An Investigation on the Reading Growth of Students with Learning Disabilities: With Special Regard to Simplex Model*

Dong-il Kim**

Department of Education, Seoul National University

Abstract: The purpose of the present study was to investigate the reading growth of students with severe reading difficulties. The subjects in the four-year longitudinal data set were school-identified students with learning disabilities (LD) from a single school district in a large midwestern city. Since a single indicator was collected at each measurement period, the simplex structure was applied. The identified model in the present study was a perfect simplex model. The results also showed that reading is very stable over time, based on relative standings. Even among students with severe reading difficulties, the scores of students with higher reading skills continued to be higher across grades. The identified simplex structure in reading growth also showed increasing within-grade-level variability with higher grades. Even among the students with severe reading difficulties, the reading levels become more heterogeneous across grade levels. Implication of educating students with learning disabilities were also discussed.

Keywords: simplex model, longitudinal data analysis, structural equation model, curriculum based measurement, reading

Ever since researchers began investigating academic growth, they have been interested in determining the extent to which differences in growth patterns are orderly and predictable from information that can be known about the individual. They also have examined the extent to which differences are under the control of ascertainable and describable environmental conditions, and the extent to which they are inherently erratic and unpredictable (Thorndike, 1966).

Much of the use of measurement in education is based on the premiseofstability. That is, a child who performs relatively higher than others at age 8 will continue to perform relatively higher at ages 10 or 12. Since this is stability with respect to relative standing,

* The previous draft of this paper was presented at several conferences. A special thanks to Dr. Victor Battistich, Dr. Daniel Solomon, and several anonymous reviewers for their comments on earlier versions.

** Correspondence about this paper should be addressed to Dong-il Kim, Ph.D. Dept. of Education, Seoul National University.
E-mail: dikimedu@snu.ac.kr; Tel: +82-2-880-7636; fax: +82-2-878-1665

it does not necessarily mean that a higher performing student has continued to grow rapidly. Bloom (1964) systematically developed the idea that much of the observed stability of intellectual status can be accounted for by a model that says ability at a later date consists of ability at an earlier date plus a growth increment. This increment is a change for either an individual or group.

There is abundant evidence that ability and achievement of an individual are highly stable over time. However, this stability is based on relative standing and includes little information about absolute changes in performance. Traditionally, intelligence and achievement testing are not based on absolute units (i.e., number or duration), but on the relative standing of an individual compared to his or her peers. While such information may be useful for classification and placement decisions (Salvia & Ysseldyke, 1991), it tends to be unrepresentative of changes in performance over time. Derived in part from single case research methodology, standardized performance measurement procedures have been developed by Deno and his colleagues, with a focus on "generalized outcome measurement" (Fuchs & Deno, 1991). When applied to a specific curriculum, these procedures have become known as Curriculum-Based Measurement (CBM) (Deno, 1985). This measurement procedure is an alternative approach to assess and monitor the growth in academic skills of students in reading, math, written expression, and spelling. When using this performance measurement to monitor student progress in reading, change in the number of words correctly read on repeated one minute samples of reading from a text of constant difficulty is used as a growth indicator of the student's general reading proficiency. The standard procedures of these oral reading measures have been shown to provide reliable and valid data indicative of student performance in reading (Shinn, Good, Knutson, Tilly, & Collins, 1992; Fuchs, Fuchs, & Maxwell, 1988). Additionally, conceptual models of reading development addressing the relationship between decoding and comprehension provide a theoretical basis for using these passage reading measures (LaBerge & Samules, 1974; Potter & Wamre, 1990).

The number of words a student is able to read in a one-minute period is used as an indicator of reading proficiency in the performance measurement procedures for formative evaluation system that teachers could use to evaluate student programs, proposed by Deno and his colleagues (Deno, 1985; Marston & Magnusson, 1988). Deno's performance measurement procedure is an empirically derived, standardized form of reading assessment, partly based on the research studies at the University of Minnesota Institute for Research on Learning Disabilities (IRLD) from 1977 to 1983. The work of the Minnesota IRLD was rooted in a model for developing individual special education programs called Data-Based Program Modification (Deno & Mirkin, 1977). This performance measurement procedure also was influenced by "the observational and analytical methodology of applied behavior analysis" and "the techniques and methods of Precision Teaching" (Deno, 1991, p. 10). The last major influence was the application of conventional test theory to the development of performance measures. The development of the performance measures used in the formative evaluation model focused on reading, spelling, and written expression. Since the present study has dealt with reading, that will be the focus of the following discussion.

Deno's approach is not to measure specific skills, but to provide an index of general reading proficiency. Through a review of the literature on reading, several indicators of reading fluency were selected, including supplying words deleted from text (cloze), saying the meaning of words underlined in text, reading aloud from isolated word lists, and reading aloud from text passages (Deno, Mirkin, & Chiang, 1982). Then, a series of studies was conducted on various measurement parameters (format, duration, source of stimuli). Based on the findings from these various investigations of the criterion-related validity of alternative formats of performance measurement on reading, the developers concluded that counting the number of words read aloud correctly and incorrectly in 1 minute form either isolated word lists or text passages produced reliable and valid data on a student's reading proficiency.

Correlations between the oral reading samples and various published reading criterion measures (e.g., the Literal and Inferential subtests of the Stanford Achievement Test, the Woodcock Reading Test, and the Reading Comprehension subtest from the Peabody Individual Achievement Text), including decoding and comprehension, ranged from .73 to .91 with most coefficients above .80 (Deno, Mirkin, & Chiang, 1982). Internal consistency, test-retest, and interscorer reliability estimates ranged from .89 to .99 (Fuchs, Fuchs, Maxwell, 1988; Marston, 1982; Tindal, Marston, & Deno, 1983).

Assessing reading comprehension was a major consideration in developing performance reading measures from the initial stages of test development, Thus, the criteria used in the validity studies included comprehension tests. The findings indicate the validity of the reading performance measure is high with respect to comprehension (Fuchs, Fuchs, Maxwell, 1988; Shinn et al., 1992).

These passage reading measures provide the teacher with an efficientrecording, using a constant metric over time, of the changes in a child's academic skills (Deno, 1991). For the teacher, regular monitoring of changes in a constant metric provides a data base for making inferences regarding the impact of various classroom interventions. For the researcher, counting repeated occurrences of a behavior in a fixed sample provides a constant metric that can be used for both between and within subject comparisons (Fuchs & Deno, 1991).

Previous research (Deno, Marston, Shinn, & Tindal, 1983; Marston & Magnusson, 1988) using the performance measure of passage reading showed that, across grade levels, the level of academic performance of regular education students was quite different from that of students who were referred to a special education program for reading (school-identified learning disabilities). Learning disabilities (LD) is a term that refers to a category of students whose under achievement is not accounted for by other handicapping conditions. Reading is the most common problem in instruction (Carlisle, 1990). The concept of LD has been controversial from its inception and the rapid proliferation of special education programs for students classified as learning disabled has resulted in widespread concern for the validity and significance of the construct of LD (Keogh, 1987). In the normative development studies using the performance measure of passage reading (Deno, Marston, Shinn, & Tindal, 1983; Marston & Magnusson, 1988), the reading growth from grade 1 to grade 6 of students with learning disabilities as a group was flatter than the reading growth of children in regular education or in Chapter I. Oral reading fluency was sensitive enough to differentiate students identified as LD from other low achieving peers as a group (Deno, Marston, Shinn, & Tindal, 1983). Standardized performance measurement procedures using passage reading have been adopted in five decision areas in the assessment of learning disabilities, including screening, identification, program planning, progress monitoring, and program evaluation (Marston & Magnusson, 1988). However, because these data come from cross-sectional studies, they may not accurately represent the actual patterns in reading growth.

The two conventional designs used for the examination of an age-functional relationship are generally known as the longitudinal and cross-sectional methods. The longitudinal method follows the same persons through all age levels with repeated observations. The cross-sectional method compares different age groups observed at one point in time (Baltes, 1968; Baltes & Reinert, 1969). When it comes to identification of intra-individual change, the cross-sectional design is only an imperfect approximation to the use of a longitudinal design. Although it is frequently used, the cross-sectional design is rarely a defensible "short-cut". Under limiting conditions, the cross-sectional method can be used to examine an "average" growth function, which is, however, open to empirical examination.

The tradition of describing individual growth (intra-individual level) as a function of time using a longitudinal design dates back at least to the eighteenth century (Willett, 1985). In both the biological and academic growth literature, a pair of competing yet complementary strategies for selecting growth models have been proposed. These are the rational and empirical strategies (Guire & Kowalski, 1979). When little is known of the mechanism governing the growth process, the empirical approach must be applied, and the researcher selects a mathematical function that simply fits the growth record adequately. In many descriptive and empirical approaches, the class of polynomial regression is apt to be satisfactory and have convenient properties (Guire & Kowalski, 1979).

A higher-order polynomial is always better-or-equal in fit to a lower order model. Thus, a trade-off between goodness of fit and the order of the polynomial is involved in deciding on a particular member of the class of polynomial functions as an appropriate growth model. In general, the higher-order models beyond the cubic function may be difficult to interpret in terms of their biological and academic significance (Willett, 1985). Among the lower-order polynomials, the straight-line (first-order) is often the growth model of first choice. If observed growth shows evidence of curvature, then a non-linear model (quadratic and/or cubic models) is required. Using statistical and practical criteria, a parsimonious and well-fitting representation of growth is chosen.

On the other hand, a rational alternative to the empirical selection of a suitable growth model requires the growth function to mirror accurately the mechanism driving the growth process, so that the parameters of the model have an interpretation in terms of internal processes responsible for the observed relationships (Willett, 1985). As with the empirical approach, the intention is to select a parsimonious and well-fitting representation, the statistical estimation of which acts to smooth and summarize the individual time course. In this way, the parameter estimates can indicate crucial features of the data.

Among the rational approaches, when the same variable is measured repeatedly on the same people over several occasions, a multiwave-one-variable growth model generated by a first-order autoregressive process can be used (Jöreskog, 1979). This particular model is referred to as a simplex model (Guttman, 1954). The typical feature of a simplex correlation structure is that the entires in the correlation matrix decrease monotonically as one moves away from the main diagonal. The simplex model appears to be particularly appropriate for studies of academic growth (Jöreskog, 1979; Jöreskog & Sörbom, 1989; Lunneborg & Lunneborg, 1970).

Several statistical models have been proposed for such simplex structures. Jöreskog (1970) discussed several different models. First, depending upon the reliability of the measures, a distinction is made between a perfect simplex and a quasi-simplex. A perfect simplex model is reasonable only if measurement errors in test scores are negligible. A quasi-simplex model, on the other hand, allows for sizable errors of measurement. Second, simplex models can be formulated through Wiener and Markov stochastic processes. The Wiener simplex is a scale-dependent model and is appropriate only when the unit of measurement is the same for all tests. The Markov simplex model, on other hand, is scale-free model. Because of its flexibility and fewer restrictions (scale-free), the Markov model is frequently used in time-series models (Collins, 1991; Frederiksen & Rotondo, 1979).

The term Markov refers to an interesting property of the probability structures of these models that results in "memoryless" behavior of the system as it moves through time. A model has the Markov property if past states have no influence on the probability of any future states of the system, given the present state of system. Therefore, Markov models have no "memory" of the states the system has already passed through (the past history of the system). Only the present state of the system has any effect on the probability of future states.

These different models are special cases of a very general model considered by Jöreskog & Sörbom (1989), and can be estimated by analysis of covariance structures. In this paper, the goodness of fit of the different models to observed reading performance data is tested, and the parameters of the most appropriate model are estimated. Particularly, in the present investigation using the CBM reading procedure, the research questions regarding the developmental nature of reading are as follows:

1. Is a simplex structure appropriate for the reading growth model of students with learning disabilities?
2. Does the CBM reading procedure show the same reliability across years? Furthermore, is it plausible thattheprocedure is reliable enough to be assumed to have no measurement error?
3. Are the absolute increments in reading the same across grade levels?

The above research questions can be answered by investigating the several alternatives and finding out the most parsimonious model. From the least constrained to the most constrained, the proposed models are:

Model 1. A quasi-simplex model: The correlation among the true measures has the simplex property (the correlation's decrease as one moves away from the main diagonal) and sizable measurement errors are allowed.
Model 2. A quasi-simplex model with equal reliabilities: The reliabilities of the measures across occasions are always same.
Model 3. A perfect simplex model: The measurement errors are negligible in the simplex model.
Model 4. A perfect simplex model with equal regression weights: Annual growth rates (actual increase in reading scores) are constant across occasions.

Ⅰ. Method

A. Subjects and setting

The subjects in this four-year longitudinal data set were 65 school-identified students with learning disabilities (LD) from a single school district in a large midwestern city. The distribution breakdown included 29 boys/15 girls (by gender) and 20 African Americans, 11 Caucasians, 4 Native Americans, 1 Asian, and 1 Hispanic (by ethnicity), based on available demographic information. All participants were reported as students with Reading Individual Educational Plans (Reading IEP) across public schools in the district. This longitudinal data set was a subset of the cross-sectional data sets from 1986 to 1989. That is, all of the students with Reading IEP in this district completed the standardized reading task as part of district-wide program evaluation process. In the cross-sectional data sets, there were around 200 students with Reading IEP at each grade level. The average growth patterns among the cross-sectional data sets and the longitudinal data set were similar (Kim, 1993).

The procedures for identifying students for special education service as mildly disabled in this school district have been described in detail elsewhere (Marston & Magnusson, 1988). In general, students who are referred are extensively screened using Curriculum-Based Measurement (CBM) procedures. The students placed in special education are those who function two grade levels below their age/grade placement on the curriculum-based reading measures, and who have been determined by their IEP team to have special educational needs. The mean level and standard deviations of the standard CBM reading measure for the students participating in this study are presented in Table 1. Along with that, the grade norms for the same measure (Doss & Deno, 1987) were reported to show a substantial difference in mean levels between the general education students and students with learning disabilities.

Table 1.
Means and Standard Deviations of the Standardized Reading Measure (Numbers of Words Read Correctly)

		Grade			
		2	3	4	5
LD Cohort	M	14.6	34.6	64.5	80.5
	SD	8.3	20.3	30.0	27.0
	n	61	61	65	45
Grade Norm[a]	M	107.1	124.4	145.1	161.4
	SD	47.8	45.9	41.6	30.5

B. Materials

Testing material was a third-grade passage derived from the students' basal reading series, Ginn 720, Rainbow Edition (Clymer & Fenn, 1979) and yields the following descriptive information:
Length: 303 syllables/248 words/21 sentences
Word size: 52 of six or more letters/3 of three or more syllables/
average letters per word=4/average syllables per word=1.2
Readability: Fry=Grade 3

C. Administration and Scoring

Data were collected in the spring (first week of May) of each year by special education resource teachers (SERT). The standardized directions and scoring procedures proposed by Deno, Mirkin, and Wessen (1984) were employed. Words were counted as correct if they were pronounced correctly. An error was scored when a word was mispronounced, omitted, or substituted.Words pronounced incorrectly but self-corrected within 3 seconds were scored as correct. Repetitions and insertions were not counted as errors. When students took longer than 3 seconds to read a word, SERTs supplied the correct word and counted an error. A teacher and a student copy of the passage were provided to SERTs. The number of words was totaled cumulatively to the right of each line on the teacher copy; student copies were unnumbered. The number of words read correctly from the passage was served as the dependent measure for this study.
For the four-year period, 37 students in the cohort had four data points; another 28 students had three data points with one missing.

D. Training

The SERTs were trained in administration and scoring of the CBM measure in 2 half-day workshops before the study began. The training was designed to educate SERTs on the philosophy and research behind CBM, and includes extensive experience in administration and scoring. Intra-observer agreement index estimate exceeded .90 (see Marston & Magnusson, 1988).

E. Data Analysis

A path diagram for the simplex model of the present four-year longitudinal study is presented in Figure 1. The observed scores (y_i) are assumed to be related to their corresponding true scores (η_i) by the following equation:

$$y_i = \eta_i + \varepsilon_i \qquad i=1,2,3,4 \text{ (years 1 to 4)}$$

(1) where all ε_i are independent of each other and all true scores. The simplex structure among the true scores can be stated as

$$\eta_i = \beta_i \eta_{i-1} + \zeta_i \qquad i=2,3,4$$

(2) where all ζ_i are independent and β_i is the true regression weight. This equation (2) implies that the partial correlation between η_{i-2} and η_i is zero with η_{i-1} controlled. That is, the present state (η_i) depends only on the state of the system just one interval earlier (η_{i-1}).

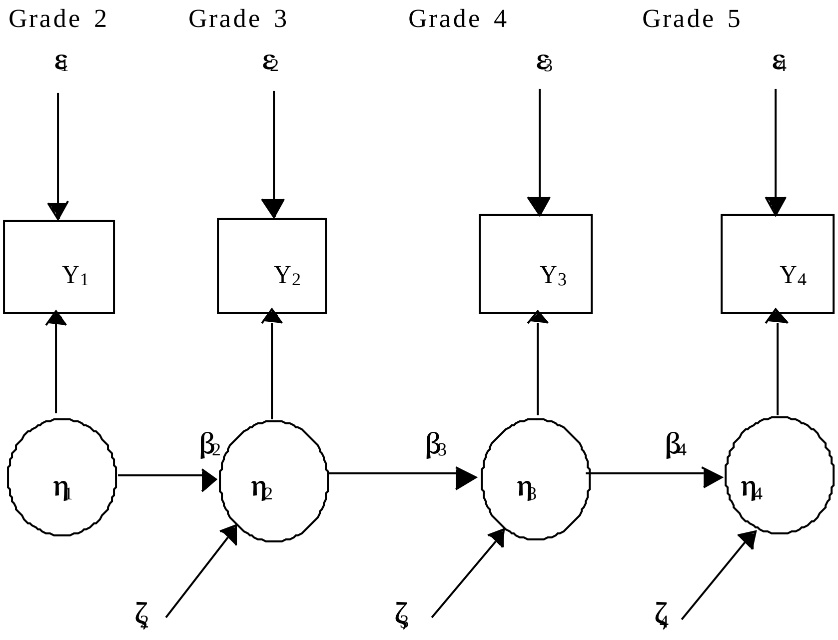

Figure 1. A simplex model.

In order to estimate the parameters of the simplex model, LISREL 7 (Jöreskog & Sörbom, 1989) and EQS (Bentler, 1989) were used. Jöreskog and Sörbom (1989) provide a detailed description of covariance structures and ways to eliminate indeterminacies by restricting the conditions of parameters (p. 185). To avoid indeterminacies, $\varepsilon_1 = \varepsilon_4 = 0$ and $\eta_1 = \zeta_1$ were defined.

Ⅱ. Results

The mathematics grounding the calculation of maximum likelihood estimates assumes that we have a variance-covariance matrix created by recording the value of each individual on all the variables included in the input data matrix (a listwise matrix). In Table 2, the observed variance-covariance matrix of the standard CBM reading measure from the 37 students with reading difficulties in grades 2 to 5 is given. Variances increased as the grade level of the students increased (up to grade 4). This matrix was analyzed under four different hypotheses (quasi-simplex, quasi-simplex with equal εs, perfect simplex, and perfect simplex with equal βs). In Table 3, it is seen that Model 1 (χ^2=.29, df=1, p>.05), Model 2 (χ^2=.45, df=2, p>.05), and Model 3 (χ^2=4.71, df=3, p>.05) demonstrate a good fit to the data, but Model 4 (χ^2=25.95, df=5, p<.01) has a poor fit. Nonsignificant χ^2 values indicate good fitting models and values of Goodness of Fit (GFI) and Bentler-Bonett Fit Index (NFI) greater than .90 are desirable.

Table 2.
Listwise Observed Variance-Covariance Matrix

Grade 2 (y1)	Grade 3 (y2)	Grade 4 (y3)	Grade 5 (y4)
54.55			
113.36	415.80		
151.299	455.17	894.92	
128.04	366.43	693.61	793.13

Table 3.
Summary of Analyses

Model	df	Chi-Square	p	GFI[a]	NFI[b]
1. Quasi-simplex	1	.29	.59	.997	.998
2. Quasi-simplex with Equal Reliabilities	2	.45	.80	.995	.997
3. Perfect Simplex	3	4.71	.20	.956	.966
4. Perfect Simplex with Equal Regression Weights	5	25.95	.00	.804	.812

[a]Goodness of Fit Index (Joreskog & Sorbom, 1989)
[b]Bentler-Bonett Normed Fit Index (Bentler, 1989)

These results implied that the present growth data had a simplex property but the annual magnitudes of growth were not the same across grade levels. The maximum-likelihood solution under Model 3 was chosen, because this model was parsimonious. The parameters in Model 3 were estimated as follows (parameters not listed were not identified):

a. Parameters in β_2=2.08, β_3=1.10, and β_4=0.78
b. Parameters in standardized β (β*): β_2*=0.75, β_3*=0.75, and β_4*=0.82
c. Parameters in y: Var(ζ_1)=54.55, Var(ζ_2)=180.23, Var(ζ_3)=396.65, and Var(ζ_4)=255.55
d. Parameters in ε are zero (perfect simplex model).

While the true unstandardized regression weights (β's) decreased over time, the standardized regression weights (β*'s) were similar. These results implied that, for this basic academic skill (reading fluency), the subsequent influence declined each year from grade 2 to grade 5, but the relative magnitude of contributions from the previous year to the following year was very stable across adjacent grade levels.

Table 4 shows the typical property of the perfect simplex structure. The correlations decrease as one moves away from the main diagonal. Under the perfect simplex structure, every correlation, r_{ij} with $|i-j|>1$, is the product of correlations just below the diagonal (Jöreskog, 1978). For instance, $r_{14} = \rho(\eta_1, \eta_2) * \rho(\eta_2, \eta_3) * \rho(\eta_3, \eta_4) = 0.753 * 0.746 * 0.823 = 0.462$.

Table 4.

Correlation Matrix among True Variables in a Perfect Simplex Model

Grade 2 (η_1)	Grade 3 (η_2)	Grade 4 (η_3)	Grade 5 (η_4)
1.00			
.753	1.00		
.562	.746	1.00	
.462	.614	.823	1.00

So far, the analyses are done based on the listwise covariance matrix based on 37 students' data. It is also conceivable to use a pairwise matrix (65 students, in all), in which each covariance is based on all the cases having information available for only the relevant pair of variables. In the present study, the sample size is relatively small due to attrition, so that comparison of solutions between two different matrices enable to, in a way, demonstrate stability of the model. The last analysis was to fit the perfect simplex model (Model 3) to the two samples (listwise covariance and pairwise covariance) with all Bs constrained to be equal over two groups. As a result, the comparability between two matrices indicates a reasonably good fit to the data (χ^2=10.81, df=5, P>.05, NFI=.960)

Ⅲ. Discussion

The purpose of this study was to investigate the reading growth for the students with severe reading difficulties through the simplex model, since a single indicator was collected at each measurement period. The identified model in the present study was a perfect simplex model. This model was based on a Markov process. That is, given the present state of the system, past states have no influence on any future states of the system. This identified property of the growth data has an implication for prediction. When I predict the reading scores of fifth graders with reading disabilities, I use the reading scores in the previous year (i.e., grade 4), ignoring the influence of the reading levels in grade 2 or 3.

The results are also consistent with pervious findings that reading is very stable over time, based on relative standings (Bloom, 1964; Wessen et al., 1982). Even among students with severe reading difficulties, the scores of students with higher reading skills continued to be higher across grades. Interestingly, in the sense of absolute reading scores, the subsequent influence of the previous reading achievement on the following year's reading achievement slightly declined in the elementary years. The identified simplex structure in reading growth also showed increasing within-grade-level variability with higher grades. Even among the students with severe reading difficulties, the reading levels become more heterogeneous across grade levels (increasing variability).

The professional literature has provided several conventional reliability coefficients (e.g., test-retest reliability, internal consistency, and interrater agreement) on the CBM reading measures (Fuchs, Fuchs, & Maxwell, 1988; Marston, 1989). The results of the present investigation confirm that the CBM reading procedures are very reliable. This is an interesting alternative perspective on the reliability of the CBM measures. The errors of measurement were negligible enough to be assumed (and tested) as zero in this study. The simplex model is useful in detailing the rationale behind the estimation of the reliabilities of each time (Werts, Linn, & Jöreskog, 1977). This method contrasts with the conventional split half or parallel form methods of obtaining test reliabilities which provide a single estimate of reliability. The single reliability estimate calculated by these conventional methods cannot be confirmed or rejected in terms of inconsistency with the data.

Yet, it should be noted that although goodness-of-fit is perhaps a necessary condition for the employment of a simplex model to mirror a growth process, it is not alway sufficient to accept the identified model as the truth. This is due not only to technical, statistical difficulties but also to the philosophy underlying the use of goodness-of-fit tests (Guire & Kowalski, 1979). The problem is that, often, a number of other models will fit the data equally well. Because the very nature of these tests (goodness-of-fit tests) are oriented to rejecting models, not proving them true, exploration of other plausible models and replication of the present study are required. Another caveat is that the present investigation was solely based on the group-level description of intra-individual difference. One problem with this model is the ignorance of inter-individual difference in reading growth. In reality, some students could progress faster than others. Recent methodological advances in the multi-level modeling of growth, encompassing inter-individual differences, address these issues (Collins & Horn, 1990).

Application of a simplex model for students with reading difficulties in a city school district is extremely expensive and difficult because of the small target population (around 5-10% of whole student population) and high attrition rate. Although the number of subjects was relatively small (less than 100), this study was based on all of the available students with severe reading difficulties from a large city school district. The results of this investigation imply the generalization of the simplex model to readers with unique needs (i.e., students identified as reading disabilities).

Again, relative stability in growth across adjacent grade levels was found in this study. However, the influence of reading achievement on the following year's reading achievement declined slightly across elementary years and the relationship between grades became substantially lower as the gap increased (e.g., the correlation coefficient for Grade 2 and Grade 5 was much lower than the correlation coefficient for Grade 2 and Grade 3). This result implies that teaching students with reading difficulties is not doomed to failure. Thereadinggrowth even for students with severe reading difficulties was evident from elementary to intermediate grades, in terms of absolute scores. Undoubtedly, one of the major attributions of their reading progress can be schooling. In the reading instruction of such unique readers, along with the confidence in reading progress, teachers should realize that the present level of reading is the most important consideration for future achievement, probably, not the student's whole past.

Footnote

Under maximum likelihood criterion, with samples of less than 100 cases, there are several cautions, such as convergence failures or inaccuracy of estimates (Anderson & Gerbing, 1984; Boomsma, 1985). In additions, Bentler and Chou (1987) mention as a rule of thumb that the ratio of sample size to the number of free parameters should not go lower than 5:1 if the variables are multivariate normally distributed. Fortunately, the perfect simplex model in this study does not show a failure of convergence and the ratio of sample size to the number of free parameters I higher than 5:1. The results are also almost identical from several different estimation strategies such as maximum likelihood (ML), generalized least squares (GLS), and arbitrary distribution general least squares (AGLS).

REFERENCES

Anderson, J., & Gerbing, D. (1984). The effect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika, 49,* 155-173.

Baltes, P. (1968). Longitudinal and cross-sectional sequences in the study of age and generation effects. *Human Development, 11,* 145-171.

Baltes, P., & Reinert, G. (1969). Cohort effects in cognitive development of children as revealed by cross-sectional sequences. *Developmental Psychology, 1,* 169-177.

Bentler, P. (1989). *EOS: Structural equations program manual.* Los Angeles, CA: BMDP Statistical Software.

Bentler, P., & Chou, C. (1987). Practical issues in structural modeling. *Sociological Methods and Research, 16,* 78-117.

Boomsma, A. (1985). Nonconvergence, improper solutions, and starting values in LISREL maximum likelihood estimation. *Psychometrika, 50,* 229-242.

Bloom, B. S. (1964). *Stability and change in human characteristics.* New York: John Wiley & Sons.

Carlisle, J. F. (1990). Diagnostic assessment of listening and reading comprehension. In H. L. Swanson & B. Keogh (Eds.), *Learning disabilities: Theoretical and research issues.* Hillsdale, NJ: Lawrence Erlbaum Associate.

Collins, L. (1991). Measurement in longitudinal research. In L. Collins & J. Horn (Eds.), *Best methods for the analysis of change.* Washington, DC: American Psychological Association.

Collins, L., & Horn, J. (1991). *Best methods for the analysis of change.* Washington, DC: American Psychological Association.

Clymer, T., & Fenn, T. (1979). *Reading 720 rainbow edition.* Lexington, MA: Ginn.

Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52,* 219-232.

Deno, S. L. (1991). *The nature and development of curriculum-based measurement.* Unpublished manuscript. University of Minnesota, Minneapolis.

Deno, S., Marston, D., Shinn, M., & Tindal, G. (1983). Oral reading fluency: A single datum for scaling reading disability. *Topics in Learning and Learning Disability, 2,* 53-59.

Deno, S. L., & Mirkin, P. K. (1977). *Data-Based program modification: A manual.* Reston, VA: Council for Exceptional Children.

Deno, S. L., Mirkin, P. K., & Chiang, B. (1982). Identifying valid measures of reading. *Exceptional Children, 18,* 16-26.

Deno, S. L., Mirkin, P. K., & Wesson, C. (1984). How towrite effective data-based IEP's. *Teaching Exceptional Children, 16,* 99-104.

Doss, L. S., & Deno, S. (1987). *Aggregated and analyzed Curriculum-Based Measurement data used by a group of Minnesota school districts.* Unpublished manuscript.

Frederiksen, C., & Rotondo, J. (1979). Time-series models and th study of longitudinal change. In J. Nesselroade & P. Baltes (Eds.), *Longitudinal research in the study of behavior and development.* New York: Academic Press.

Fuchs, L. S., & Deno, S. L. (1991). Paradigmatic distinctions between instructionally relevant measurement models. *Exceptional Children, 57(6),* 488-500.

Fuchs, L., Fuchs, D., & Maxwell, L. (1988). The validity of informal reading comprehension measures. *Remedial and Special Education, 9(2),* 20-28.

Guire, K., & Kowalski, C. (1979). Mathematical description and representation of developmental change functions on the intra- and interindividual levels. In J. Nesselroade & P. Baltes (Eds.), *Longitudinal research in the study of behavior and development.* New York: Academic Press.

Guttman, L. A. (1954). New approach to factor analysis: The radex. In P. F. Lazarsfeld (Ed.), *Mathematical thinking in the social sciences* (pp. 258-348.) New York: Columbia University Press.

Jöreskog, K. (1970). Estimation and testing of simplex models. *The British Journal of Mathematical and Statistical Psychology, 23,* 121-145.

Jöreskog, K. (1979). Statistical estimation of structural models. In J. Nesselroade & P. Baltes (Eds.), *Longitudinal research in the study of behavior and development.* New York: academic Press.

Jöreskog, K., & Sörbom, D. (1989). *LISREL 7: A guide to the program and applications.* Chicago, IL: SPSS Inc.

Kim, D. (1993). *Trends of reading growth for students with severe reading difficulties: A four-year longitudinal study* (Unpublished Doctoral Dissertation). Minneapolis, MN: University of Minnesota.

Keogh, B. (1987). Learning disabilities: In defense of a construct. *Learning Disabilities Research, 3(1),* 4-9.

LaBerge, D., & Samules, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6,* 293-323.

Lunneborg, C. E., & Lunnenborg, P. W. (1970). Relations between aptitude changes and academic success during college. *Journal of Educational Psychology, 61,* 169-173.

Marston, D. (1982). *The technical adequacy of direct, repeated measurement of academic skills in low achieving elementary students* (Unpublished Doctoral Dissertation). Minneapolis, MN: University of Minnesota.

Marston, D. (1989). A Curriculum-Based Measurement approach to assessing academic performance: What it is and why do it. In M. Shinn (Ed.), *Curriculum-Based Measurement: Assessing special children.* New York: Guilford.

Marston, D., & Magnusson, D. (1988). *Curriculum-Based Measurement: An Introduction* (Unpublished manuscript). Minneapolis, MN: Minneapolis Public Schools,Department of Special Education.

Potter, M., & Wamre, H. (1990). Curriculum-based measurement and developmental reading models: Opportunities for cross-calidations. *Exceptional Children, 57,* 16-25.

Salvia, J., & Ysseldyke, J. (1991). *Assessment (5th ed.).* Boston: Houghton Mifflin.

Shinn, M. R., Good, R. H., Knutson, N., Tilly, W. D., & Collins, V. L. (1992). *Curriculum-based measurement reading fluency: A confirmatory factor analysis of itsrelationship to reading.* Unpublished manuscript.

Thorndike, R. L. (1966). Intellectual status and intellectual growth. *Journal of Educational Psychology, 57*, 121-127.

Tindal, G., Marston, D., & Deno, S. (1983). *The reliability of direct and repeated measurement (Research Report No. 109).* Minneapolis, MN: University of Minnesota, Institute for Research on Learning Disabilities.

Werts, C. E., Linn, R. L., & Jöreskog, K. G. (1977). A simplex model for analyzing academic growth. *Educational and Psychological Measurement, 37,* 745-756.

Wesson, C., Deno, S., Mirkin, P., Sevcik, B., Skiba, R., King, R., Tindal, G., & Maruyama, G. (1982). *Teaching structure and student achievement effects of Curriculum-Based Measurement: A causal (structural) analysis (Research Report No.105).* Minneapolis, MN: University of Minnesota, Institute for Research on Learning Disabilities.

Willett, J. (1985). *Investigating systematic individual differences in academic growth.* Unpublished doctoral dissertation, Stanford University.