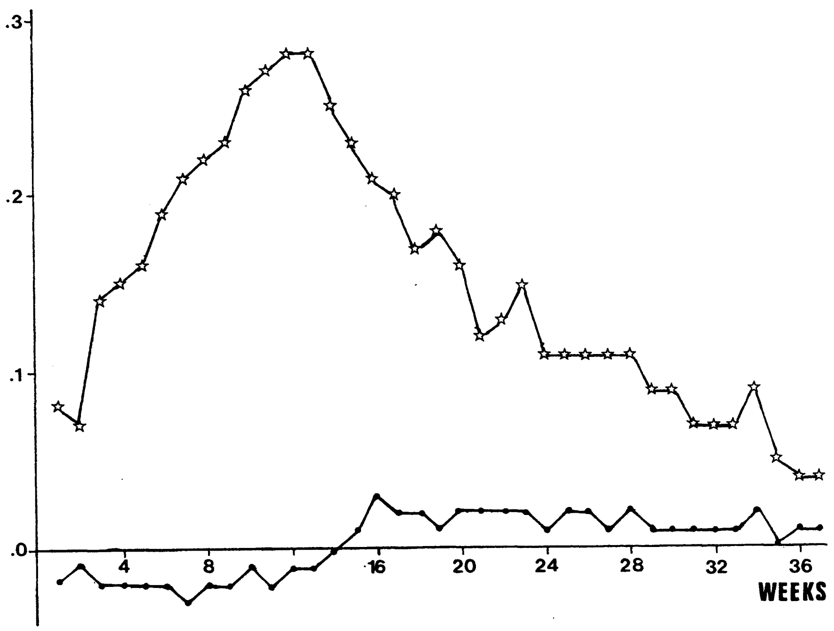
GROWTH CURVES

14

learning-to-read study



14.1 INTRODUCTION

In this chapter we will discuss the three-mode analysis of growth curves. Various proposals to deal with growth curves (reference, learning, or response curves) have been made. Either the exact form of the curve is of interest, or the differences in parameters to describe these curves in different sub-populations which are central to an investigation, or both. For instance, Snee, Acuff, & Gibson (1979) discuss a series of models to deal with univariate growth curves in several populations using Mandel's (1971) proposals. The essence of these proposals, as discussed in section 6.4, is that they consist of an additive main effect model with multiplicative interactions.

The growth curve proper can be analysed by postulating a functional model, and fitting it to the data. After fitting the curves, an analysis-of-variance procedure on the parameters can be used to assess differences between groups, a procedure already proposed by Wishart (1938). For the data to be described logistic regression models (i.e. regression models with a binary counted variable as response variable) with time as independent variable and *slope* and *intercept* as parameters can be fitted to the individual learning curves (Jansen & Bus, 1982); we will relate some of their results to the outcome of the three-mode analysis.

The analysis performed here can best be seen as a multivariate generalization of the growth (learning) curve approach tackled by Tucker (1958, 1966b and Weitzman, 1963). They performed a singular value decomposition (see section 2.2) on the observed raw data. Applications of this approach can, for instance, be found in Kouwer & Hartong (1961), Van Egeren, Headrick, & Hein (1972), McCall, Appelbaum, & Hogarty (1973), Hamel & Netelenbos (1976), Svendsrød & Ursin (1974). Van de Geer (1962) showed how orthogonal polynomials can be used instead of singular value decomposition, and Van Maanen-Feijen (1968) gives a theoretical and empirical comparison between the latter two approaches. That three-mode principal component analysis is a generalization of this approach to learning curves, follows from the observation that the technique is a generalization of singular value decomposition.

14.2 DATA AND PREPROCESSING

In a study to investigate the process of learning to read seven first-grade children were tested weekly (except for holidays) with five different tests (see Table 14.1 for a description), which were designed to measure different aspects of reading ability. We will not discuss the theoretical rationale behind the test the details of the design, the testing procedures, and the overall quality of the data (see Bus, 1982, Jansen & Bus, 1982, and Bus & Kroonenberg, 1982). Of the seven children which took the tests, one is not included in the present analysis, as he was added to the study at a later moment, and accounted for a large part of the missing data.

Table 14.1 *Learning to read study: Description of tests*

Test	Description
P	regular orthographic short words
Q	regular orthographic long words
R	irregular orthographic long and short words
S	regular orthographic long and short words within context
L	letter knowledge test

Because the tests had different ranges, either 10, 15 or 47 items, the data were rescaled so that all the tests ranged from 0 to 1. In this way all the differences in variation were maintained in the data, while making the tests comparable. Subsequently we constructed the average learning curve by averaging over pupils and tests for each occasion. Thus, in effect, we will use a mixed additive and multiplicative model for the data (see section 6.4 for a discussion of such models) analysing the common part of the learning curve additively, and the residuals, ε_{ijk}, multiplicatively:

$$z_{ijk} = \mu + \gamma_k + \varepsilon_{ijk} \qquad (i=1,\ldots,\ell;j=1,\ldots m;k=1,\ldots n)$$

$$= \mu + \gamma_k + \sum_{p=1}^{s} \sum_{q=1}^{t} \sum_{r=1}^{u} g_{ip}h_{jq}e_{kr}c_{pqr},$$

with μ the overall average, and γ the occasion main effect with the restriction $\Sigma\gamma_k = 0$. The least squares estimates for $\mu = \bar{z}_{...}$, and for $\gamma_k = \bar{z}_{..k} - \bar{z}_{...}$ $(k=1,\ldots,n)$ according to the standard theory of linear models. Thus the residuals $\varepsilon_{ijk} = z_{ijk} - \bar{z}_{..k}$ $(i=1,\ldots,\ell; j=1,\ldots,m;k=1,\ldots,n)$.

An advantage of the above model is that the (cor)relations between pupils and tests are no longer influenced by the average growth curve, and the interactions between tests and pupils over time can be analysed separately. Of course, averaging and interpreting the average growth curve is only meaningful if the individual curves more or less resemble one another (see e.g. Tucker, 1966b, p.480-483, and the references therein for a discussion of problems around average learning curves). Furthermore, averaging over tests is only meaningful if all tests measure essentially the same variable to a different extent. In the present case this is not unreasonable considering the high intercorrelations (average = 0. 88; taken over all time-pupil combinations).

A study by Jansen & Bus (1982) on the same data suggests that per test a logistic regression model with the same slope for all individuals is not an unreasonable model. The slopes are, however, different across tests. This implies that the average learning curve, $\mu + \gamma_k$ $(k=1,\ldots,n)$ does not necessarily represent any one test in particular. Nevertheless, it serves as a baseline for comparisons between tests and pupils.

Removing only the time main effect has the advantage that the difference in complexity between tests, and differences in reading ability between pupils remain in the analysis of the interactions. One could, of course, remove the pupil and test main effects as well, and analyse the remaining residual with three-mode principal component analysis, but this course is not pursued here.

14.3 AVERAGE LEARNING CURVE

In Fig. 14.1 the $\bar{z}_{..k}$ (k=1,...,n) have been plotted against k (time) to give a general impression of the shape of the learning curves. There is a rapid increase in the test scores until about week 16, and a gradual growth until the ceiling of one, i.e. all

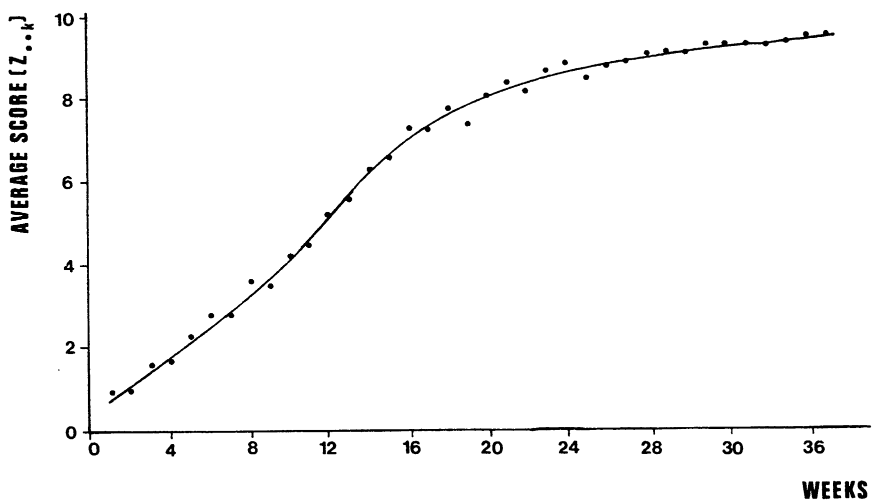


Fig. 14.1 *Learning-to-read study: average learning curve*

items of the tests correct. The actual process of learning-to-read thus roughly takes place in roughly those first 16 weeks; the later weeks are used to perfect the reading of the better pupils, and remedy that of the weaker ones.

14.4 GENERAL CHARACTERISTICS OF THE SOLUTION

The primary research questions in this study were centred around questions like: Was the performance of the pupils uniform over tests, did the tests cause uniform differentiation between pupils, or were certain pupils better on some tests, while others were better on other tests?

To answer these questions we will have to look at joint plots (see section 2.4, and 6.10) of the pupils and the tests. Before we can do this it is necessary to investigate how many dimensions are necessary for each of the three modes. At least two are needed for the pupils and the tests each (see Table 14.2). The table also shows that for occasions one component is sufficient, especially if one remembers that the most important source of variation (i.e. due to the average learning curve) has already been removed. Thus the

Table 14.2 *Learning-to-read study: proportions explained variation*

mode	proportion explained variation of components 1	2
1 pupils	.44	.33
2 tests	.43	.34
3 occasions	.70	.07

first time component already explains about 70% of the interactions. Fig. 14.2 shows the curves of the first and second components (scaled according to their relative weight). The second component seems to exhibit little (interesting) variation, and we will, therefore, not discuss it further.

The first component has a nearly perfect rank correlation with the variation per occasion measured by the fitted sum of squares. In other words, high loadings correspond to large differences between tests and/or pupils (the individual curves show that both is the case), and small loadings to small differences. From Fig. 14.2 we may conclude that in the first 12 weeks large differential growth exists between pupils on the tests and that the differences

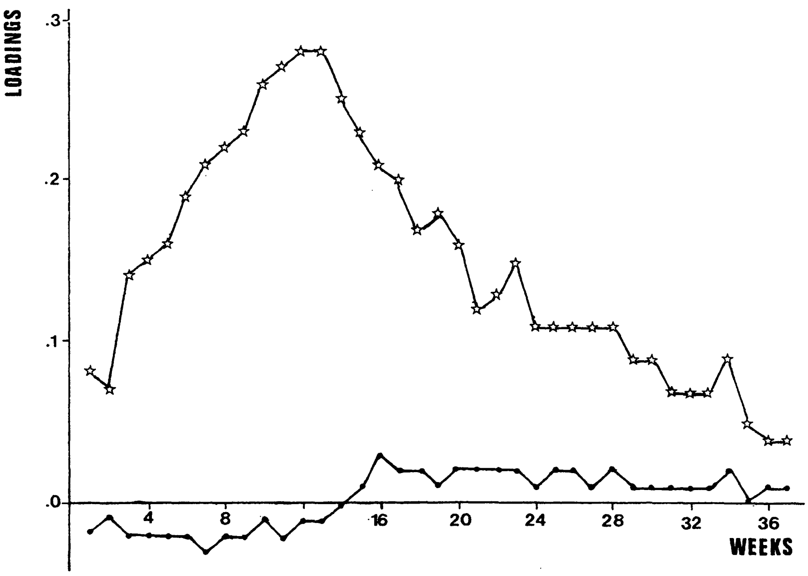


Fig. 14.2 *Learning-to-read study: Trends*
(☆ = first trend; • = second trend)

have largely disappeared at the end of the schoolyear. Inspection of the individual curves and observations in the classroom show that the decreasing differences are largely a ceiling effect, in other words, that further progress was impossible to measure with these tests as the better (quicker) pupils already obtained maximum scores on one or more tests.

14.5 ANALYSIS OF INTERACTIONS

With respect to the interactions between the tests and pupils, the foregoing shows that we only need to look at their interrelationships in the first core plane of a TUCKALS3 analysis (see Table 14.3), which shows the combinations of the components of tests and pupils responsible for the differential growth curves. The severe non-diagonality of this plane suggests that the relationships are rather complex. However, the near equality of the absolute values of the off-diagonal elements shows that the two sets of axes (of tests, and of pupils) are rotated with respect to one another over an angle of approximately 45° (see note of Table 14.3 and section 6.9).

Table 14.3 *Learning-to-read study: relationships between test and pupil components*

first frontal plane of Tuckers3 core matrix

		test components		proportion explained variation	
pupil	1	17.3	-11.4	.27	.12
components	2	11.2	14.7	.12	.70
				sum	.70

Note: direction cosine between pupil and test components $\cos \alpha = 11.2/ \sqrt{17.3 \times 14.7} = .70 \rightarrow \alpha \cong 45^{\circ}$.

In the joint plot corresponding with the first TUCKALS3 core plane (Fig. 14.3) the approximate positions of the original axes are drawn. How may we interpret Fig. 14.3 as far as the interactions of tests and pupils is concerned, or how do the combinations of components contribute to the differential growth? The interpretation is simplest when we take the component axes of the tests as reference. On the first axis we find a large differentiation between the tests, but not between pupils. In other words, the first test axis shows that all pupils have a positive differential growth (i.e. have curves above the average learning curve, see Fig. 14.1) for tests with positive loadings on this component (S,P,L), and a negative differential growth (i.e. curves below the average learning curve) for tests which have negative loadings there (R,Q). Thus, for all pupils with only minor deviations, S,P,L,Q,R are ordered from simple to difficult.

On the second test axis the situation is exactly reversed. There the tests have about equal loadings, but there are large differences between pupils. Pupil 4 is best on all tests followed

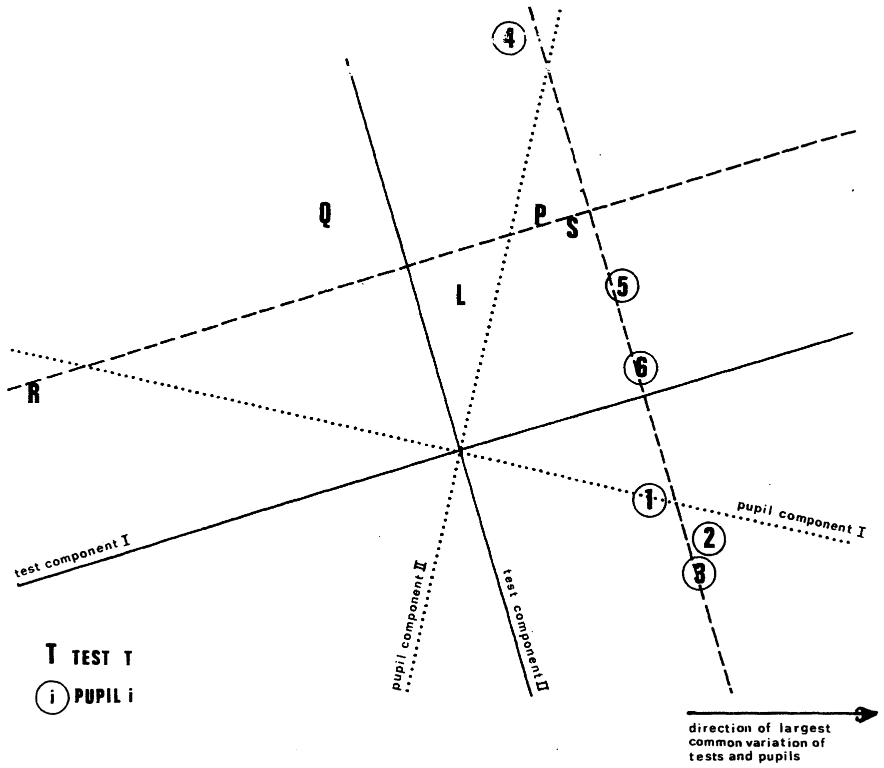


Fig. 14.3 *Learning-to-read study: Joint plot for tests and pupils*

by 5, 6, 1, 2 & 3. We see that 4 and 5 show positive differential growth on virtually all tests, while 6 lies more or less on the average curve for all tests, and 1, 2, and 3 show a negative differential growth on all tests. The slight variation in the loadings of the tests on this axis indicates that the spacing of the pupils is not exactly the same on all tests, but only approximately so. In particular, Q and L are somewhat different from the other tests.

In summary, there seems to be more or less independence of the differential growth due to differences in tests, and the differential growth due to differences in pupils. In other words, a test is easy or difficult for all pupils, and a pupil is better or worse than another pupil on all tests. This independence is nicely confirmed by Table 4 of Jansen & Bus (1982), which shows the estimated points on the time axis, $t_{ij}^{(\frac{1}{2})}$, at which a performance level of .50 (for the range 0-1) is reached. The estimates are based on logistic regression curves fitted to the scores of each pupil on each test (except for L, which was not included in their study). In Table 14.4 we have reproduced her table in rearranged form, as well as the

Table 14.4 *Learning-to-read study: estimated half-way scores* (in weeks)

	tests						residuals from two-way main effects model			
pupils	S	P	Q	R	mean	row effect	tests S	P	Q	R
4	4	4	2	15	6	-11	3	2	-2	-1
5	9	11	12	24	14	- 3	6	1	0	0
6	12	12	19	27	18	+ 1	-1	-2	3	-1
1	14	15	19	28	19	+ 2	0	1	2	-1
3	14	17	24	33	22	+ 5	-3	-1	4	1
2	18	17	23	34	23	+ 6	0	-2	2	1
mean	12	13	15	27	17					
column effect	-5	-4	-2	+10						

residuals $r_{ij}^{(\frac{1}{2})}$ after we have fitted a two-way main effect model to these half-way scores $t_{ij}^{(\frac{1}{2})}$:

$$r_{ij}^{(\frac{1}{2})} = t_{ij}^{(\frac{1}{2})} - t_{..}^{(\frac{1}{2})} - (t_{.j}^{(\frac{1}{2})} - t_{..}^{(\frac{1}{2})}) - (t_{i.}^{(\frac{1}{2})} - t_{..}^{(\frac{1}{2})}),$$

with $t_{ij}^{(\frac{1}{2})}$ the time at which pupil i scores .50 on test j; the "." indicates averaging over the index it replaces. The residuals do not seem to behave in a systematic way except that Q has maybe too many positive scores, confirming the slightly different behaviour of Q which we noted above. From the lack of pattern in the residuals we may therefore conclude that no interactions exist between pupils and tests with respect to the half-way scores.

14.6 CONCLUSION

In conclusion we may say that the difference in the scores of the pupils and the tests can be ascribed to two more or less independent factors: the varying degrees of difficulty of the tests, and the differences in ability of the pupils. Furthermore, the relationships between tests and pupils are time-independent, and the time aspect is contained in the loadings of the first time component. In other words, the structural relationships are invariant over time.

On a methodological level the example shows how three-mode principal component analysis can be used to analyse learning curves. It also gives a demonstration of the use of mixed additive and multiplicative models. Finally, the individual curve fitting as performed by Jansen & Bus (1982) and the three-mode analysis presented here, as well as a similar three-mode analysis by Bus & Kroonenberg (1982) nicely supplement each other. A more precise statement about the ways the two techniques can be used in conjunction requires another study.