DOCUMENT RESUME

ED 387 496 TM 023 598

AUTHOR Davison, Mark L.; And Others
TITLE Utilizing Profile Analysis via Multidimensional Scaling To Ascertain Patterns in Course-Taking: Mathematics and Science Course-Taking Patterns.
SPONS AGENCY Department of Education, Washington, DC.
PUB DATE Apr 95
CONTRACT R999B40010
NOTE 63p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 18-22, 1995). Additional funding provided by the Jack C. Merwin Endowment Fund of the University of Minnesota.
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Course Selection (Students); *Mathematics Instruction; *Multidimensional Scaling; National Surveys; Profiles; *Research Methodology; Research Needs; *Science Instruction
IDENTIFIERS *National Assessment of Educational Progress; *Profile Analysis Approach

ABSTRACT

This paper describes a methodology, called Profile Analysis via Multidimensional Scaling (PAMS), designed to identify major patterns of variables and to study the relationships into which those patterns enter. Patterns of variables are here called profiles. The PAMS procedure has been adapted to characteristics of the National Assessment of Educational Progress (NAEP) data. Statistical Package for the Social Sciences (SPSS) and Statistical Analysis System (SAS) templates to implement the methodology have been or are being developed. Adaptations of the general PAMS approach for the study of course-taking patterns are described. Multidimensional scaling is being used to identify patterns of course-taking in mathematics and science, and preliminary results from these analyses are reported. The direction of future research is outlined. Four tables and 12 figures illustrate the analyses. (Contains 20 references.) (Author/SLD)

Utilizing Profile Analysis via Multidimensional Scaling to Ascertain Patterns in Course-taking: Mathematics and Science Course-taking Patterns

Mark L. Davison, Ernest C.Davenport Jr., John Bielinski,

Shuai Ding, Haijiang Kuang, Fuliao Li,

and Katherine Seiden

Department of Educational Psychology

University of Minnesota

Requests for copies should be sent to Mark L. Davison and Ernest C. Davenport Jr., Department of Educational Psychology, University of Minnesota, 178 Pillsbury Dr. S.E., Minneapolis, MN 55455. This research was supported by Grant No. R999B40010 from the U.S. Department of Education and by the Jack C. Merwin Endowment Fund of the University of Minnesota.

Abstract

This paper describes a methodology, called PAMS (Profile Analysis via Multidimensional Scaling) designed to identify major patterns of variables and to study the relationships into which those patterns enter. Patterns of variables are here called profiles. The PAMS procedure has been adapted to characteristics of the NAEP data. SPSS^{R} and SAS^{R} templates to implement the methodology have been or are being developed. Adaptations of the general PAMS approach for the study of course-taking patterns are described. Multidimensional scaling is being used to identify patterns of course-taking in mathematics and science, and preliminary results from these analyses are reported.

Utilizing Profile Analysis via Multidimensional Scaling to Ascertain Course-taking Patterns: Mathematics and Science Course-taking Patterns

Mark L. Davison, Ernest C. Davenport Jr., John Bielinski,
Shuai Ding, Haijiang Kuang, Fuliao Li,
and Katherine Seiden
Department of Educational Psychology
University of Minnesota

Many of the research questions to which NAEP data can be applied involve patterns of student characteristics, patterns of instructional factors, patterns of teacher characteristics, etc. For instance, one can ask questions such as the following. "What student attitude and behavior patterns are associated with better performance?" "Do similar patterns of effort exist for minority and majority students?" "Which behavioral and home environment patterns are associated with school success?"

If behaviors, attitudes, and other characteristics exist as patterned wholes (rather than unintegrated, aggregates of characteristics), then it is worthwhile for researchers to investigate these patterns and their relationships to outcome variables. Davison (Davison & Skay, 1991; Davison, 1994; Davison, Gasser & Ding, 1995) describes a multidimensional scaling (MDS) method for identifying the major patternings of variables in a population. The method is called Profile Analysis via Multidimensional Scaling (PAMS) and the major patternings are called "prototypical profiles." In PAMS, each multidimensional scaling dimension corresponds to a prototypical profile of student (teacher, school, etc.) characteristics. Each observation is characterized by a level parameter and a set of weights, one weight for each dimension, which index the degree of resemblance between the observations' profile and a prototypical profile dimension.

The concept of a prototypical profile reminds one of typologies. Cluster analysis, rather than multidimensonal scaling, has been the major technique used to study typologies (i.e. Bailey, 1994; Lorr, 1994). Cluster analysis divides observations (people) into discrete categories. MDS and PAMS represent structure in terms of continuous dimensions corresponding to prototypical profiles. In cluster analysis, one most commonly talks about observations as belonging or not belonging to clusters. In PAMS, each observation corresponds to an observed profile of scores, and one talks about an observation's degree of resemblance to a prototypical profile. Observed profiles are represented as linear combinations of PAMS prototypical profiles; therefore, an actual profile may closely resemble a single prototypical profile or it may be a mixture of several prototypical profiles.

One area in which patterns are potentially important is course-taking in mathematics and science. Recent studies have examined the relationship between course-taking and achievement. For the most part, these studies have counted the number of formal or pre-college math and science courses taken and related that number to achievement (i.e. Davenport, 1992; Jones, Burton & Davenport, 1984; Jones, Beduis & Davenport, 1985; Noble & McNabb, 1984; Welch, Anderson & Harris, 1985). Other than distinguishing formal and informal mathematics courses, these studies have paid relatively little attention to the pattern of the mathematics and science coursework. Further progress in the study of mathematics and science course-taking requires identification of the major course-taking patterns, so that one can study the relationship of achievement to both type and number of courses. Therefore, one reason for studying course-taking patterns in mathematics and science is to refine our knowledge of those patterns. This refinement can directly lead to a better understanding of the relationship between course-taking and achievement.

The second reason for studying mathematics and science, particularly mathematics, is that mathematics has some well-defined sequences. These sequences provide a partial validity check on the method. That is, the method should successfully recover these sequences if it is functioning appropriately.

This paper outlines the basic PAMS model, which has guided much of this research, and the analysis based on that model. We then describe some modifications of the basic PAMS procedure which we have used for analysis of transcript data. Finally, we present some sample results based on math and science course-taking data. SPSS[R] for WINDOWS [R] and SAS[R] templates for fitting the PAMS model have been prepared and are available from the authors (Davison & Davenport, 1994; Davenport & Davison, 1995)[1]. A full PAMS analysis involves both deriving prototypical profiles via MDS and estimating correspondence weights. In this paper, the focus is on the derivation of the prototypical profiles.

The PAMS model and Analysis

One way to explain the PAMS model is to contrast it with the well known factor analytic approach. The description of the factor model will be rather superficial, because we present it only as a contrast with PAMS. In contrast to factor analysis, with its individual differences perspective, PAMS is more idiographic. To see the difference, let's look at Figure 1.

Insert Figure 1 about here.

Figure 1 contains a data matrix in which rows represent people (the observations) and columns represent measures. Let an element of the matrix, $\underline{m}_{si}$, be the score of subject s on individual differences measure $\underline{i}$. In factor analysis, the focus is on the columns, each of which is an individual differences variable, measure $\underline{i}$. Let $\mathbf{m}_i$ be the vector of scores in column $\underline{i}$. Factor analysis posits a small set of "latent" individual differences variables, the factors or components $\mathbf{f}_k$ ($\underline{k}$ = 1,...,$\underline{K}$), such that the observed individual differences variables

[1]SPSS is a registered trademark of SPSS, Inc. Windows is a registered trademark of Microsoft, Inc. SAS is a registered trademark of SAS Institute, Inc.

$\mathbf{m}_i$ can be represented as linear combinations of the $\mathbf{f}_k$:

$$\mathbf{m}_i = \Sigma_k \underline{a}_{ik} \mathbf{f}_k + \mathbf{e}_i \qquad (1)$$

where $\underline{a}_{ik}$ is a weight for individual differences measure $\underline{i}$ on factor $\underline{k}$, and $\mathbf{e}_i$ is a vector of deviations which includes both error, and if the common factor model rather than the components model, specific factor variance. These "latent" individual differences variables are represented to the right of the data matrix in Figure 1.

In the PAMS model, (Davison & Skay, 1991; Davison, 1994) the focus is on the rows. Let $\mathbf{m}_s$ be the row vector of scores for subject $\underline{s}$. Whereas factor analysis posits a set of latent individual differences variables, the factors, which can account for the observed column variables, PAMS posits a latent set of profiles such that each observed row profile can be represented as a linear combination of the latent row profiles:

$$\mathbf{m}_s = \Sigma_k \underline{w}_{sk} \mathbf{x}_k + \mathbf{c}_s + \mathbf{e}_s \qquad (2)$$

where $\underline{w}_{sk}$ is a weight for subject s on prototypical profile $\underline{k}$, $\mathbf{c}_s$ is a constant vector for subject $\underline{s}$, and $\mathbf{e}_s$ is a vector of deviations. Except for the use of some vector notation, Equation 2 is what has been called elsewhere the vector model (e.g. Tucker, 1960; Davison & Skay, 1991), because each subject's profile of data $\mathbf{m}_s$ can be represented as a vector in the space spanned by the prototypical profiles $\mathbf{x}_k$. As we shall see, each prototypical profile corresponds to a dimension in a multidimensional scaling analysis.

The individual differences parameters in Equation 2 deserve some explanation. The subject weight, w_{sk}, is a weight for the observed profile of subject s on prototypical profile k. It indexes the direction and degree of linear association between the actual profile of subject s and the latent, prototypical profile k. If w_{sk} is positive and large, then subject s tends to have his or her highest (and lowest scores) on the same variables as prototypical profile k. If w_{sk} equals zero, then there is no linear association between observed profile s and prototypical profile k. If w_{sk} is negative, then the <u>highest</u> scores for observed subject s tend to fall on the measures which have the <u>lowest</u> scores in prototypical profile k. The parameter c_s is a level parameter which indexes the overall height of the profile for subject s. Thus, the level of profile s is represented by a single parameter, c_s, and the shape (pattern and scatter) are represented as a linear combination of the prototypical profiles $\mathbf{x}_k$.

Both the factor model of Equation 1 and the PAMS model of Equation 2 are linear models. In the factor model, the weights are associated with observed individual differences variables, and each observed individual differences variable is represented as a linear combination of the latent individual differences variables, the factors. In the PAMS model, weights are associated with subjects, and each subject's observed profile of responses is represented as a linear combination of latent, prototypical profiles, $\mathbf{x}_k$, plus a level parameter, $\underline{c}_s$.

The difference between the factor and PAMS models are shown schematically in Figure 1. The factor model represents the columns of the data matrix as linear combinations of the factor variables shown to the right of the data matrix. The PAMS model represents the rows of the data matrix as linear combinations of the prototypical profile variables shown below the data matrix.

Estimating Parameters in the PAMS model. There are a variety of methods for estimating parameters in the PAMS model itself or closely related models (e.g. Weller & Romney, 1990). We have employed a procedure outlined by Davison and Skay (1991). It involves two steps: estimating the prototypical profiles, $\mathbf{x}_k$, using MDS and then estimating the subject parameters $\underline{w}_{sk}$, $\underline{c}_s$ through regression of the subject data, $\mathbf{m}_s$ onto the prototypical profiles $\mathbf{x}_k$.

In applying MDS to estimate the prototypical profiles, the approach outlined by Davison and Skay (1991) relies on the standard, squared Euclidean distance measure defined over pairs of measures ($\underline{i}$, $\underline{i}'$):

$$\delta^2_{ii'} = (1/\underline{S})\Sigma_s(\underline{m}_{si} - \underline{m}_{si'})^2 \qquad (3)$$

where $\underline{S}$ is the number of subjects. Let $\underline{I}$ be the number of measures (or columns in Figure 1). If one assumes homogeneity of error variances across people $\underline{s}$ and measures $\underline{i}$ and that

$$0 = \Sigma_i \underline{x}_{ik} \quad \text{for all } \underline{k}, \tag{4a}$$

$$S = \Sigma_s \underline{w}^2_{sk} \text{ for all } \underline{k}, \tag{4b}$$

$$0 = \Sigma_s \underline{w}_{sk} \underline{w}_{sk'} \text{ for all } (\underline{k}, \underline{k}'), \tag{4c}$$

$$0 = \Sigma_s \underline{e}_{si} \underline{e}_{si'} \text{ for all } (\underline{i}, \underline{i}') \tag{4d}$$

and

$$0 = \Sigma_s \underline{e}_{si} (\Sigma_k \underline{w}_{sk} \underline{x}_{ik} + \underline{c}_s), \tag{4e}$$

then

$$\delta^2_{ii'} = \Sigma_k (\underline{x}_{ik} - \underline{x}_{ik'})^2 + 2\sigma^2(e). \tag{5}$$

In short, the mean squared Euclidean distances between pairs of observed subject profiles, defined in Equation 3, is within an additive constant of squared Euclidean distances between points representing measures ($\underline{i}$, $\underline{i}'$) in a space defined by the $\underline{K}$ prototypical profiles. As a result, MDS can be used to estimate the prototypical profiles. Each dimension will correspond to a profile. In the results section below, we will illustrate the interpretation of profiles (dimensions) resulting from such an analysis.

Once the MDS has been completed, the $\underline{K}$ prototypical profile vectors, $\mathbf{x}_k$, have been estimated. Let $\mathbf{X}$ be a matrix with $\underline{I}$ rows and $\underline{K}$+1 columns. The first $\underline{K}$ columns contain the MDS solution; that is, the $\underline{K}$ prototypical profile column vectors $\mathbf{x}_k$. The last column is a vector of 1's. Let **W** be a matrix with $\underline{S}$ rows, one for each subject, and $\underline{K}$+1 columns. The first $\underline{k}$ columns contain the subject weights $\underline{w}_{sk}$ and the last column contains the level parameters $\underline{c}_s$. Then Equation 2 can be rewritten in matrix form as

$$\mathbf{M} = \mathbf{W}\mathbf{X}' \qquad (6)$$

where **M** is the $\underline{S}$x$\underline{I}$ matrix of data in Figure 2. Given the estimate of **X** provided by MDS and the raw data matrix **M**, the unknown matrix of subject weights and level parameters can be estimated as

$$\mathbf{W} = \mathbf{M}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}. \qquad (7)$$

Davison and Davenport (1994; Davenport & Davison, 1995) have created an SPSS[R] and SAS[R] templates which executes the steps outlined above.

While the full PAMS analysis consists of both steps-- estimating prototypical profiles $\mathbf{x}_k$ and subject parameters, $\underline{w}_{sk}$ and $\underline{c}_s$-- the focus in this paper is on estimation of the prototypical profiles, particularly as that estimation is applied to transcript data.

<u>Deviations from the PAMS Model</u>. The PAMS model assumes continuous variables $\underline{m}_{si}$. The transcript data on which we have been focusing is dichotomous. That is, we have coded the data so

that $\underline{m}_{si} = 1$ if student s has taken course $\underline{i}$ long enough to earn at least .20 Carnegie units, and $\underline{m}_{si} = 0$ otherwise. Furthermore, the proportions of students taking some of the courses are very small. Therefore, based in part on trial and error, we have modified the PAMS approach for estimation of prototypical course-taking profiles as follows.

Instead of using the squared Euclidean distance measure of proximity, we have been using $(1 - \underline{J}_{ii'})$ where $\underline{J}_{ii'}$ is the Jaccard measure, also known as the similarity ratio, defined over all possible pairs of courses ($\underline{i}$, $\underline{i}'$). Figure 2 shows the usual 2x2 contingency table for two courses. Let $(1 - \underline{J}_{ii'}) = \tau_{ii'}$ be the proximity measure used to obtain the results below:

$$\tau_{ii'} = (\underline{b} + \underline{c})/(\underline{a} + \underline{b} + \underline{c}). \qquad (8)$$

In words, the numerator of $\tau_{ii'}$ is the number of people who took one and only one of the two courses ($\underline{i}$, $\underline{i}'$)'; the denominator is the number who took at least one. Of those who took at least one of the two courses, $\tau_{ii'}$ is the proportion who took only one.

Insert Figure 2 here.

When the data are dichotomous, $\tau_{ii'}$ is related, but not identical to, the squared Euclidean distance in Equation 3. In Equation 3, $\underline{S}$ is the total number of subjects. If $\underline{S}$ on the right side of Equation 3 were replaced by $\underline{S}_{ii'}$ (the number of people who took at least one of the two courses) and the variable $\underline{m}_{si}$ is dichotomous, then the right side of Equation 3 would equal $\tau_{ii'}$.

Using $\tau_{ii'}$ as the proximity measure in our analyses below is quite ad hoc. But, as we shall see, it yields generally interpretable results. Theoretically, however, we have no model of the data $\underline{m}_{is}$ to suggest as a justification for $\tau_{ii'}$. That is, we have no counterpart of Equation 2 to offer from which one can derive $\tau_{ii'}$ as an appropriate proximity measure.

Since $\tau_{ii'}$ would seem to have an absolute zero, we have assumed ratio level data in our MDS analyses, also an option in the PAMS SPSS[R] template. Furthermore, where a Varimax rotation (Kaiser, 1958) yielded a more interpretable solution, we have performed such a rotation.

Standard Errors. Few applications of MDS include estimates of standard errors around scale values. We have developed a computationally intensive approach to estimation of errors using resampling. In this approach, we begin with an estimate of the MDS coordinate matrix **X** derived from the full sample; that is, in computing the proximity matrix on which **X** is based, the population weights are used. Here **X** is an unaugmented matrix with $\underline{I}$ rows and $\underline{K}$ columns.

Next, using the NAEP transcript resampling weights, we repeat the MDS scaling process $\underline{B}$ times. Let $\mathbf{X}^b$ be the MDS solution from replication $\underline{b}$. Like **X**, $\mathbf{X}^b$ will have $\underline{I}$ rows and $\underline{K}$ columns. To derive the proximity matrix for the $\underline{b}$th replication, the $\underline{b}$th set of sampling weights is applied to the raw data in computing the proximity matrix on which the bth MDS solution is based. Let x_{ik} be the coordinate of course i along prototypical dimension k in $\mathbf{X}$ and let x^b_{ik} be the corresponding element of $\mathbf{X}^b$. Then, the sampling variability of x_{ik} is estimated as follows:

$$s_{ik} = [\Sigma_b(x_{ik} - x^b_{ik})^2/(B - 1)]^{1/2}.$$

Because MDS solutions are determined only up to a rotation and uniform stretching or shrinking of the axis, we have applied a rotation to congruence (Davison, 1983; Gower, 1971, 1975; Schonemann & Carroll, 1970) in the process of deriving matrix $\mathbf{X}^b$. That is, let $\mathbf{X}^{b*}$ be the raw MDS coordinate matrix derived from the proximity matrix using the bth resampling weights. Let $\mathbf{T}$ and c be the orthonormal transformation matrix and scalar constant relating $\mathbf{X}^b$ and $\mathbf{X}^{b*}$: $\mathbf{X}^b = c\mathbf{X}^{b*}\mathbf{T}$. Let $\boldsymbol{\Omega}$ and $\mathbf{V}$ be the characteristic roots and vectors of $\mathbf{X}'\mathbf{X}^{b*}\mathbf{X}^{b*\prime}\mathbf{X}$. Then $\mathbf{T} = \mathbf{X}^{b*\prime}\mathbf{X}\mathbf{V}\boldsymbol{\Omega}^{-1/2}\mathbf{V}'$. If we let $\mathbf{X}^{b**} = \mathbf{X}^{b*}\mathbf{T}$, then $c = (\Sigma_{(i,k)}x^{b**}_{ik}x_{ik})/(\Sigma_{(i,k)}x^{b**2}_{ik})$.

Methods

The analyses reported below are based on the 1990 High School Transcript Study conducted by Westat, Inc. for the National Center for Education Statistics.

Sample.

Over 23,000 students who graduated from American high schools in 1990 were selected as the sample; 21,784 transcripts were received. The transcripts, along with other school materials, such as course catalogs and student handbooks, were solicited from approximately 283 public, Catholic, and other private schools representing a range of geographic areas and school sizes. The sample is diverse in terms of cultural background and ethnic composition. Legum, Caldwell, Goksel, Haynes, Hynson, Rust, and Blecher (1993) contains a description of the sample.

Course Coding System. Courses were coded using the Classification for Secondary School Courses (CSSC) coding system. Tables 1 and 2 show the courses included in the analyses below, along with the first two digits of its CSSC course code. The first two digits indicate the general content area of the course; e.g 26 = life sciences, 27 = mathematics, 40 = physical sciences, etc. We have assigned an abbreviation to each course used in later tables and figures, and these abbreviations are shown in the last columns of Tables 1 and 2.

Insert Tables 1 and 2 here.

In course titles, certain descriptive terms are used in the CSSC system:

- **Resource**--remedial; courses which serve as tutoring vehicles or subject area services
- **General Skills**--functional level
- **Basic**--simplified, fundamental, slower paced
- **Applied**--emphasis on application over theory,

application to work setting

- **Introductory**--basic, survey-level course
- **Informal**--more practical and occupational applications oriented
- **Formal**--middle-level; equal emphasis on theory and application
- **Unified**--special, inter/multi-disciplinary study
- **Review**--special preparations for SAT and ACT college entrance examinations
- **Honors**--courses which prepare students for post-secondary academic study
- **Advanced**--equivalent to Advanced Placement; preparation for College Board Advanced Placement examinations

Math Courses. In the CSSC, mathematics courses are designated 27XXXX. In addition, some mathematics related courses can be found in other academic areas, such as business (07XXXX) and computer science (11XXXX). There are also functional level courses (54XXXX) and resource courses (56XXXX).

With some exceptions, primarily resource courses, courses were dropped if they had less than 300 students. We began by looking only at courses in mathematics, 27XXXX, and then we expanded our search to include mathematics related courses with a significant enrollment. A few computer science and business courses were included as a result. Courses were eliminated or disregarded based on several criteria. Linear Algebra, Mathematics Tutoring, and Science Mathematics were all dropped due to low frequencies. Independent Study Mathematics, Other Mathematics, and Other Pure Mathematics were considered too vague to be meaningful in a scaling procedure. After reading the course description, Mathematics in the Arts was discarded as not being very mathematically oriented.

Where it seemed sensible, small enrollment courses were combined with other courses to avoid dropping them altogether. Students who took a course entitled Calculus and Analytic Geometry were credited with having studied Calculus and having studied Analytic Geometry. Similarly, Algebra and Analytic Geometry were combined with Algebra 3 and with Analytic Geometry. Trigonometry and Solid Geometry were combined with Trigonometry and with Solid Geometry. The individual courses Probability and Statistics were combined with the single course Probability and Statistics. Computer Math 1 and 2 were combined into Computer Math. Business Math 1 and 2, Agricultural Math, and Financial Math were all combined under the heading of Business Math. Functional Consumer Math, Functional Vocational Math, and Special Education Math were all collapsed into General Math Skills. Vocational Math and Technical Math were combined under the heading Technical Math. Basic Math 2, 3, and 4 were combined under the heading Basic Math 2. Resource General Math includes those who did and did not take it for credit.

Science Courses. Life sciences courses are designated 26XXXX and physical sciences courses are designated 40XXXX, but some science related courses can be found under other headings;

e.g engineering, agricultural science, etc. Courses outside the 26XXXX and 40XXX designations (including resource science courses) were dropped, primarily due to low enrollments in this sample. Three "miscellaneous" categories were created: Other Formal Physical Sciences, Other Formal Life Sciences, and Other Informal Physical Sciences. Advanced Physiology was combined with Physiology.

The analyses below are based on 53 variables for each student, 31 corresponding to the math courses in Table 1 and 22 corresponding to the science courses in Table 2. Each variable was scored dichotomously: 1 if the student had earned at least .25 Carnegie units and 0 otherwise.

Results

To illustrate coursework prototypical profiles, this section presents the MDS dimensions from an analysis of the mathematics and science coursework data. The proximity measure employed was $(1 - J_{ii'})$ where $J_{ii'}$ is the Jaccard coefficient for the courses (i, i'). That is, the proximity measure equals the number of people who took only one of the two courses expressed as a proportion of those who took at least one.

In MDS, dimensions can be reflected without loss of fit. It is completely arbitrary as to which end is positive and which end is negative. This means that there are potentially two profiles for each dimension. The first profile for a dimension is marked by courses at the positive end; we call this the dimension's profile. The second profile for a dimension is marked by courses at the negative end; we call this the dimension's mirror image profile. To illustrate the relationship between a dimension profile and its mirror image, we have plotted both (Figures 3 and 4) for Dimension 1 in Table 3. For all other dimensions in Tables 3 and 4, however, we have plotted only the dimension profile.

In what follows, dimension profiles (and mirror images) are interpreted in terms of their highest points or peaks. Particular attention has been paid to courses with scale values above 1.00; and these courses are identified in the figures below. Given the normalization used by ALSCAL whereby the sum of squared scale values in a solution equals IK, the root mean square of all scale values equals 1.00. Therefore, courses with scale values above one are above the root mean square of all scale values. As will become apparent, however, some of the courses with scale values above 1.00 figure more prominently in our interpretations than do others.

Mathematics Courses

Table 3 shows the four dimensional solution for the mathematics data obtained from the ALSCAL program (Young & Lewyckyz, 1979). The fit measures for this solution were S-STRESS = .363, STRESS = .258, and R^2 = .349. The first two of these are least squares, badness-of-fit measures which range between 0 (perfect fit) and 1 (total lack of fit). The last of these fit measures is the squared correlation between the model predicted distances and the data. Because the number of courses is relatively large and because the data were treated as ratio scaled data, rather than ordinal data, these fit measures are poorer than those obtained from most applications of MDS. The data were treated as ratio scaled data because this proximity measure seems to have an absolute zero (i.e. nobody takes one of the two courses without taking both); furthermore, inspection of the proximities suggested that the sizes of the intervals contained meaningful information.

Figure 3 shows the first dimension in Table 3 plotted as a profile. The highest three points in this profile correspond to Algebra 1 (Alg1), Geometry (Geo), and Algebra 2 (Alg2). This profile is most prominently marked by the standard three formal, high school mathematics courses. The other peaks in Figure 3 correspond to Trigonometry (Trig), Algebra 3 (Alg3), Analytic Geometry (AnGeo), and Technical Mathematics (Tech, which consists of applied algebra and geometry; technical training; numerical trigonometry). Because this profile is marked primarily by Algebra 1, Geometry, and Algebra 2, we have called it the Standard Formal Sequence.

Insert Figures 3 and 4 here.

Figure 4 shows the mirror image profile for Dimension 1 in Table 3. It has three prominent peaks, Unified Mathematics 1 (M1Unif), Unified Mathematics 2 (M2Unif), and Unified Mathematics

3 (M3Unif). These three courses present logic, algebra, geometry, trigonometry, and probability in a unified approach. We call this Dimension 1 Mirror Image Profile the Unified Sequence. Not surprisingly, the Unified Sequence and the Standard, Formal Sequence fall at opposite ends of a dimension, suggesting that those who study these subjects in a unified approach by taking M1Unif, M2Unif, and M3Unif do not then take separate courses in algebra, geometry, and trigonometry. A fourth peak in the profile, Basic Math 1 (BasM1) does not seem to fit with the remaining prominent courses in the profile.

The Dimension 2 Profile, shown in Figure 5, is marked most prominently by two calculus courses, Advanced Placement Calculus (CalAdv) and Calculus (Calc). Introduction to Analysis (IntAnal) and M3Uni, courses which commonly precede calculus, are the next most prominent courses in this profile. We call this the Calculus Profile, because it contains the calculus courses and courses which would precede them.

Insert Figure 5 here.

The three most prominent courses shown in Figure 5 at the negative end of Dimension 2, which form the mirror image profile, have little in common except that they fall outside any formal mathematics sequence. They are Pre-algebra (PreAlg), Business Mathematics (BusM), and General Mathematics 1 (M1Gen). We call this the Pre-formal Profile. The remaining three courses marked in this profile constitute a sequence in which algebra 1 is spread over two years; Algebra 1 Part 1 (AlgP1), and Algebra 1 Part 2 (AlgP2); and Geometry is studied in a more applied, simplified fashion, Informal Geometry (GeoInf). One or more of these three courses; AlgP1, AlgP2, and GeoInf; can follow one of the Preformal Sequence course (most probably PreAlg or M1Gen) in a students course of study.

The Dimension 3 Profile, Figure 6, is marked most prominently by three courses Computer Mathematics (CompM), Review Mathematics (MRev), and Probability and Statistics (ProbStat). Review Mathematics is described as college entrance exam practice, an overview course with reading questions. Computer Mathematics covers simple calculators, flow charts, elementary programming and mathematical applications. These appear to be courses which go beyond basic mathematics, but which do not cover standard topics in algebra, geometry, or trigonometry. We call this the Beyond the Basics Profile.

Insert Figure 6 here.

The Dimension 3 Mirror Image is marked by basic, informal mathematics sequences: General Mathematics 1 and 2 (M1Gen and M2Gen) and Basic Math 1 and 2 (BasM1 and BasM2) as well as Consumer Math (ConM). We call this the Informal Sequence Profile. Some of these courses tend to emphasize applied business or personal finance. Others are used to prepare for competency exams.

From the standpoint of interpretation, Dimension 4 (Figure 7) poses some difficult problems. The positive end of Dimension 4 is most prominently marked by marked by Resource General Mathematics (ResGM), General Mathematics Skills (GenMS), and Resource Consumer Math (ResCM). We call this the Functional Math Skills Profile. The remaining two peaks in the profile, Algebra 3 and Analytic Geometry, simply do not fit the profile.

Insert Figure 7 here.

The Dimension 4 Mirror Image Profile is also a bit messy. It is marked most prominently by Algebra/Trigonometry (AlgTrig). After that, it is marked by Algebra 1 Part 1, Algebra 1 Part 2, Informal Geometry, and Plane Geometry. The majority of these courses; Algebra 1 Part 1, Algebra 1 Part 2, and Informal Geometry; constitute a variation on the standard sequence in which Algebra 1 is spread over two years, and geometry is covered in a more applied, somewhat simplified fashion. We call this the Basic Formal Sequence.

The four dimensions define eight profiles, three of which appear to be college preparatory: the Standard Sequence of Algebra 1, Algebra 2, Geometry, Trigonometry, and Analytic Geometry; the Unified Mathematics Sequence in which algebra, geometry, logic, trigonometry and statistics are taught in a unified fashion, and the Calculus Profile. Three of the profiles appear to be non-college preparatory: the Pre-Formal Profile consisting of General Math 1, Pre-algebra, and Business Math 1; the Informal Sequences Profile consisting of General Math 1 and 2, Business Math 1 and 2, and Consumer Math, and the Functional Math Profile consisting of Resource General Math, General Math Skills, and Resource Consumer Math. The remaining two profiles are more difficult to classify on a college preparatory basis. The Alternative, Standard Profile includes courses which cover algebra and geometry, but at a slower pace or in a less formal fashion. The Beyond Basic Profile includes one course clearly designed for a college bound audience, Review Mathematics, as well as Computer Mathematics and Probability and Statistics

<u>Science</u>

Table 4 and Figures 8-12 show the five dimensional solution for the science data obtained from the ALSCAL program (Young & Lewyckyz, 1979). The fit measures for this solution were S-STRESS = .277, STRESS = .191, and R^2 = .482. Because the number of courses is somewhat large and the data were treated as ratio-scaled data, rather than ordinal, these fit measures are poorer than those obtained from most applications of MDS. A Varimax rotation (Kaiser, 1958) was applied to this solution to improve interpretability. As will become apparent below, the mirror image science profiles are poorly defined, a result which we attribute to the Varimax rotation. That is, the Varimax rotated science dimensions in Table 4 do not have the strong bipolar character of the unrotated mathematics dimensions in Table 3.

Insert Figures 8-12 here.

Figure 8 shows the profile plot of Dimension 1. The profile is most prominently marked by Biology Honors 1 (HonBio1), Biology Advanced (AdvBio), Chemistry 2 (Chem2), and Physics 2 (Phys2). It should be noted that Chemistry 2 and Physics 2 include Advanced Placement Chemistry and Physics. While this represents the classic high school sequence (biology, chemistry, and physics) it is the most advanced, rigorous form of the sequence. We refer to this as the Advanced Standard Sequence.

There are three marginally prominent (scale values slightly over 1.00) courses in the mirror image profile, Biology Basic (BasBio1), Science Unified (UniSci), and Earth Science (Esci). This combination contains a common ninth grade course plus two rather basic courses, one in biology and one covering both physical and life sciences. We do not refer to this combination further, and so we do not label it.

Figure 9 shows the Dimension 2 Profile, marked most prominently by Biology General 1, Chemistry 1, and Physics 1. We call this the Standard Sequence. It is the same sequence as Dimension 1, but at a less advanced level.

The mirror image profile is prominently marked by only one course, Other Informal Physical Science. To a lesser extent, it is marked by Other Formal Physical Science and Physical Science Applied. This appears to be a profile marked by physical science courses other than the standard ones.

The Dimension 3 profile, Figure 10, is marked most prominently by Chemistry Introductory, Earth Science College Preparatory, and Physics General. This would appear to be an alternative to the standard sequence which replaces biology with a college preparatory earth science, and which covers physics and chemistry at a more basic level than does the Standard Sequence. We call this the Alternative Sequence. The mirror image is marked only by Science Unified.

The Dimension 4 profile, Figure 11, is marked by Ecology and Marine Biology. This appears to be a combination of courses which would interest someone primarily interested in the Life Sciences and which is rare in the high school curriculum. The Dimension 4 Mirror Image profile seems to parallel, in some respects, the Dimension 1 Mirror Image. That is, it contains a biology course, General Biology, and a course covering physical and life sciences in an integrated, basic fashion, Unified Science. It also contains the Other Informal Physics category.

The last Dimension profile, Figure 12, is marked by Biology Basic, Physical Science Applied, and Earth Science. This appears to be a combination which is composed of Earth Science, commonly taught in the Grade 9, a life science course and one course covering both chemistry and physics. Both Physical Science Applied and Biology Basic appear to provide an applied, and possibly simplified presentation of the biology, chemistry, and physics taught in the standard sequence. We call this the Basic

In addition, there are some deviations from the common perception. First, Earth Science College Preparatory is an alternative to biology, chemistry and physics. Finally, a number of students appear to be taking courses more commonly associated with college curricula; e.g. human physiology, marine biology, and the courses lumped into our Other Formal Life Sciences and Other Formal Physical Sciences. We speculate that many of these are indeed college courses taken by students in districts (or states such as Minnesota) which allow high school students to take college courses, either on local college campuses or in the high school, and use them to meet high school requirements.

Summary and Conclusions

The methodological developments in this project are proceeding on two levels of generality. At the most general level, we are developing SAS® and SPSS® templates and macros to implement a Profile Analysis via Multidimensional Scaling (PAMS) with NAEP data. These templates are not limited to course-taking behavior. The analysis in these templates is based on the model of Equation 2. The model posits a small number of $\underline{K}$ "latent" profiles, called prototypical profiles, and it assumes that observed profiles can be represented as linear combinations of the prototypical profiles. On one level, the model is simple in that it posits a small number of prototypical profiles. But it is also powerful in that it can account for a very large number of different observed profiles as linear combinations of the prototypes.

Sequence. It should be noted that there appears to be an alternative to this Basic Sequence consisting of Biology Basic, Chemistry Introductory, and Physics General which appears in the unrotated solution (but not shown here). This alternative to the Basic Sequence seems to present science in a more applied, simplified manner but splits the physical sciences into two separate courses in physics and chemistry. The mirror image of this profile is prominently marked by Other Formal Life Sciences, Human Physiology, and Other Formal Physical Science. This appears to be a combination of courses, probably taken by the college oriented with more emphasis in the Life Sciences than the Physical Sciences. Some courses in this group are more commonly found in college curricula than high school curricula.

There is, we think, a common perception that high school science consists of biology, chemistry and physics. Our analyses support this perception but add some complexity to it. There does appear to be a Standard Sequence; Biology General 1, Chemistry 1, and Physics 1. There is also a more rigorous counterpart to that sequence composed of Biology Honors or Biology Advanced, Chemistry 2 (which includes Advanced Placement Chemistry) and Physics 2 (which includes Advanced Placement Physics). There also appears to be a more applied, simplified counterpart to the basic sequence composed of Biology Basic plus one course covering chemistry and physics, Physical Science Applied, or two courses, Chemistry Introductory and Physics General.

The model leads to a multidimensional scaling analysis of squared Euclidean distances among all possible pairs of measures. The dimensions can then be plotted as profiles, such as those in Figures 3 - 12. These profiles are interpreted in terms of prominent scores.

While the analyses reported here have proceeded no further than the MDS, the next step can be the estimation of subject correspondence weights, $\underline{w}_{sk}$, in Equation 2. The weight $\underline{w}_{sk}$ indexes the correspondence between the observed profile of subject $\underline{s}$ and prototypical profile $\underline{k}$. If $\underline{w}_{sk} = 0$, there is no linear correspondence between the observed profile of subject $\underline{s}$ and prototypical profile $\underline{k}$. If the weight is positive, then there is a positive linear association between the observed profile of subject $\underline{s}$ and prototypical profile $\underline{k}$. A negative weight signifies a negative association between the observed and latent profiles. One can use the correspondence weights to study a variety of issues. For instance, one can use the correspondence weights to study gender differences. If there is a gender difference in mean correspondence weights along dimension $\underline{k}$ favoring males, then on average, male profiles more closely resemble the prototypical profile than do female profiles. Further, one can study correlates of a profile. For instance, if there is a positive correlation between correspondence weights along dimension $\underline{k}$ and SES, then that suggests the observed profiles of upper SES subjects tend to resemble the prototypical profile more closely than do the observed profiles of lower SES subjects.

At this point, we have written SPSS[R] (Davison & Davenport, 1994) and SAS[R] (Davenport & Davison, 1995) templates to compute an appropriate proximity matrix, apply MDS to the proximity matrix to estimate the dimensions, and compute the subject correspondence weights. Our next step is to develop SPSS[R] and SAS[R] macros which will estimate the standard errors on the MDS scale values through resampling.

At a less general level, we have been adapting the PAMS approach for the study of course-taking patterns. Course-taking tends to be dichotomous: one has or has not taken a course. Linear models, such as that in Equation 2, tend to provide only rough approximations of dichotomous variables. Furthermore, the marginal in course-taking behavior can be very extreme. We have begun our adaptation of the PAMS approach by experimenting with proximity measures which are appropriate for dichotomous data and which intuitively seem suitable for variables with extreme marginal. This has led us to the Jaccard coefficient, $J_{ii'}$, or more accurately $(1 - J_{ii'})$ which expresses the number of people who have taken exactly one of the two courses (i, i') expressed as a proportion of those who took at least one of the two courses. The measure itself seems intuitively appealing for course-taking behavior and the MDS solutions based on it make sense, for the most part. If there were a response model, analogous to Equation 2, which led to the proximity measure $(1 - J_{ii'})$, just as Equation 2 leads to the squared Euclidean distance proximity measure in Equation 3, then our course-taking developments would rest on a much firmer theoretical foundation.

Not all of the developments in this project are or will be methodological. To date, our preliminary results are the MDS analyses reported in Tables 3 and 4 and the various figures. Major results from the mathematics analyses suggest three contrasting sequences. The first is composed first and foremost of Algebra 1, Algebra 2, and Geometry; these three core courses can be supplemented by Trigonometry and Analytic Geometry. To our surprise, Technical Mathematics emerged in this grouping, a course which seems to be taken with one or more of the basic three. A second sequence involves the same basic subject areas--algebra, geometry, and trigonometry--taught in a unified fashion. Logic and statistics can be included in this unified approach. Thirdly, there seems to be a somewhat simplified, slower paced sequence in which Algebra 1 is taught over two years and geometry is taught in a more simplified, slower paced, and/or applied approach.

There are also groups of courses which do not explicitly include algebra and geometry. For the seemingly college oriented, there are courses to review for entrance exams (Mathematics Review), courses in Computer Mathematics, and courses in Probability and Statistics. For students with special needs, there are courses in General Math Skills, Resource Consumer Math, and Resource General Math. And finally, there are groups of courses in Pre-Algebra, General Math, Basic Math, Consumer Math, and Business Math which can be terminal courses, preparation for competency exams, or preparation for algebra.

High school science is most commonly perceived as biology, chemistry, and physics. However, there appear to be three variations on this sequence; a more advanced sequence consisting of Advanced or Honors Biology, Chemistry 2, and Physics 2; a standard sequence consisting of General Biology 1, Chemistry 1, and Physics 1; and a basic sequence consisting of Basic Biology, plus a single course in Applied Physical Science or separate courses in Introductory Chemistry and General Physics. Furthermore, it would appear that some students have the opportunity to depart from these sequences with courses, such as Human Physiology or Marine Biology, or the many courses we have "lumped" together under Other Formal Physical Sciences, Other Formal Life Sciences, and Other Informal Physical Sciences. Some of these, such as Human Physiology and Marine Biology, are courses more commonly associated with college curricula, rather than high school curricula.

We were inconsistent in our collapsing of courses in math versus science, and on hindsight, our decisions in math were probably more sound. First, in math (but not science) we retained various resource and functional courses. Given the emphasis on inclusion of special needs students in educational data, we should have done the same in science. Second, in mathematics, we avoided using miscellaneous categories because they are vague and can lead to uninterpretable profiles. In science, we included three miscellaneous categories; Other Formal Physical Sciences, Other Formal Life Sciences, and Other Informal Physical Sciences. As feared, they led to some uninterpretable profiles, the Dimension 2 and 5 mirror image profiles, composed solely or primarily of miscellaneous courses. We need to redo our science analysis adding in resource and functional courses and eliminating the vague, miscellaneous categories.

Our next steps in this project are as follows. Methodologically, we are developing SPSS® and SAS® templates which use resampling to estimate standard errors on MDS scale values. Substantively, we need to redo our science solutions deleting miscellaneous categories and adding in the resource and functional categories. The stability of math and science MDS solutions across gender and ethnic groups needs investigation. Ultimately, we need to examine differences in NAEP achievement scores across different types of sequences, and we need to examine differences in participation rates by gender and ethnicity.

References

Bailey, K. D. (1994). Typologies and taxonomies: An introduction to classification techniques. Thousand Oaks, CA: Sage.

Davison, M. L. (1983). Multidimensional scaling. New York: Wiley. (Later New York: Krueger.)

Davison, M. L. (1994). Multidimensional scaling models of personality responding. In S. Strack & M. Lorr (Eds.), Differentiating normal and abnormal personality. New York: Springer Publishing.

Davison, M. L. & Davenport Jr., E. C. (1994). PAMS: SPSS[R] template. Unpublished software. Minneapolis, MN: Department of Educational Psychology, University of Minnesota.

Davison, M.L., Gasser, M., & Ding, S. (1995). Identifying major profile patterns in a population: An exploratory study of WAIS and GATB patterns. Unpublished Manuscript, Department of Educational Psychology, University of Minnesota, Minneapolis, MN 55455.

Davison, M.L. & Skay, C.L. (1991). Multidimensional scaling and factor models of test and item responding. Psychological Bulletin, 110, 551 - 556.

Davenport, Jr., E. C. (1992, April 21). The making of minority scientists and engineers. Invited adress presented to the American Educational Research Association. San Francisco, CA.

Davenport Jr., E.C. & Davison, M.L. (1995). PAMS: SAS[R] template.

Unpublished software. Minneapolis, MN: Department of Educational Psychology, University of Minnesota.

Gower, J.C. (1971). Statistical methods of comparing different multivariate analyses of the same data. In F. R. Hodson, D. G. Kendall, and P. Tatu (Eds.), Mathematics in the archaeological and historical sciences. Edinburgh: Edinburgh University.

Jones, L. V., Bekhuis, T., & Davenport, Jr., E.C. (1985). College-bound high school graduates of 1972 and 1980 compared by college major. Research in Higher Education, 23, 319-331.

Jones, L. V., Burton, N. W., & Davenport, Jr., E. C. (1984). Monitoring the mathematics achievement of black students. Journal for Research in Mathematics Education, 15, 154-164.

Kaiser, H. F. (1958). The varimax criterion for analytic rotation to simple structure in factor analysis. Psychometrika, 23, 187-200.

Legum, S., Caldwell, N., Goksel, H., Haynes, J., Hynson, C., Rust, K., & Blecher, N. (1993). The 1990 high school transcript study data file user's manual. Los Angeles, CA: Westat, Inc.

Lorr, M. (1994). Cluster analysis: Aims, methods, and problems. In S. Strack & M. Lorr (Eds.), Differentiating normal and abnormal personality. New York: Springer Publishing.

Noble, J. & McNabb, T. (1989, August). Differential coursework and grades in high school: Implications for performance on the ACT assessment. Iowa City, IA: ACT Research Reports Series.

Schonemann, P. H. & Carroll, R. M. (1970). Fitting one matrix to another under choice of a central dilation and a rigid motion. Psychometrika, 35, 245-255.

Tucker, L. R. (1960). Intra-individual and inter-individual multidimensionality. In H. Gulliksen and S. Messick (Eds.), Psychological scaling: Theory and applications. New York: Wiley.

Welch, W. W., Anderson, R. E. & Harris, L.. (1982). The effects of schooling on mathematics achievement. American Educational Research Journal, 19, 145-153.

Weller, S. C. & Romney, A. K. (1990). Metric scaling: Correspondence analysis. Thousand Oaks, CA: Sage.

Young, F. W. & Lewyckyz, R. (1979). ALSCAL 4 User's Guide(2nd ed.). Chapel Hill, NC: Data Analysis and Theory Associates.

Table 1. Mathematics Courses

Initial CSSC Digits	Course Title	Abbreviation
27	Algebra 1	Alg1
27	Geometry	Geo
27	Algebra 2	Alg2
27	Mathematics 1, General	M1Gen
27	Pre-Algebra	PreAlg
27	Analysis, Introductory	IntAnal
27	Consumer Mathematics	ConM
27	Basic Math 1	BasM
27	Trigonometry	Trig
27	Algebra and Trigonometry	AlgTrig
07	Business Mathematics	BusMa
27	Algebra 1, Part 1	Alg1P1
27	Mathematics, Unified	M1Unif
27	Calculus, Advanced Placement	CalAdv
27	Algebra 3	Alg3
27	Geometry, Informal	GeoInf
27	Mathematics 2, General	M2Gen
27	Geometry, Plane	GeoPl
27	Analytic Geometry	AnGeo
27	Algebra 1, Part 2	Alg1P2
27	Mathematics 2, Unified	M2Unif
27	Basic Math 2	BasM2
27	Mathematics 3, Unified	M3Unif
27	Calculus	Calc
27	Technical Mathematics	Tec.
56	Resource General Math	ResGM
11	Computer Mathematics	CompM
27	Mathematics Review	MRev
27	Probability and Statistics	ProbStat
54	General Math Skills	GenMS
56	Resource Consumer Math	ResCM

Note: 07 = Business and Office-related Mathematics, 11 = Computer and Information Sciences, 27 = Pure Mathematics, 54 = Functional Math Skills, and 56 = Subject Area Services

Table 2. Science Courses

Initial CSSC Digits	Course Title	Abbreviation
26	Other Formal Life Sciences	OFLSci
26	Biology, Basic 1	BasBiol
26	Biology, General 1	GenBiol
26	Biology, General 2	GenBio2
26	Biology, Honors 1	HonBiol
26	Biology, Advanced	AdvBio
26	Ecology	Ecol
26	Marine Biology	MarBio
26	Physiology, Human	HumPhys
30	Science, Unified	UnifSci
40	Other Formal Physical Sciences	OFPSci
40	Other Informal Physical Sciences	OIPhys
40	Physical Science	PhySci
40	Physical Science, Applied	AppPSci
40	Chemistry, Introductory	IntChem
40	Chemistry 1	Chem1
40	Chemistry 2	Chem2
40	Earth Science	ESci
40	Earth Science, College Preparatory	EScoCP
40	Physics, General	GemPhys
40	Physics, 1	Phys1
40	Physics, 2	Phys2

Note: 26 = Life Sciences, 30 = Unified Sciences, 40 = Physical Sciences.

Table 3. Multidimensional Scaling Solution for Mathematics Courses

Course	Dim 1	Dim 2	Dim 3	Dim 4
Alg1	1.6423	-0.1456	-0.7540	-0.2737
Geo	1.7248	0.3496	-0.5362	-0.2557
Alg2	1.7885	0.3161	-0.3435	-0.2513
M1Gen	0.0049	-1.3662	-1.3875	0.4357
PreAlg	0.3392	-1.5347	-1.1691	0.1974
IntAnal	0.7808	1.4431	-0.2771	-0.9734
ConM	-0.2283	-1.0459	-1.5805	0.5802
BasM	-0.8386	-0.8513	-1.5851	-0.0204
Trig	1.5578	0.7366	0.2271	0.9894
AlgTrig	0.0057	1.0524	-0.4276	-1.6211
BusM1	-0.3016	-1.4807	-0.3654	-1.2560
Alg1P1	-0.2377	-1.2087	0.7322	-1.5163
M1Unif	-1.3988	1.3168	-0.5517	-0.5601
CalAdv	-0.0786	1.8720	-0.3363	-0.5154
Alg3	1.4034	0.5120	0.0706	1.3696
GeoInf	-0.0041	-1.2253	0.6674	-1.4772
M2Gen	-0.8381	-0.5489	-1.5322	0.8994
GeoPl	0.6733	-0.7844	1.1898	-1.2553
AnGeo	1.5474	0.5317	0.5059	1.0185
Alg1P2	-0.1596	-1.1141	0.9818	-1.4184
M2Unif	-1.4590	1.2963	-0.3127	-0.5707
BasM2	-1.1875	-0.5154	-1.3704	0.7292
M3Unif	-1.4561	1.3158	-0.2362	-0.5349
Calc	-0.6819	1.8351	0.1667	0.0485
TechM	1.5223	0.7726	0.2496	0.7753
ResGM	-1.0697	-0.2491	0.5549	1.6468
CompM	-0.1543	-0.2354	1.9290	-0.4433
MRev	-0.8080	0.0594	1.8405	0.2625
ProbStat	-0.3525	-0.4659	1.7880	0.7615
GenMS	-0.8482	-0.2798	0.8872	1.6580
ResCM	-0.8879	-0.3579	0.9748	1.5710

Table 4. Multidimensional Scaling Solution for Science Courses

Course	Dim 1	Dim 2	Dim 3	Dim 4	Dim 5
OFLSci	-0.3714	-0.8662	-0.5782	-0.2361	-2.1339
BasBio1	-1.0971	-0.5394	-0.9531	0.3926	1.5936
GenBio1	-0.7130	1.8264	-0.1648	-0.2414	0.0283
GenBio2	-0.3965	-0.2486	-0.2290	-2.2611	0.2746
HonBio1	2.0932	0.4904	0.1804	0.0382	-0.2075
AdvBio	2.0918	-0.0173	-0.3252	0.0026	0.8150
Ecol	-0.3576	-0.8756	-0.6250	2.0806	0.5604
MarBio	-0.3436	-0.1544	0.1842	2.3580	-0.1945
HumPhys	-0.4804	0.5801	-0.3249	0.1738	-2.0053
UnifSci	-1.0329	-0.0681	-1.5460	-1.0380	-0.0165
OFPSci	-0.4114	-1.1728	-0.0907	0.8151	-1.8937
OIPhys	-0.3377	-1.9912	-0.0065	-1.4040	-0.0323
PhySci	-0.6538	1.8041	-0.4112	0.3485	0.1245
AppPSci	-0.6814	-1.3177	-0.7502	0.5460	1.5744
IntChem	-0.7929	-0.2766	2.0767	-0.1186	-0.0581
Chem1	0.0418	1.7628	-0.4366	-0.2817	0.0114
Chem2	2.1454	-0.2559	-0.4333	-0.2356	0.0656
ESci	-1.1138	1.1361	0.2984	-0.4714	1.1354
ESciCP	0.1513	-0.6069	2.2932	0.0325	0.1309
GenPhys	-0.4791	-0.3725	2.2770	-0.1254	0.2868
Phys1	0.5425	1.7246	-0.3431	-0.4096	-0.0331
Phys2	2.1968	-0.5614	-0.0923	0.0350	-0.0261

Figure 1. Factor and MDS representation of structure

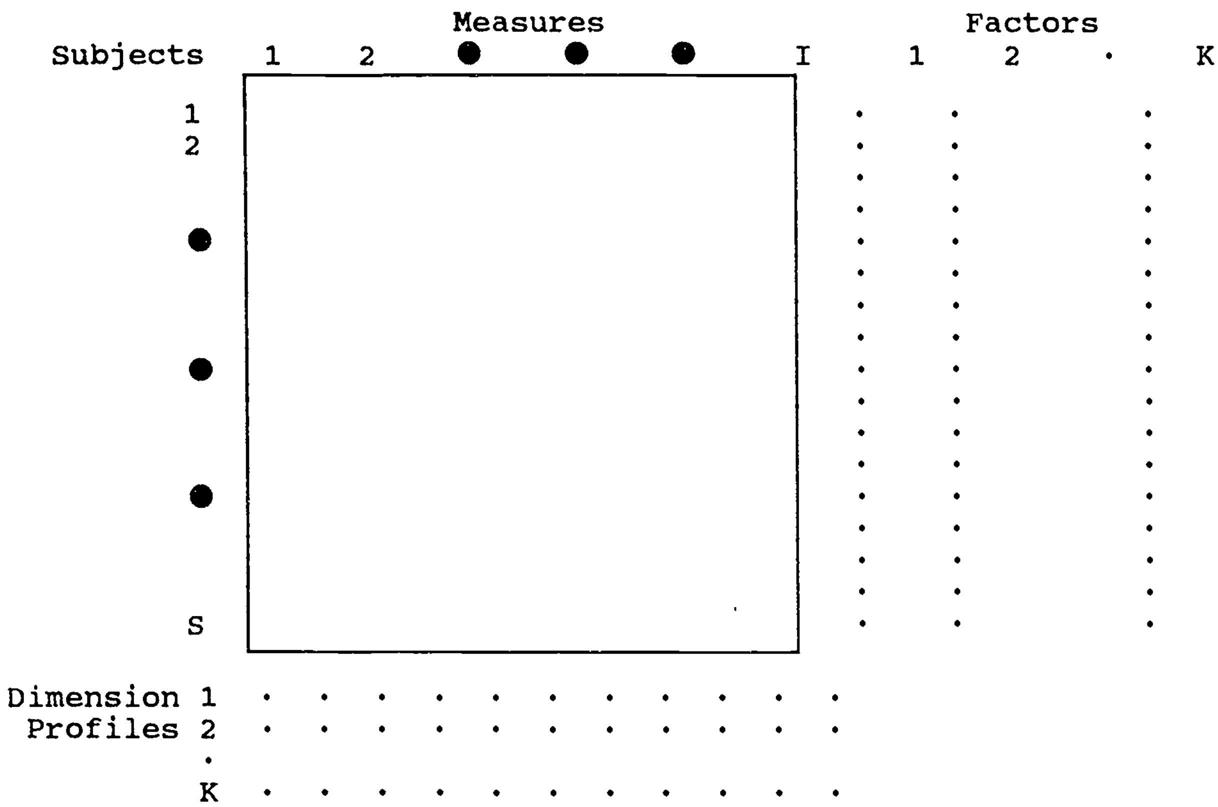


Figure 2. 2 X 2 Contingency Table of Course-taking Behavior

		Course i′	
		Taken	Not Taken
Course i	Taken	a	b
	Not Taken	c	d

Figure 3. Profile of Mathematics Dimension 1

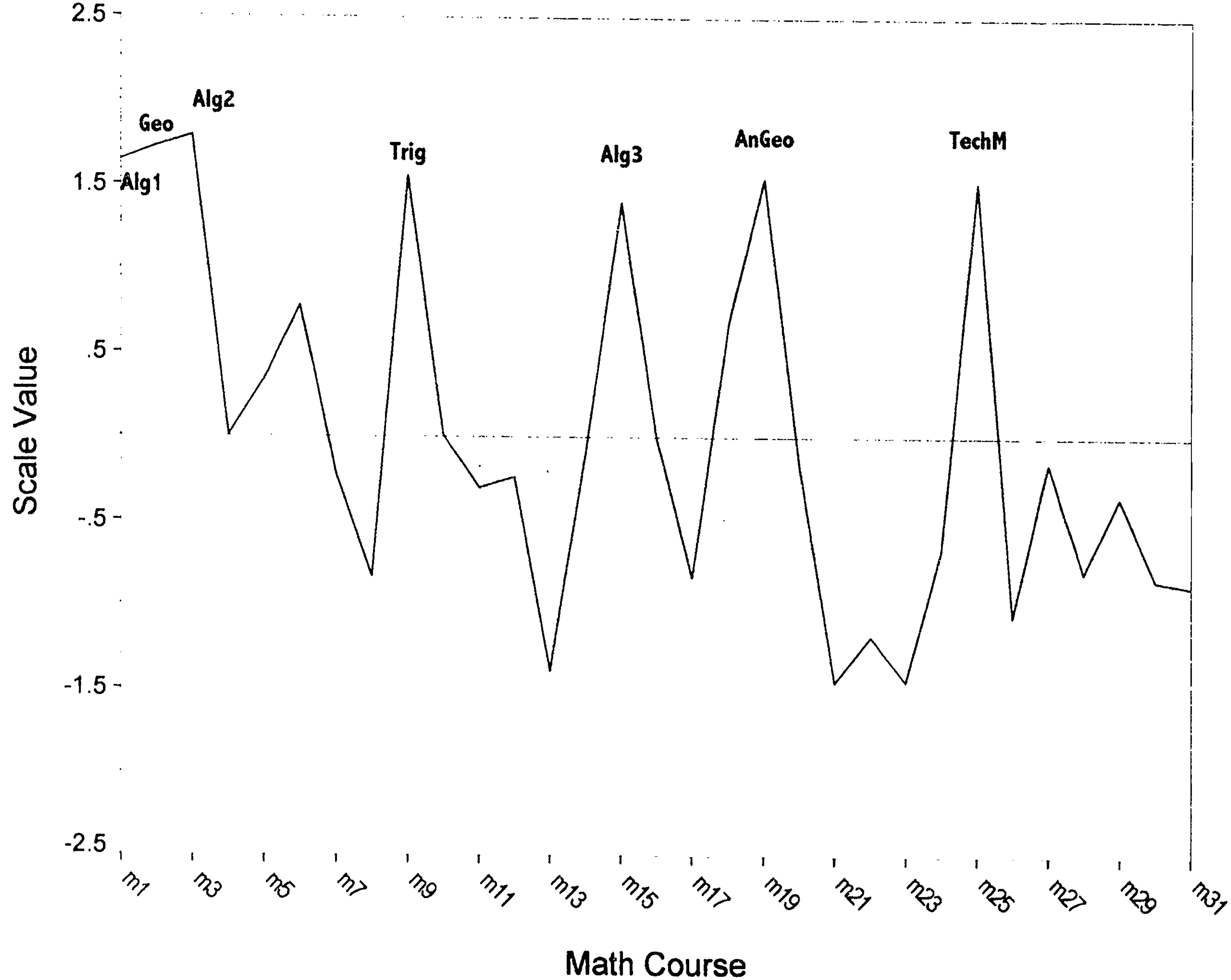


Figure 4. Mirror Image of Mathematics Dimension 1 Profile

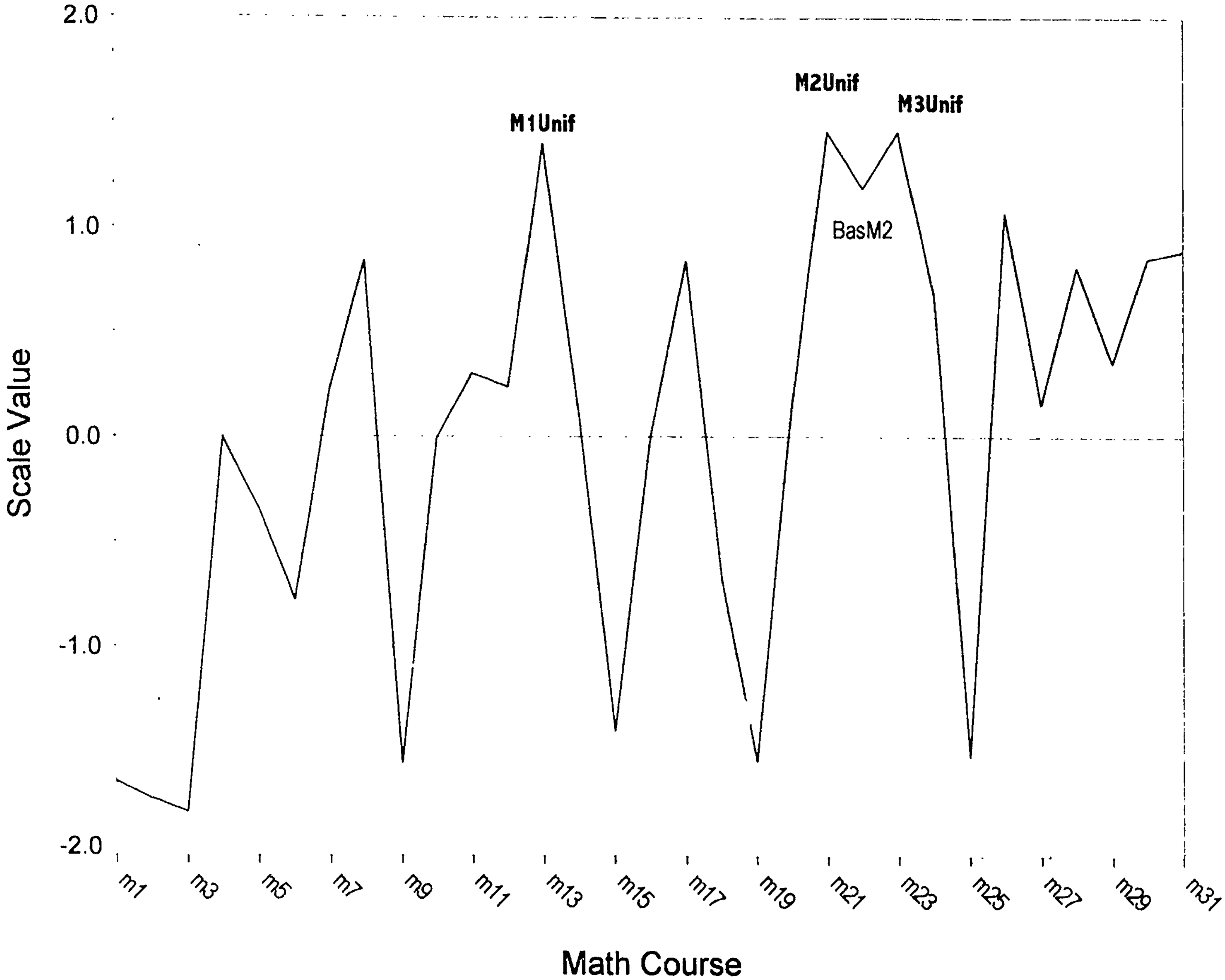


Figure 5. Profile of Mathematics Dimension 2

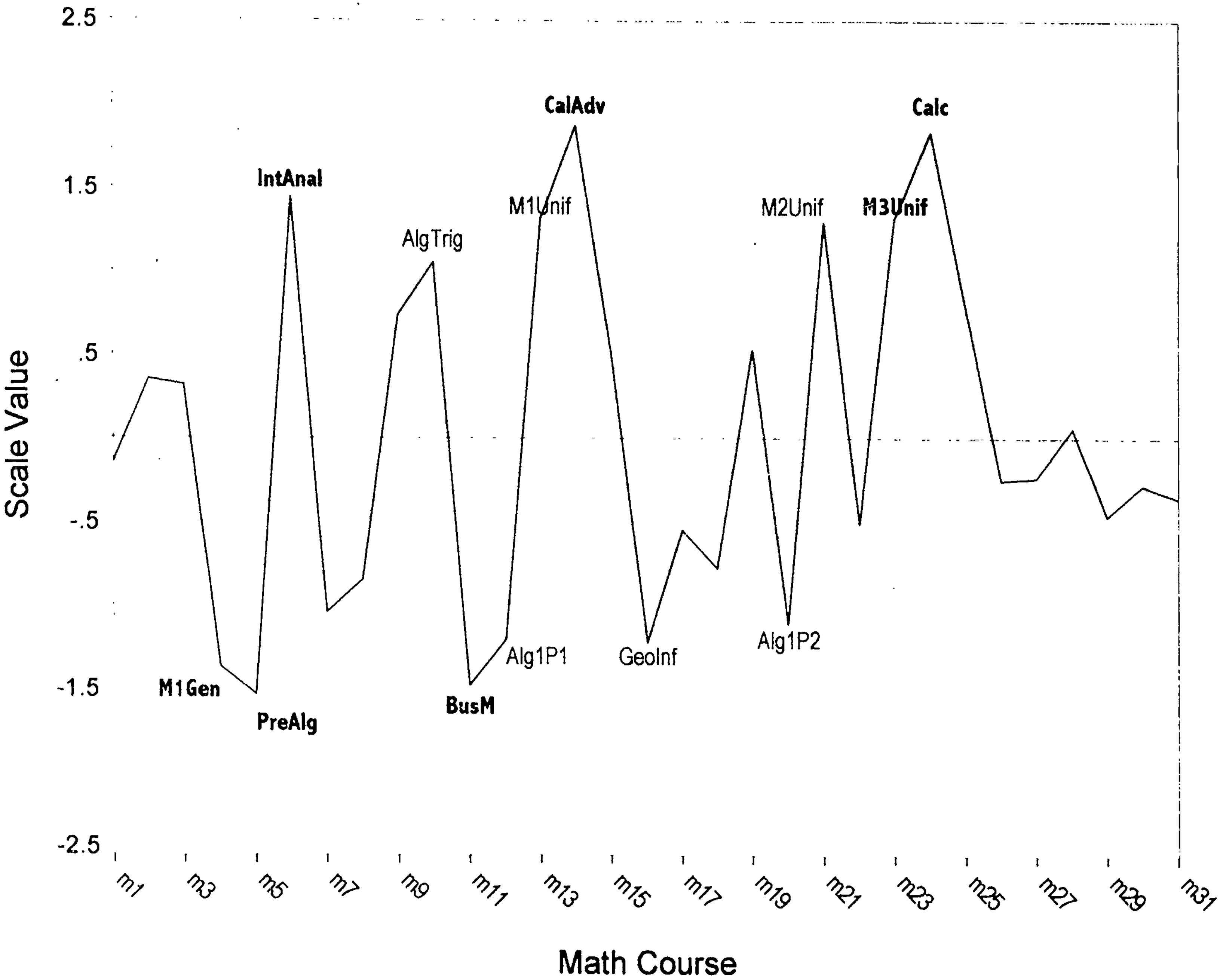


Figure 6. Profile of Mathematics Dimension 3

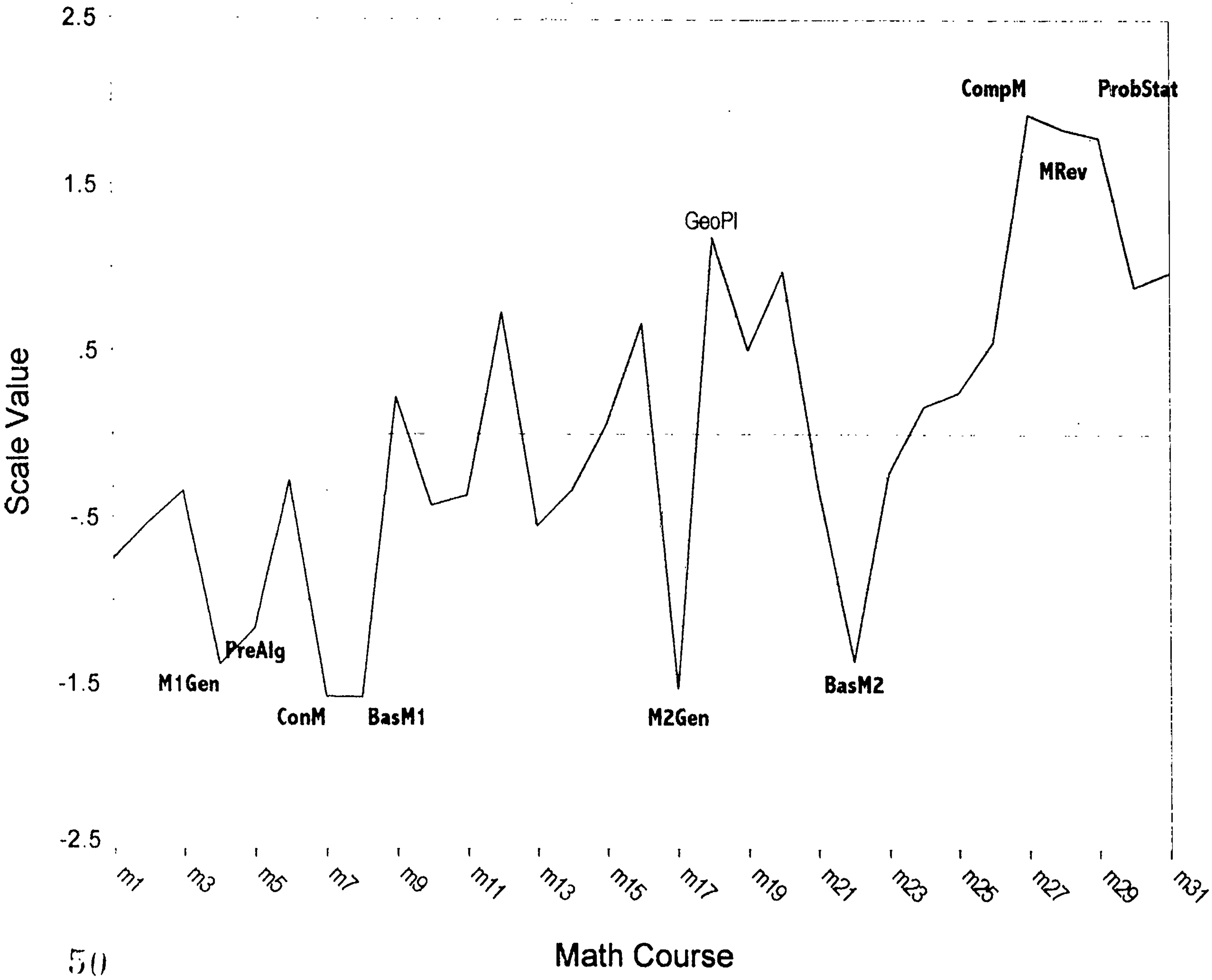


Figure 7. Profile of Mathematics Dimension 4

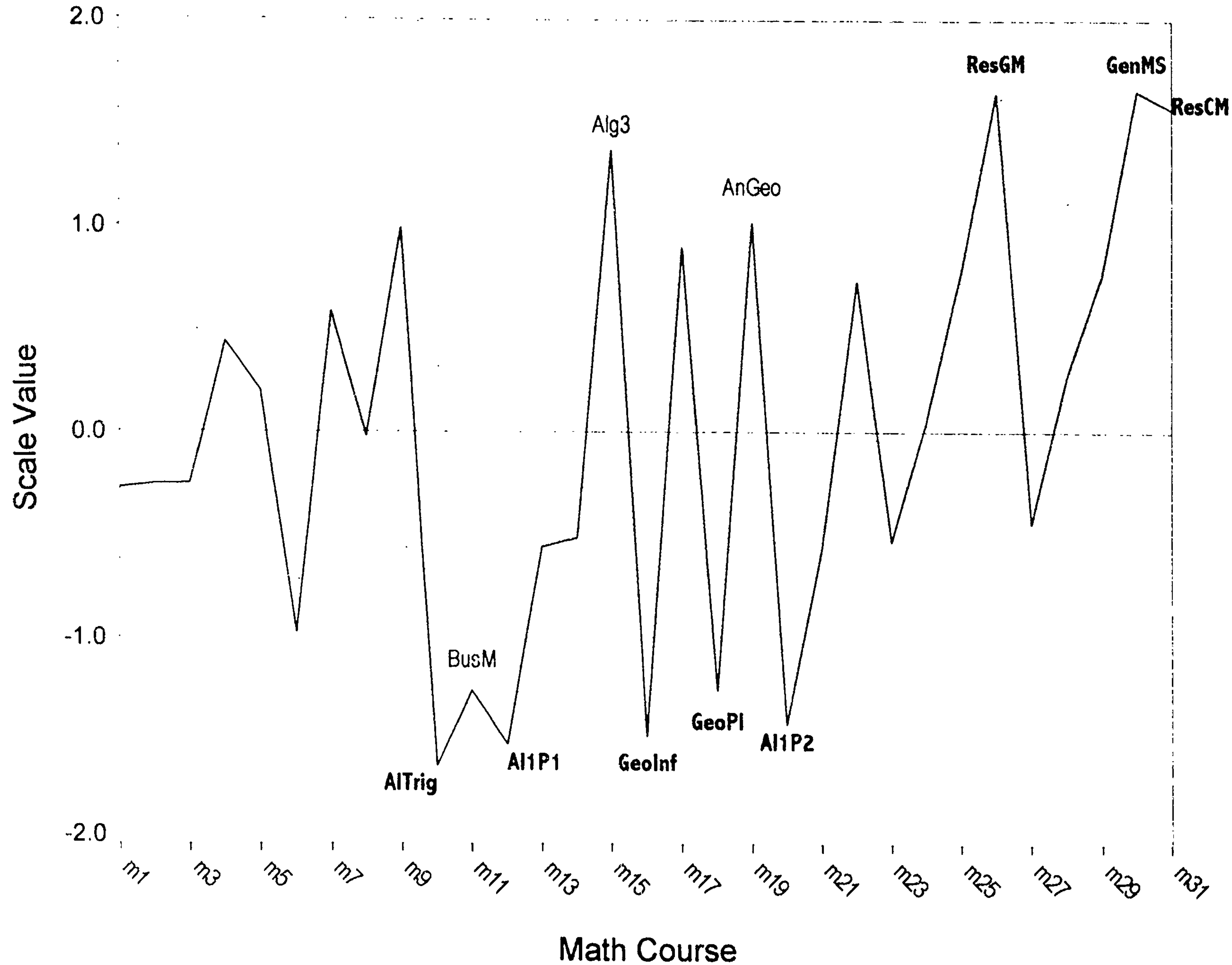


Figure 8. Profile of Science Dimension 1

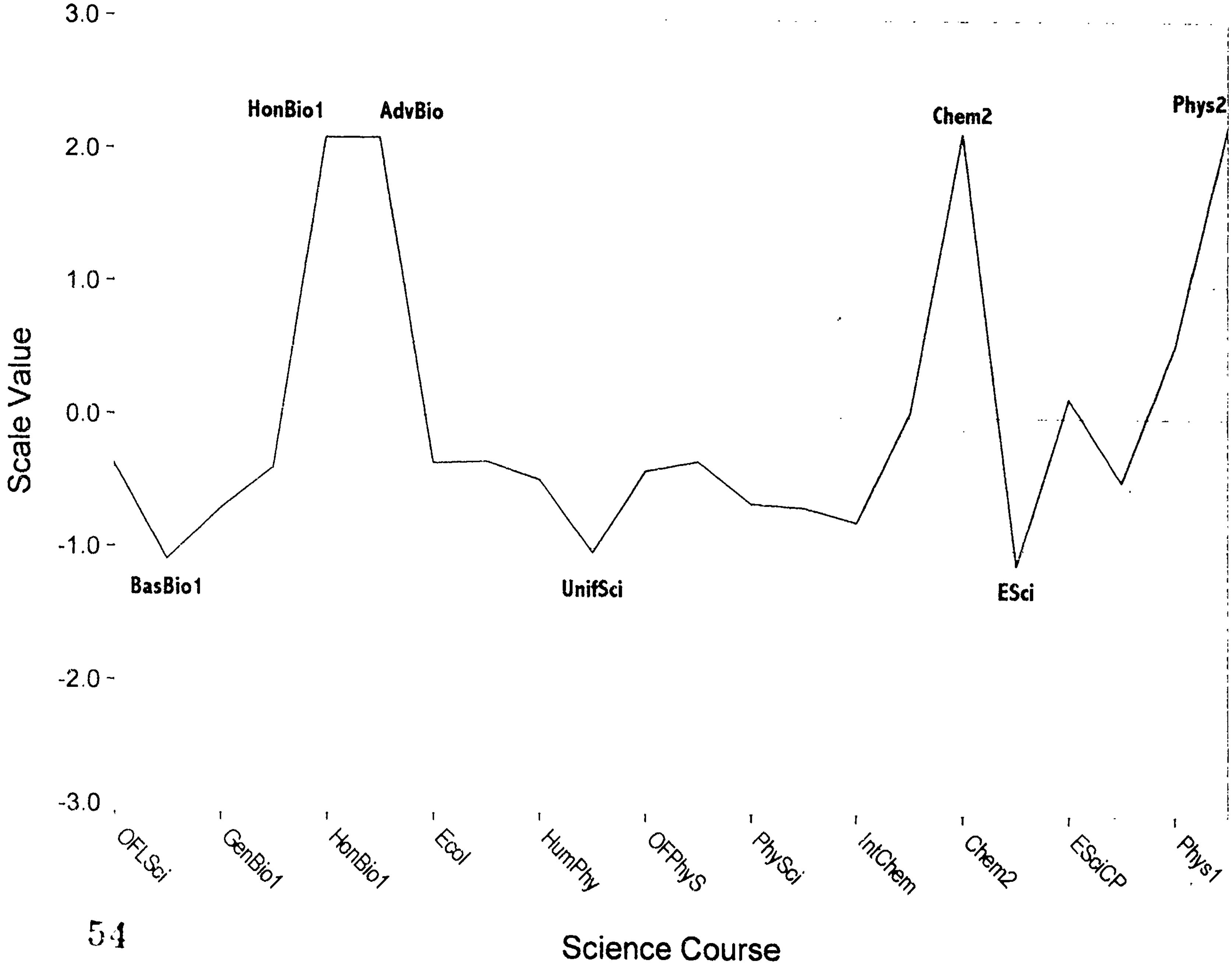


Figure 9. Profile of Science Dimension 2

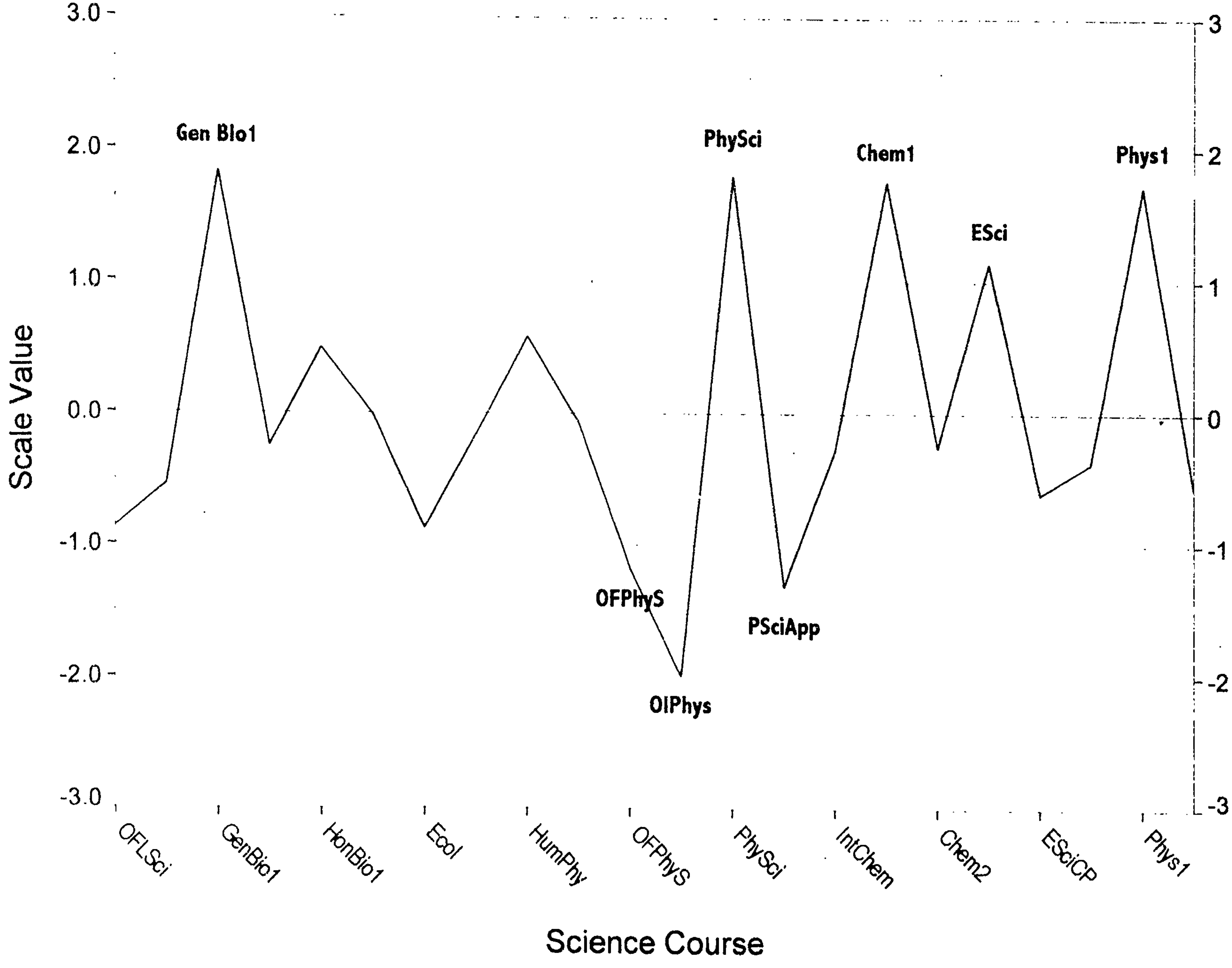
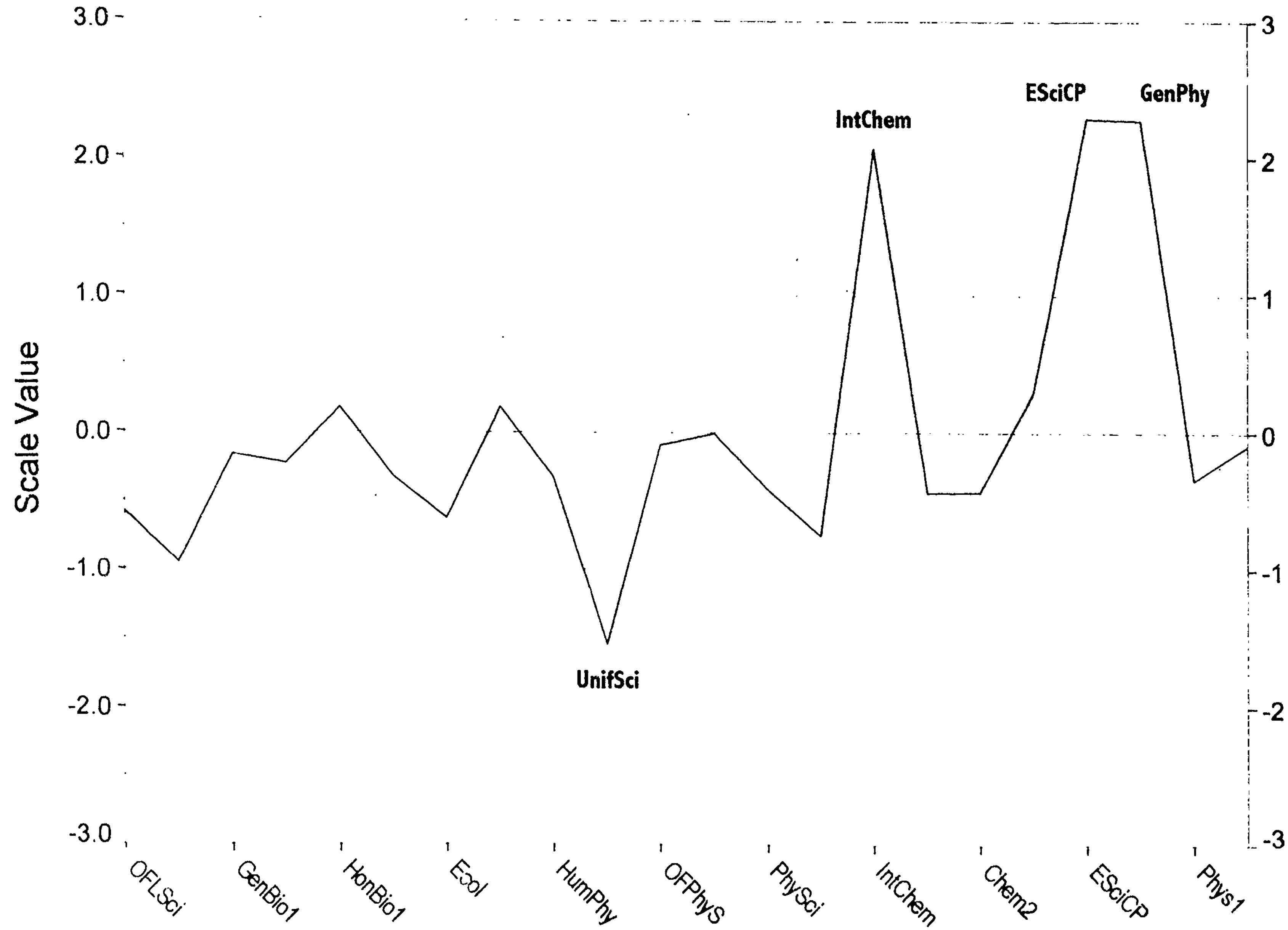


Figure 10. Profile of Science Dimension 3



58

Figure 11. Profile of Science Dimension 4

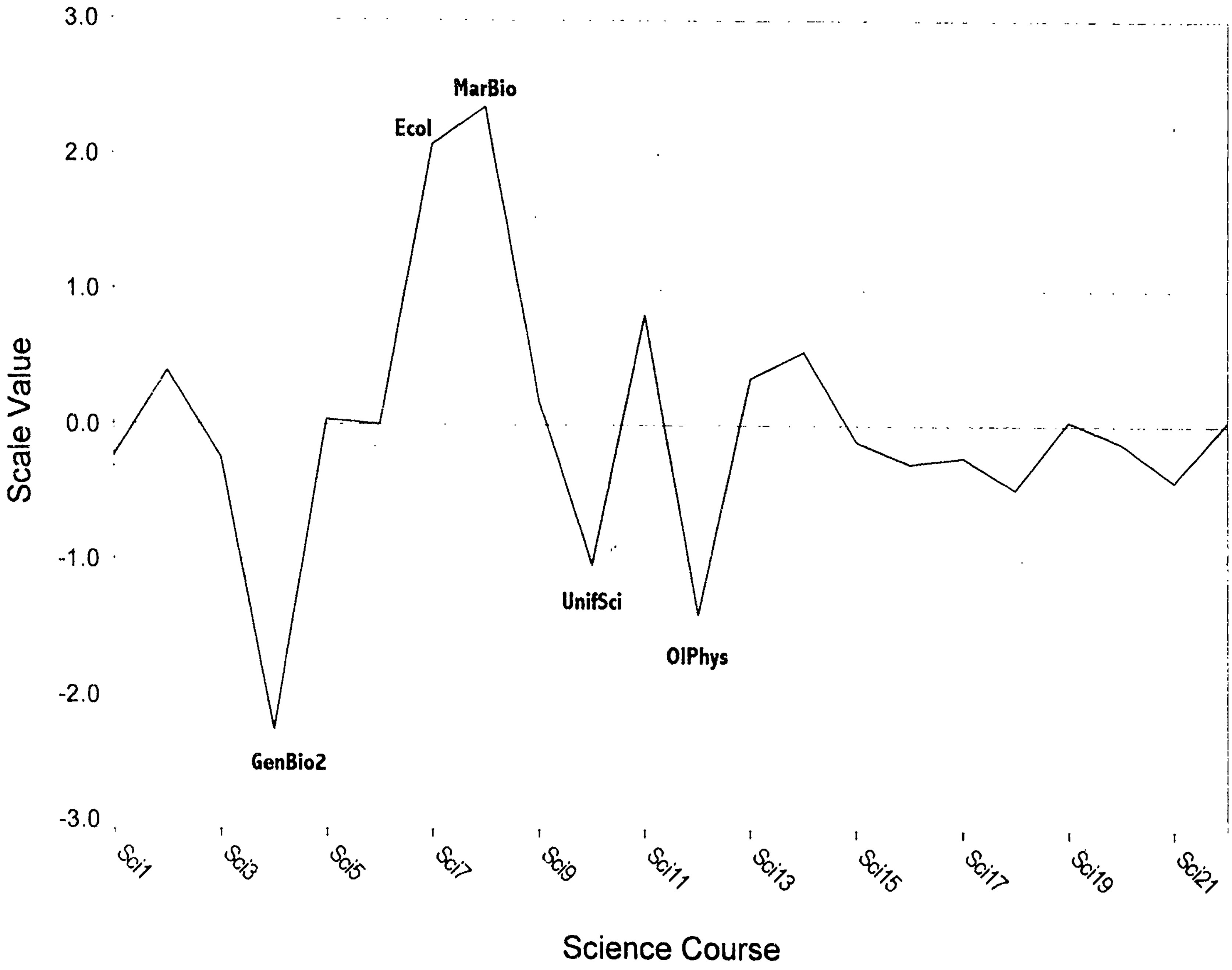


Figure 12. Profile of Science Dimension 5