EARLY PREDICTION OF READING ACHIEVEMENT. A FOUR YEAR LONGITUDINAL STUDY

Susan R. Butler & Herbert W. Marsh
The University of Sydney
Marlene J. Sheppard
New South Wales Department of Education
John L. Sheppard
Cumberland College of Health Sciences

Children with severe reading problems which are not diagnosed and treated at an early stage are likely to develop learning disabilities with concomitant disturbances. Unless remediation is instituted at an early stage, the gap between the disabled and his/her peers may continue to widen. The incidence reported varies between 7 and 15 per cent depending upon the criterion. The problem may well represent the major child and adult mental health concern of today (Satz, 1977).

The need for an efficient and accurate means of predicting academic failure at an early age has been raised by many studies (Book, 1975; Butler, 1979; de Hirsch, Jansky and Langford, 1966; Feshbach, Adelman and Fuller, 1977; Pope, Lehrer and Stevens, 1980; Satz and Friel, 1974). A note of caution must be observed as the potential value of any screening device may be marred through inappropriate evaluations and employment of results, thus placing the child at an even greater risk (Keogh and Becker, 1973; Satz and Fletcher, 1979).

There are many theories as to the cause of reading disability encompassing a wide range of aetiological variables. To examine the precursors of the disability further longitudinal studies are indicated. Several such studies have been invaluable to the field. Feshbach, Adelman and Fuller (1977) using two large samples of kindergarten children (N's = 403 and 364) found correlations ranging from .4 to .5 between IQ scores, the de Hirsch-Jansky Predictive Index and teacher ratings at the kindergarten level, and reading achievement scores in grades 1, 2 and 3. Satz, Taylor, Friel and Fletcher (1978) used 16 variables to test kindergarten children and compared their results with reading tests in grades 2 (N = 458) and 5 (N = 442) finding that they could predict relatively well especially at the extremes.

The present study is being carried out to assess the ability of a broad, comprehensive battery of tests to identify at an early age children judged to have a high risk of reading failure. A battery of tests was administered to a large representative sample of kindergarten pupils with reading achievement being later assessed in grades 1, 2 and 3. The major objective was to determine how well reading achievement could be predicted. More specific questions included: 1) do the tests predict reading achievement beyond that which can be explained by IQ alone? 2) does the level of predictability in the three grades differ? 3) does the impact of the different predictors vary in the three grades? and 4) can the use of an initial screening device substantially reduce the amount of testing that needs to be done without sacrificing the ability to detect children with reading problems?

Method

Subjects

The initial sample consisted of all kindergarten pupils in seven schools from the St. George area of metropolitan Sydney. This public school district was chosen because it was representative of the diversity of socio-economic levels and cultural backgrounds found in the state of New South Wales. At the start of the study, tests were administered to a total of 392 children (204 boys and 188 girls). The average age of these children was five years seven months, and ages ranged from five years to seven years two months. The majority of the children in the study (78 per cent) came from homes in which at least one parent spoke English. The largest non-English speaking group was Greek migrants. A more detailed description of the population is presented by Butler (1979).

For each child, diagnostic pretest data were collected in kindergarten (1974), and various reading tests were administered at the end of first (1975), second (1976), and third (1977) grades. Data described in this study are based upon only the 320 children that could be located for the 1977 testing at the end of third grade, representing an attrition rate of 18 per cent. In many cases the children had changed school districts within the city or had moved to another location within the state. Whenever the child could be located, materials were sent to the child's teacher with a request for cooperation and a tester visited the classroom to collect data from the child. Only in cases in which the child had moved out of the state (or out of the country) or there was no forwarding address was the child dropped from the sample.

Materials and Administration

The materials considered in this study are part of a larger, on-going project described in more detail in Butler (1979). Measures included:

NOVEMBER/DECEMBER 1974:

SESSION A:

I. Southern California Sensory Integration Tests (Ayres, 1972).

II. Fine Motor Coordination

A. Matchsticks (Gollnitz, 1961).

B. Tying a Knot (Stanford - Binet). (Terman, Merrill, 1973).

III. Hilda Santucci Copying Geometric Forms Test (Santucci, 1964).

IV. Marianne Frostig Developmental Test of Visual Perception (Shortened Form) (Frostig, 1963).

SESSION B:

I. Letter Recognition Test

II. Auditory Attention Span for Related Syllables (Baker, Leland, 1967).

III. Picture Screening Test of Hearing (Reed, 1960).

IV. Predictive Screening Test of Articulation (Van Riper, 1968).

V. Mira Stambak Rhythm Test (Stambak, 1964).
Test B (2) (Reproduction of Rhythmical Structures)

VI. The Bureau Auditory Comprehension Test (N.S.W. Department of Health).

VII. Peabody Picture Vocabulary Test (Dunn, 1959).

SESSION C:

I. Slosson Intelligence Test (Slosson, 1963).

II. Illinois Test of Psycholinguistic Abilities (Kirk, 1961).
A. Grammatic Closure B. Auditory Closure C. Sound Blending

SESSION D:

I. A Teacher's School Entry Screening Test (Sheppard, 1975).

NOVEMBER/DECEMBER 1975:

Tests administered individually or in groups:

I. Australian Council for Educational Research

A. Lower Grades Reading Test (Level I Group Test).

B. Lower Grades Number Concepts (Level I Group Test). (A.C.E.R., 1962).

II. Daniels and Diack, Standard Test of Reading Skill (Daniels, Diack, 1958). (Test No. I Individual Test)

III. Schonell Graded Word Reading Test (R.I.) (Individual Test) (Schonell, 1955).

OCTOBER/NOVEMBER/DECEMBER 1976:

Individually administered Tests:

I. Doren Diagnostic Reading Test of Word Recognition Skills (Adapted to the Australian population) (Doren, 1973).

II. Schonell Graded Word Reading Test (R.I.) (Schonell, 1955).

OCTOBER/NOVEMBER/DECEMBER 1977:

I. Stanford Diagnostic Reading Test Level I (Karlson, et al., 1966).

II. Schonell Graded Word Reading Test (R.I.) (Schonell, 1955).

Preliminary Data Reduction and Statistical Analysis

Preliminary analysis consisted of determining what variables could be combined, primarily through the use of factor analysis, to produce a more manageable set of variables. After standardization, the two IQ scores were combined to obtain a single IQ variable. Factor analysis of the three reading tests administered in grade 1 (Schonell, Daniels & Diack, and ACER) indicated that a single factor could account for 79 per cent of the variance and that a second factor only had an eigenvalue of 0.52 (an eigenvalue of 1.0 or more is normally required before an additional factor is considered). Consequently, factor score coefficients (Nie, Hull, Jenkins, Steinbrenner & Bent, 1975) were used to weight each of the original scores after standardization to create a single Reading score.

Similarly, factor analysis of the separate subscales of the Doren test and the Schonell (tests administered in grade 2) also indicated that only a single factor existed, which could account for 68 per cent of the variance (the second eigenvalue was 0.79). Factor analysis of the Stanford subscales and the Schonell test (the tests administered in grade 3) also indicated that a single reading score was sufficient; the first factor accounted for 62 per cent of the variance and the second eigenvalue was only .71.

As it is possible that specific information may have been lost by reducing the reading scores for each year to a single summary score, an additional factor analysis was performed on all the different reading tests from all three years (a total of 21 variables including each of the different subscales of the Stanford and Doren tests). The first factor again accounted for over 60 per cent of the variance, though the second and third eigenvalues were over 1.0 (1.29 and 1.02 respectively). However, these last two factors only accounted for an additional 6 per cent and 5 per cent of the variance. Furthermore, even with an oblique rotation, every variable loaded more highly on the first general factor than on either the second or third factor. The second and third factors were highly correlated with the first factor, and no meaningful interpretation could be put on the last two factors. This supplemental analysis offers further support for the decision to consider only a combined reading score for each year.

Factor analysis of the set of predictor variables resulted in six predictor factors that accounted for 62 per cent of the variance in these variables. The predictor factors are as follows:

PRDF I -- Psycholinguistic Abilities. This factor was defined by the three subscales from the Illinois Test of Psycholinguistic Abilities (Grammatic Closure, Auditory Closure, and Sound Blending).

PRDF II -- Figure Drawing. This factor was primarily defined by three tests from the Sheppard School Entry Screening Test that require pupils to copy single geometric patterns presented visually, to draw them from memory when given only components of them, and to draw a human figure. The Santucci copying Geometric Forms Test also loaded on this factor.

PRDF III -- Language. This factor was primarily defined by items from the Sheppard Test that require pupils to repeat nonsense syllables, to distinguish between word pairs, to describe a story, to repeat a short paragraph/single words/sentence, and retell a story. The Bureau Auditory Comprehension Test (a measure of comprehension of spoken language) and the Predictive Screening Test of Articulation also had moderate loadings on this factor.

PRDF IV -- Rhythm. This factor was defined by the Stambak Rhythm Test (a measure of the child's ability to reproduce a tapping pattern) and items from the Sheppard Test that measured the same ability.

PRDF V -- Perceptual Motor. This factor was defined by items from the Sheppard Test requiring pupils to demonstrate proficiency in hopping, jumping, skipping, throwing and catching a ball, walking a beam, etc.

PRDF VI -- Spatial/Form Perception. This factor was primarily defined by three of the four tests from the Frostig Developmental Test of Visual Perception (Position in Space, Eye Motor Coordination, and Spatial Relationships). The Santucci Copying Geometric Forms Test also loaded on this factor.

The Letter Recognition Test did not load substantially on any one factor, but had its highest loadings on the Psycholinguistic, Figure Drawing and Language Factors (PRDF I, PRDF II, and PRDF III). The fourth Frostig Test, Figure Ground, did not load substantially on any factor, but had its highest loading on PRDF V.

Subsequent analyses are based upon relationships between three criterion variables (reading scores for grades 1, 2, and 3), the eight predictor variables (Sex, Parent's Language, and the six predictor factors), and the one control variable (IQ). In the first analysis simple correlations were determined between the set of nine variables, and then the effect of IQ was partialled out of each relationship. In the second analysis, multiple regression was used to determine how well each of the criterion variables could be predicted by the set of predictor variables. Next, a path analytic model was developed to account for the relationships among the predictor and criterion variables. Finally, the effectiveness of using a preliminary screening device in conjunction with a comprehensive battery of tests was explored.

Results

Correlations Among Predictor, Criterion and Control Variables

Correlations among the eight predictor variables - the six predictor factors and the two background variables - and reading scores in grades 1, 2, and 3 are presented in Table 1. The relationship between each of these variables and IQ is also presented, and then the effects of IQ are partialed out of each of the other relationships.

The two background variables, Sex and Parent's Language, are related to each of the three reading scores. Female pupils tend to have higher reading ability, and this effect is relatively independent of IQ. Children coming from families in which neither parent speaks English score much lower on the three reading scores, but this effect

is almost completely eliminated after controlling for the effect of IQ. However, this finding must be viewed with caution in that the IQ tests given in kindergarten are heavily biased against children with little or no English and no attempt was made to test these children in their native language.

Each of the six predictive factors are significantly correlated with each of the reading scores, and these correlations tend to be higher than those between the background variables and the reading scores. Partialling out the effect of IQ reduced the magnitude of the correlations between the reading scores and the predictive factors, but each of the correlations is still statistically significant. The predictive factors are able to predict substantial portions of the variance in reading beyond that which can be accounted for by IQ alone.

Table 1

Correlations Between Predictor Variables and Reading Criterion Scores Before and After (correlations in parentheses) Controlling for Initial IQ

Predictor Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
(1) Sex	-														
(2) Parent's Language	.04	-													
	(.01)														
(3) PRDF I	.03	.45	-												
	(.11)	(.12)													
(4) PRDF II	.02	.18	.23	-											
	(.04)	(.06)	(.09)												
(5) PRDF III	.07	.34	.41	.16	-										
	(.13)	(.04)	(.04)	(.02)											
(6) PRDF IV	.05	.17	.41	.41	.38	-									
	(.08)	(-.04)	(.23)	(.36)	(.22)										
(7) PRDF V	.10	.16	.25	.46	.31	.38	-								
	(.13)	(-.04)	(.02)	(.41)	(.15)	(.29)									
(8) PRDF VI	.00	.28	.31	.55	.21	.41	.47	-							
	(.02)	(.08)	(.04)	(.51)	(-.04)	(.30)	(.38)								
Actual Reading Scores															
(9) Actual Reading Grade 1	.07	.27	.44	.35	.40	.40	.23	.37	-						
	(.11)	(.08)	(.27)	(.29)	(.24)	(.30)	(.12)	(.26)							
(10) Actual Reading Grade 2	.15	.23	.40	.38	.47	.44	.32	.46	.69	-					
	(.20)	(.03)	(.20)	(.32)	(.33)	(.35)	(.21)	(.36)	(.64)						
(11) Actual Reading Grade 3	.16	.29	.49	.42	.49	.49	.41	.48	.69	.88	-				
	(.22)	(.05)	(.27)	(.35)	(.31)	(.38)	(.31)	(.35)	(.63)	(.86)					
Predicted Reading Scores[a]															
(12) Predicted Reading Grade 1	.12	.46	.76	.60	.68	.69	.39	.64	.58	.45	.68	-			
	(.24)	(.13)	(.54)	(.41)	(.48)	(.64)	(.23)	(.54)	(.48)	(.54)	(.56)				
(13) Predicted Reading Grade 2	.24	.36	.62	.59	.72	.68	.49	.71	.56	.65	.69	.96	-		
	(.36)	(.02)	(.32)	(.58)	(.57)	(.62)	(.37)	(.63)	(.45)	(.57)	(.58)	(.44)			
(14) Predicted Reading Grade 3	.23	.41	.70	.59	.69	.69	.59	.68	.56	.65	.70	.97	.98	-	
	(.38)	(.06)	(.44)	(.60)	(.51)	(.64)	(.52)	(.60)	(.45)	(.55)	(.60)	(.43)	(.97)		
Control Variables															
(15) IQ	-.07	.55	.69	.24	.57	.38	.34	.41	.38	.39	.46	.41	.70	.64	-
	(00)	(00)	(00)	(00)	(00)	(00)	(00)	(00)	(00)	(00)	(00)	(00)	(00)	(00)	

[a]Predicted Reading Scores are the reading scores that are predicted on the basis of the eight predictor variables. They were obtained with multiple regression as summarized in Table 2. For example, the multiple correlation between the set of predictor variables and actual reading in grade 3 is .70 (.60 after correcting for IQ).

Combined Effect of Predictor Variables

Stepwise multiple regression (Nie, et al., 1975) was used to determine how well the combined set of eight predictor variables could predict reading ability in grades 1, 2 and 3. These results (see Table 2) indicate that 34 per cent (grade 1), 42 per cent (grade 2) and 49 per cent (grade 3) of the variance in the reading scores is predictable. Identifying the most important predictors - whether based upon simple correlations, beta weights, or order of entry into the regression equation - varies somewhat for the different reading scores, but several generalizations are possible. The most important variables tend to be predictive factors I (Psycholinguistic Abilities), III (Language), and VI (Spatial/Form Perception). Predictive factors II (Figure Drawing) and IV (Rhythm) also entered significantly into each of the regression equations, but PRDF V (Perceptual Motor) did not. Neither of the two background variables, Sex and Language, contributed much to the regression equations beyond that which could be explained by the six predictor factors.

It is particularly interesting to note that reading scores in this study grow substantially more predictable with the passage of time. Measures collected in kindergarten do a better job of predicting reading scores in third grade than in second, and better in second grade than in first. Furthermore, with a few minor exceptions, this pattern holds for each of the predictor variables as well (see Table 1). This pattern is in marked contrast to

most validity studies where the longer the time interval between the collection of predictor and criterion variables, the poorer the prediction. The apparent explanation for this phenomenon is with the measurement of reading. The predicted reading scores, as opposed to the actual reading scores, for each of the three years all correlate .96 or higher with each other. Consequently, the predictor variables predict virtually the same relative performance in each of the three years. This suggests that the criterion is becoming more predictable rather than the predictors doing a better job of predicting, and perhaps the explanation lies in our ability to measure reading in the first grade. There is also the possibility that the particular selection of reading tests used in the present study are responsible for this finding. However, similar phenomena were reported by Clarke et al. (1978) and Feshback et al. (1977).

Table 2

Stepwise Multiple Regressions Used to Predict Reading in Each Grade

Grade 1 Reading			Grade 2 Reading			Grade 3 Reading		
Predictor Variables	Change in Mult R^2	Final Beta Weight	Predictor Variables	Change in Mult R^2	Final Beta Weight	Predictor Variables	Change in Mult R^2	Final Beta Weight
PRDF I	.20	.23**	PRDF III	.22	.31**	PRDF I	.24	.22**
PRDF II	.06	.17**	PRDF VI	.14	.26**	PRDF VI	.12	.20**
PRDF III	.05	.22**	PRDF IV	.03	.13*	PRDF III	.08	.26**
PRDF VI	.01	.15**	Sex	.02	.12*	PRDF II	.02	.13*
PRDF IV	.01	.12*	PRDF I	.01	.13*	Sex	.02	.12*
PRDF V	.01	-.10	PRDF II	.01	.14*	PRDF IV	.01	.13*
Sex	.00	.05	PRDF V	.00	-.05	PRDF V	.00	.06
Language	.00	-.01	Language	.00	.04	Language	.00	.01
Mult R_2	.58		Mult R_2	.65		Mult R_2	.70	
Mult R^2	.34		Mult R^2	.42		Mult R^2	.49	

Note: The order in which the variables entered the regression equation was determined by which variables could contribute the most towards explaining residual variance, so the order varied somewhat for each of three grades. Multiple correlation coefficients based upon a sample are positively biased, depending upon the sample size and the number of variables. In this application, correction for this bias would lower the Mult R^2 values by about .01.

Path Analysis

A path-analytic model was developed to further explore the relationships among the eight predictor variables and the three reading scores (see Figure 1). For purposes of this model, the two background variables were assumed to precede the predictor factors. This analysis allows the researcher to decompose correlations into direct, indirect and spurious effects (see Wolfe, 1980). Sex, for example, has a direct effect on second grade reading; the effect, along with the standardized path coefficient, is represented as a straight arrow connecting the two variables. Sex has no indirect effect since it is not significantly related to any of the predictor factors. Parent's Language, in contrast, has no direct effect on any of the reading scores, but has indirect effects through each of the predictor factors.

The predictor factors, unlike the background variables, have large and substantial direct effects on the reading scores. Even here, however, the direct effect occurs primarily in the first year, a little in the second, and almost none in the third year. The variables measured by the predictor factors influence earlier reading ability, which in turn influences latter reading. Nearly all of the influence of the predictor factors on third grade reading is indirect, through first and second grade reading. Poorer reading in the third grade is indirectly related to these variables in that they predict reading achievement in earlier years which in turn leads to poorer reading in later years.

Expectancy Tables

The correlation coefficient, while being a convenient index of the relationship between predicted and actual scores, has several disadvantages. First, it fails to describe how accurately decisions can be made at a particular point (e.g., how accurately children with reading scores below a given cut off can be identified). Second, the correlation cannot be easily translated to nontechnical people. An alternative way of representing the predictor-criterion relationship that overcomes these disadvantages is an expectancy/classification table and the calculation of hit and miss ratios. In the present application, for example, the target population might be the 10 per cent of the children that score lowest on third grade reading. A criterion must then be selected to classify children "at risk" (i.e., likely to have reading problems) on the basis of the predicted third grade reading scores. Using 10 per cent as the selection criterion, 20 children (out of a target population of 32) were correctly identified as having reading problems. Of the 288 who actually did not have reading problems, 272 were correctly classified. The resulting "hit rate", the percentage of children correctly classified, was 95 per cent ((16 + 272)/320). The

miss rate, the percentage of children incorrectly classified, is one minus the hit rate.

The hit rate and the number of children in the target population who are correctly classified will vary with the size of the target population and the selection criterion. The optimal values will depend upon the particular application and the costs involved for various types of misses. For example, using a stringent criterion (e.g., identifying only the bottom 5 per cent of the predicted scores as being at risk) will minimize the number of children incorrectly predicted to have reading problems, but only at the expense of missing more children who do. Results for four possible target groups (bottom 5 per cent, 10 per cent, 15 per cent and 20 per cent based on actual reading scores) and four criterion groups (bottom 5 per cent, 10 per cent, 15 per cent and 20 per cent based on predicted reading scores) are summarized in Table 3 (Strategy I).

Table 3

Classification of Children as Being "At Risk" Based Upon Three Strategies: Hit Rate and Number of Target Population Correctly Classified

Strategy I - Comprehensive Testing of All Children

Target Population	Selection Criterion: Bottom 5%	Bottom 10%	Bottom 15%	Bottom 20%
Bottom 5% (N = 16)	95% (8)	93% (12)	88% (13)	84% (14)
10% (N = 32)	93% (13)	93% (20)	89% (23)	86% (25)
15% (N = 48)	89% (14)	90% (24)	88% (28)	84% (31)
20% (N = 64)	84% (14)	86% (25)	85% (32)	83% (37)

Strategy II - Eliminating 50% of Sample on Basis of Letter Recognition Test and Comprehensive Testing of Remaining 50%

Target Population	Selection Criterion: Bottom 5%	Bottom 10%	Bottom 15%	Bottom 20%
Bottom 5% (N = 16)	94% (7)	92% (11)	87% (12)	83% (13)
10% (N = 32)	93% (12)	92% (19)	89% (22)	85% (24)
15% (N = 48)	88% (13)	89% (23)	87% (27)	84% (30)
20% (N = 64)	83% (13)	85% (24)	84% (31)	82% (36)

Strategy III - Elimination of 75% of Sample on Basis of Sheppard Test and Comprehensive Testing of Remaining 25%

Target Population	Selection Criterion: Bottom 5%	Bottom 10%	Bottom 15%	Bottom 20%
Bottom 5% (N = 16)	95% (8)	93% (12)	88% (13)	84% (14)
10% (N = 32)	93% (13)	93% (20)	89% (23)	86% (24)
15% (N = 48)	89% (14)	90% (24)	88% (28)	84% (30)
20% (N = 64)	84% (14)	86% (25)	85% (32)	82% (35)

Effectiveness of Preliminary Screening

Results presented thus far have emphasized the ability of a broad, comprehensive set of predictor variables to forecast reading scores. However, limitations in time, money and professional staff may render the administration of such a large battery as unfeasible in many settings. An alternative strategy would be to administer a preliminary screening device, to select children who are likely to be at risk on the basis of the preliminary screening, and then to administer a more comprehensive battery to these children. Results presented in this section test the effectiveness of this alternative strategy with two possible screening devices.

The first screening device is the Letter Recognition Test. The results of this test alone correlated .56 with third grade reading, compared to .70 for the entire battery. Previous work with a different population of children has also shown this to be an effective predictor of reading. In this test, children are simply asked to identify the letters of the alphabet when presented in a random order. Consequently the test can be quickly and easily administered by a person without any special training.

To test the effectiveness of this test as a screening device, children scoring in the top 50% on the Letter Recognition Test were automatically classified as not being at risk. The predicted scores for the remaining 50% of the children were then used to make further classifications. The results (see Table 3 - Strategy II) demonstrate that classification of children into the at risk group using this strategy is nearly the same as based upon the comprehensive testing of all 320 children.

The second screening device to be considered is the Sheppard School Entry Screening Device. This test (Sheppard, 1972) is specifically designed to provide early detection of children likely to have reading problems. Previous research with this same sample of children indicated that the Sheppard Test can be described by three factors (Language, Figure Drawing, and Perceptual Motor), and that these factors correlate .59 with reading in

grade 3 (Butler, Marsh, Sheppard & Sheppard, (in press). The Sheppard Test, while providing a more accurate assessment of reading than the Letter Recognition Test, is also designed to be administered by a classroom teacher.

To test the effectiveness of the Sheppard Test as a screening device, children predicted to fall in the top 75 per cent of the sample on the basis of the Sheppard Test were automatically classified as not being at risk. The predicted scores for the remaining 25 per cent of the children were then used to make further classification. The results (see Table 3 - Strategy III) again show that there is little or no loss in the ability to identify children at risk when using this strategy.

Results presented in this section indicate that the identification of children who are "at risk" does not require that all children be given the entire battery of tests. Alternatively, a preliminary screening device can be used to identify children who are particularly unlikely to have reading problems, and then the comprehensive battery can be given to the remaining children. Depending upon the sensitivity of the screening device, 50 per cent or 75 per cent of the original sample can be excluded from the comprehensive testing with little or no loss in the identification of children who are likely to have reading problems.

Discussion

A set of six predictor factors underlay a broad comprehensive set of psychological tests administered to kindergarten children. Each of these six factors was substantially correlated with reading performance in grades 1, 2 and 3, even after partialling out the effect of IQ. Two background variables, child's sex and whether or not at least one parent spoke English, were not so highly correlated with reading achievement and contributed little beyond what could be predicted by the six predicted factors. The multiple correlations relating the predictor variables to reading achievement were .58 (grade 1), .65 (grade 2), and .70 (grade 3).

Correlations between the predictor variables and reading achievement increased from grades 1 to 3. Similar findings were reported by Clark et al., (1978). Feshbach et al. (1977) also found that the correlations tended to increase from grade 1 to grade 2, but not for grade 3. This pattern is unusual in that the predictability of criteria normally drops as the time between the collection of the predictor and criterion variables increases. The findings in this study suggest that the measurement of reading achievement becomes more reliable, and not that any of the predictor variables become more relevant. This pattern suggests an important limitation in studies that are based upon shorter periods of time.

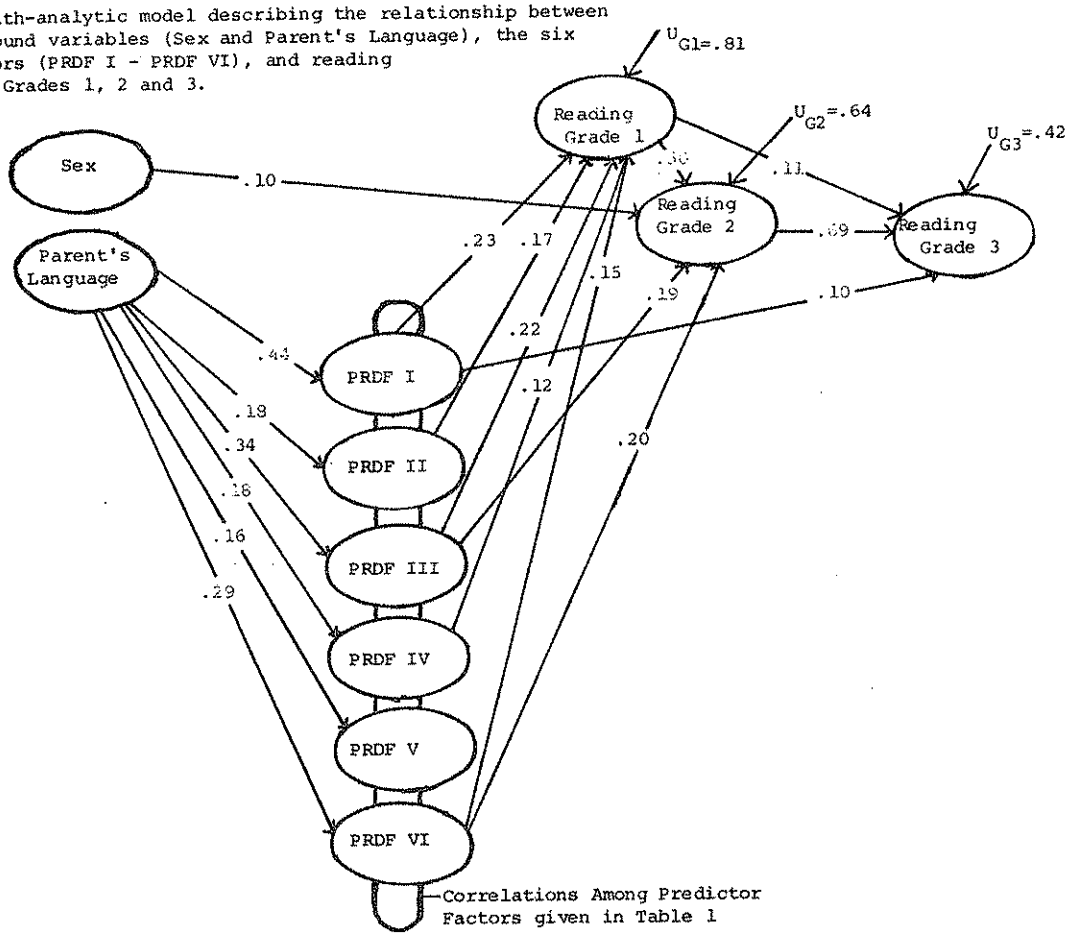
Path analysis indicated that none of the predictor variables had a large direct effect on second or third grade reading, even though each was substantially correlated with these reading scores. The largest direct effects on third grade reading was for first and second grade reading, and the largest direct effect on second grade reading was first grade reading. This illustrates that the skills measured by the tests administered in kindergarten are not directly related to later reading achievement, but are related to earlier reading achievement which then impacts later reading. These findings argue for a developmental understanding of reading problems. Other investigators also argue that the acquisition of reading skills (Satz et al., 1976; de Hirsch, Jansky & Langford, 1966) and language in general (de Hirsch, 1974) is a developmental process.

The results of this study indicate that reading achievement can be accurately predicted on the basis of a broad, battery of psychological tests administered in kindergarten. However, if the emphasis is on the early identification of children particularly likely to have reading problems, then a much more economical strategy is available. Instead of comprehensively testing all children, a preliminary screening device can be used to identify children who are particularly unlikely to have reading problems and they can be excluded from further testing. Then, the comprehensive test battery can be administered to the remaining children. Results of the present investigation indicate that, depending upon the sensitivity of the preliminary screening device, administration of the test battery can be limited to 50 per cent or even 25 per cent of the children with little or no loss in the ability to identify children with reading problems.

In summary, the results of this study indicate that a broad, comprehensive battery of tests administered in kindergarten can effectively predict reading achievement in grades 1, 2 and 3. These tests predict substantial variance beyond that which can be explained by IQ, sex or parent's language. Furthermore, the accuracy of the predictions actually increase from first to third grade. Nevertheless, the skills measured by the battery are directly related to earlier reading, and this early reading achievement then directly impacts latter reading achievement. When the emphasis is upon the identification of reading problems, administration of a preliminary screening device can substantially reduce the amount of testing that needs to be done and still not hurt the ability to

identify reading problems.

Figure 1. A path-analytic model describing the relationship between the two background variables (Sex and Parent's Language), the six predictor factors (PRDF I - PRDF VI), and reading achievement in Grades 1, 2 and 3.



References

Australian Council for Educational Research, Lower Grades Reading Test (Level 1 Group Test).

----, Lower Grades Number Concepts (Level 1 Group Test), 1962.

Ayres, A.J., Southern California Sensory Integration Tests, California Western Psychological Services, 1972.

Baker, H.J. and Leland, B., Detroit Test of Learning Aptitude, Indiana: Bobbs-Merrill Company, 1967.

Book, R.M., Predicting reading failure: a screening battery for kindergarten children. Journal of Learning Disabilities, 1975, 7, 52-56.

Butler, S.R., Predictive antecedents of reading disability in the early years of schooling. British Columbia Journal of Special Education, 1979, 3, 263-274.

Butler, S.R., Marsh, H.W., Sheppard, M.J. and Sheppard, J.L., Early prediction of reading achievement with the Sheppard School Entry Screening Test: A four year longitudinal study (in press).

Clark, C.R., Bruininks, R.H. & Glaman, G.V., Kindergarten predictors of three aspects of reading achievement. Perceptual and Motor Skills, 1978, 46, 411-419.

Daniels, J.C. and Diack, H., The Standard Test of Reading Skill (Test No. 1 Individual Test). London: Chatto and Windus Ltd., 1958.

de Hirsch, K., Jansky, J.J. & Langford, W.S., Predicting Reading Failure. A Preliminary Study. New York: Harper & Row, 1966.

de Hirsch, K., Early language development. In S. Arieti (Ed.), American Handbook of Psychiatry. Vol. 1. The Foundations of Psychiatry. New York: Basic Books, 1974, pp. 352-367.

Department of Health (New South Wales). Bureau Auditory Comprehension Test.

Doren, M., Doren Diagnostic Reading Test of Word Recognition. Minnesota: American Guidance Service, Inc., 1973.

Dunn, L.M., Peabody Picture Vocabulary Test, American Guidance Service, 1959.

Feshbach, S., Adelman, H. & Fuller, W., Prediction of reading and related academic problems. Journal of Educational Psychology, 1977, 69, 299-308.

Frostig, M., Developmental Test of Visual Perception, California: Consulting Psychologists Press, 1963.

Gollnitz, G., Oseretsky and Kwint Test. Unpublished, 1961.

Karlson, B., Madden, R. and Gardner, E.F., Stanford Diagnostic Reading Test. New York: Harcourt, Brace and World, Inc., 1966.

Keogh, B.K. and Becker, L.D., Early detection of learning problems: questions, cautions, and guidelines, Exceptional Children, September 1973.

Kirk, S.A., McCarthy, J.J. and Kirk, W., The Illinois Test of Psycholinguistic Abilities. Urbana, Illinois: Institute for Research in Exceptional Children, 1961.

Nie, N.H., Hull, C.H., Jenkins, J.G., Steinbrenner, K. and Bent, D.H., Statistical Package for the Social Sciences. New York: McGraw Hill, 1975.

Pope, J., Lehrer, B. & Stevens, J., A multiphasic reading screening procedure. Journal of Learning Disabilities, 1980, 13, 98-102.

Reed, M., Picture Screening Test of Hearing. London: Royal National Institute for the Deaf, 1960.

Santucci, H., Épreuve Graphique D'Organisation Perceptive Pour Enfants de 4 à 6 ANS. In R. Zazzo (Ed.), Manuel Pour l'Examen Psychologique de l'Enfant. Delachaux et Niestlé, 1964.

Satz, P. and Friel, J., Some predictive antecedents of specific reading disability: A preliminary two-year follow-up. Journal of Learning Disabilities, Sept. 1974, 7, 7.

Satz, P., Friel, J. and Rudegeair, J., Some predictive antecedents of specific reading disability: a two-, three-, and four-year follow-up. In J. Guthrie (Ed.), Aspects of Reading Acquisition. Baltimore: Johns Hopkins University Press, 1976, pp. 111-140.

Satz, P., Reading problems in perspective. In W. Otto, N.A. Peters and C.W. Peters (Eds), Reading Problems: A Multidisciplinary Perspective, Addison-Welsley Publishing Co., 1977.

Satz, P., Taylor, H.G., Friel, J. and Fletcher, J., Some developmental and predictive precursors of reading disabilities: a six year follow-up. In A.L. Benton and D. Pearl (Eds), Dyslexia. An Appraisal of Current Knowledge. New York: Oxford University Press, 1978, pp. 313-347.

Satz, P., Fletcher, J., Early Screening Tests: Some Uses and Abuses. Journal of Learning Disabilities, Jan. 1979, 12, 1.

Schonell, F.J. and Schonell, F.E., Schonell Reading Test. London: Oliver and Boyd, 1955.

Sheppard, M.J., A Teacher's School Entry Screening Test. Sydney: Robert Burton Printers 1972. (a) 3rd Edn., 1975.

Slosson, R.L., Slosson Intelligence Test, Slosson Educational Publications, Inc., 1963.

Stambak, M., Trois Epreuves de Rhythme. In R. Zazzo (Ed.), Manuel Pour l'Examen Psychologique de l'enfant. Neuchatel, Suisse: Delachaux et Niestlé, 1964.

Terman, L.K. and Merrill, M.A., Stanford - Binet Intelligence Scale. (Third Revision Form L - M (1960), Houghton Mifflin Co., 1973.

Van Riper, C., Erickson, R.L., Predictive Screening Test of Articulation. Michigan: Western Michigan University, Continuing Education Office, 1968.

Wolfe, L.M., Strategies of path analysis. American Educational Research Journal, 1980, 12, 183-209.