percentage in every space (including 100% representation in the 1-, 5- and 6-factor spaces). Furthermore, SBK is incorporated in each of the best combinations in Table IV. We assume that SBK is more nearly equivalent to the first abstract eigenvector than any other physically significant vector. (From the values of the abstract eigenvalues —25.0, 0.035, 0.0023, etc.— the first eigenvector dominates the space.)

Examination of the data in Table IV also allows one to speculate on associations between particular abstract eigenvectors and the physical vectors. When going from one factor space to the next larger one, if certain physical vectors exhibit a marked increase in percentage, then those vectors quite possibly are more nearly equivalent than the other physical vectors to the abstract eigenvector added to the space. SBK, with 100% representation in the one-factor space, is clearly associated with the first eigenvector. Going from one to two factors, we conclude that boiling point (°K) or boiling point (°C), for which the percentages increase from 0 to 30 and 24, respectively, or perhaps sum adjacent (SA) is most probably analogous to the second eigenvector. None of the physical vectors except SBK (which is already accounted for) shows a large increase in percentage going from the two- to the three-factor space. Methyl uniqueness appears to be most nearly associated with the fourth eigenvector, molar refraction, carbon number, boiling point (°C) or *tert.*-butyl uniqueness, with the fifth eigenvector, and chain ratio or possibly carbon number squared, with the sixth eigenvector. The dependence of this approach upon the cut-off and the nature of the set of vectors selected is yet to be established.

If two vectors are equivalent, reproductions involving the two vectors and the same sets of other vectors should lead to equivalent reproduction errors. We illustrate this approach by referring to the summary of better reproductions in Table IV (though lack of equivalence might also show from a study of the poorer reproductions). If the numerical patterns for two vectors in Table IV are quite similar, we have evidence for the equivalence of two vectors which can be substantiated, *e.g.*, by comparing in a six-factor space the reproductions of the two vectors in combination with common sets of five other vectors. From Table IV, the sets TN and CO, and RTN and RCO contain quite equivalent vectors; and the sets STN, SCO and SCN, and RCN and LCN appear to contain fairly equivalent vectors. Reproductions with selected sets of six vectors confirm these conclusions. That so many reproductions gave quite similar a.r.a.e. values further implies that there is considerable equivalence among the set of physical vectors in Table II.

At the bottom of Table IV we list those vector combinations that gave the most accurate reproductions. In most cases, vectors required for the best combinations have relatively high percentages. For example, with six factors, of the eight vectors having the highest percentage, five vectors are incorporated in the best reproduction.

The a.r.a.e. for the best six-factor reproduction (5.3) is less than twice the average experimental error in the data points. Thus a reasonably complete empirical solution to the solute part of the ether–solvent problem has been formulated from factor analysis. Verification or rejection of such solutions must await the development of considerably expanded theories of solution. If factor analysis is applied to data that can be predicted adequately from present theory[7,11], even better predictions of experimental data are possible.

Using solute-associated row vectors from original matrix

Another way of finding real vectors that span the factor space is to utilize all combinations of rows or columns of data from the original matrix. In the factor analysis of hydrocarbon retention indices[5], for example, one set of solute-associated data vectors gave an a.r.a.e. smaller than the precision of the data points. For the ether space the best reproduction employing rows of data from the original matrix (each row being associated with a particular solute) had an a.r.a.e. for six factors of 2.8 (see Table V), within experimental error for this problem. The best reproduction

TABLE V

PERCENTAGE OF BETTER REPRODUCTIONS CONTAINING SPECIFIC ROW VECTORS FROM DATA MATRIX

Solute row*	Number of factors used in reproduction					
	1	2	3	4	5	6
1	0	20	0	32	23	100
2	6	30	26	33	45	1
3	6	30	26	14	3	29
4	6	0	0	15	29	63
5	6	20	31	13	13	25
6	6	10	0	0	13	37
7	6	0	0	17	0	8
8	6	0	0	0	52	28
9	6	0	0	35	16	22
10	6	0	17	35	55	39
11	6	0	13	3	26	28
12	6	40	17	18	36	40
13	6	30	35	63	65	52
14	6	0	4	24	16	16
15	6	0	35	21	10	24
16	6	0	0	8	52	33
17	6	20	61	42	36	41
18	6	0	35	27	10	17
Total number of combinations	18	153	816	3060	8568	74256
Cut-off (a.r.a.e.)	50.0	11.0	6.8	5.0	3.8	3.3
Number of reproductions less than cut-off	17	10	23	78	31	233
Best combination:						
a.r.a.e.	26.5	9.7	6.1	4.5	3.8	2.8
Vectors involved*	6	3, 11	11, 17, 18	2, 10, 13, 17	2, 8, 10, 13, 16	1, 4, 6, 10, 13. 16
Largest error	266	37	35	24	21	21

* See Table I for designations.

included rows representing a range of ether types typical of a set a chemist would pick to span the space; the best set includes the smallest ether as well as ethers containing isoalkyl, *tert.*-alkyl and vinyl groups. As was the case with reproductions using physical vectors, many combinations gave nearly equivalent reproductions. The space does not have a unique solution from the standpoint of solute types.

An analysis of the trends in percentage with the number of factors in Table V leads one to tentatively associate certain data rows with particular abstract eigenvectors. The first eigenvector appears to be equally represented by all the solutes (excluding dimethyl ether); no one ether plays a dominant role, as expected from the similarity of the solutes. Based on the major increases in percentages as we go from left to right in Table V, the dipentyl ether-associated row seems most nearly equivalent to abstract eigenvector number two, the butyl vinyl, allyl ethyl or 2-ethyl-1-hexyl vinyl ethers, to eigenvector three, the isopropyl propyl, dimethyl or diisopentyl ethers, to eigenvector four, the dipropyl or isobutyl vinyl ethers, to eigenvector five, and the dimethyl or *tert.*-butyl methyl ethers, to eigenvector six. Four of the ethers represented in the best combination for six factors are included in the above list. The uniqueness of dimethyl and *tert.*-butyl methyl ethers for six factors (but not in smaller spaces) illustrates a possible correlation between uniqueness and association with an abstract eigenvector. Excepting the data for butyl vinyl and allyl ethyl ethers, no two sets of percentages in Table V were very similar.

Using solvent-associated columns from original matrix

Reproductions on the transpose of the data matrix involving all combinations of vectors associated with the stationary-phase solvents were performed to ascertain which solvent sets best span the interaction space. The best combination in the six-factor space (see Table VI) had an a.r.a.e. of 2.9, within experimental error. As with the solutes, many combinations produce essentially equivalent reproductions. From Table VI we see that, based upon the high-percentage criteria, diglycerol, Dow Corning FS 1265 fluid, Hyprose SP80 and Zonyl E7 are particularly important in the six-factor space. The same solvents are also the most unique solvents for six factors (see Table I).

Equivalences between abstract eigenvectors and solvent-associated data vectors, as well as the equivalences within the set of data vectors, can be estimated from the patterns in Table VI. For one factor, all columns are essentially equivalent. For two factors, the data vectors most probably equivalent to the second eigenvector are associated with diglycerol or Hyprose SP80, those equivalent to the third eigenvector with Igepal CO 890 (assuming diglycerol is already accounted for), those equivalent to the fourth eigenvector with Quadrol especially, or Carbowax 20M, those equivalent to the fifth eigenvector with Dow Corning FS 1265 fluid or Zonyl E7, and those equivalent to the sixth eigenvector with Flexol 8N8 or diisodecyl phthalate. Some groups of chemically similar coatings, such as the set di-(2-ethyl) sebacate and dioctyl sebacate, and the set SE-30, SE-31 and SE-52 exhibited nearly the same pattern going across Table VI, but, for example, the percentage data for Carbowax 600 and Carbowax 20M are quite dissimilar. An extended study of this type should be quite useful for characterizing coatings and for choosing best sets of columns for general separations, both problems of great practical importance[12,15].

TABLE VI

PERCENTAGE OF BETTER REPRODUCTIONS CONTAINING SPECIFIC COLUMN VECTORS FROM DATA MATRIX

Solvent column[*]	Number of factors used in reproduction					
	1	2	3	4	5	6
A	4	0	4	4	9	3
B	4	14	21	7	14	16
C	4	0	17	52	38	30
D	4	6	10	4	6	2
E	4	8	7	15	13	9
F	0	27	86	100	100	85
G	4	6	4	0	0	20
H	4	8	5	4	8	4
I	4	5	8	0	10	9
J	4	2	0	0	69	48
K	4	3	4	0	0	36
L	4	22	0	15	39	56
M	4	0	21	22	29	16
N	4	6	8	15	16	11
O	4	0	17	19	14	17
P	4	3	0	0	1	14
Q	4	3	0	85	40	30
R	4	5	10	19	17	14
S	4	6	12	11	17	13
T	4	6	10	15	19	27
U	4	3	10	0	4	8
V	4	13	13	11	7	6
W	4	14	9	0	3	1
X	4	6	12	4	6	13
Y	4	19	13	0	28	49
Total number of combinations	25	300	2300	12650	53130	177100
Cut-off (a.r.a.e.)	50.0	15.0	8.0	4.8	4.2	3.5
Number of reproductions less than cut-off	24	64	92	27	118	338
Best combination: a.r.a.e.	28.2	10.6	6.0	4.5	3.5	2.9
Vectors involved[*]	W	F, X	B, F, V	C, F, Q, S	C, F, J, L, S	C, F, K, L, T, Y
Largest error	275	37	26	26	16	13

[*] See Table I for designations.

REFERENCES

1 P. H. Weiner and D. G. Howery, *Can. J. Chem.*, 50 (1972) 448.
2 P. H. Weiner and D. G. Howery, *Anal. Chem.*, 44 (1972) 1189.
3 P. H. Weiner, C. J. Dack and D. G. Howery, *J. Chromatogr.*, 69 (1972) 249.
4 P. H. Weiner and J. F. Parcher. *Anal. Chem.*, 45 (1973) 302.
5 D. G. Howery, *Anal. Chem.*, 46 (1974) 829.
6 D. G. Howery, P. H. Weiner and J. S. Blinder, *J. Chromatogr. Sci.*, 12 (1974) 366.
7 P. H. Weiner, H. L. Liao and B. L. Karger, *Anal. Chem.*, 46 (1974) 2182.
8 J. H. Kindsvater, P. H. Weiner and T. J. Klingen, *Anal. Chem.*, 46 (1974) 982.
9 P. H. Weiner, E. R. Malinowski and A. R. Levinstone. *J. Phys. Chem.*, 74 (1970) 4537.
10 R. A. Keller, *J. Chromatogr. Sci.*, 11 (1973) 49.
11 P. H. Weiner and E. R. Malinowski, *J. Phys. Chem.*, 75 (1971) 3160.
12 P. H. Weiner and J. F. Parcher, *J. Chromatogr. Sci.*, 10 (1972) 612.
13 D. G. Howery, *Bull. Chem. Soc. Jap.*, 45 (1972) 2643.
14 M. C. Weiner and P. H. Weiner, *J. Med. Chem.*, 16 (1973) 655.
15 P. H. Weiner, *J. Amer. Chem. Soc.*, 95 (1973) 5845.
16 W. O. McReynolds, *Gas Chromatographic Retention Data*, Preston Technical Abstracts Co., Niles, Ill., 1966.
17 O. Exner, *Collect. Czech. Chem. Commun.*, 37 (1966) 3222.
18 D. H. McCloskey and S. J. Hawkes, *J. Chromatogr. Sci.*, 13 (1975) 1.