A Longitudinal Analysis of Sex Differences in Latent Achievement.

While previous studies investigating sex differences in the growth of academic abilities have been inconsistent in their conclusions, sex differences have often been reported, particularly in the areas of verbal and mathematical achievement. Results from several longitudinal studies have suggested that these measured differences in achievement are negligible at the elementary school level but are significant thereafter. The objective of the current investigation was to examine the possibility of sex differences with respect to true-score achievement in the verbal and mathematics domains over time, employing a longitudinal model for achievement which included both a covariance structure and a mean structure. The data were taken from the Growth Study at Educational Testing Service. Four measures of verbal achievement and two measures of quantitative achievement were examined at four time periods in consecutive 2-year intervals for 1,433 boys and 1,594 girls from grades 5 to 11. The analysis indicated that there were only slight differences in latent variable achievement means on verbal measures, but widening differences on quantitative measures. The means indicated that boys' mathematical skills increased faster than girls' skills from grades 5 to 11. The results do not support the contention that sex differences exist with respect to verbal achievement, but do support the notions that sex differences exist in the mathematics domain and that these differences are increasingly divergent across time. (Author/Note)
A Longitudinal Analysis of
Sex Differences in Latent Achievement
Wendy J. VanLoy and Robert Cudeck
University of Minnesota
August 10, 1990

We are grateful to T. L. Hilton and Educational Testing Service for use of these data. We would also like to acknowledge the support of Lisa L. O'Dell, Kelli J. Klebe, and Joseph A. VanLoy throughout the development of this paper. The first author was partially supported by an APA Graduate Student Travel Grant to the Annual Convention in Boston, MA, August, 1990.
Sex Differences 2

Abstract

While previous studies investigating sex differences in the growth of academic abilities have been inconsistent in their conclusions, sex differences have often been reported, particularly in the areas of verbal and mathematical achievement. Results from several longitudinal studies have suggested that these measured differences in achievement are negligible at the elementary school level but are significant thereafter. The objective of the current investigation is to examine the possibility of sex differences with respect to true-score achievement in the verbal and mathematics domains over time, employing a longitudinal model for achievement which includes both a covariance structure as well as a mean structure. The data are taken from the Growth Study at Educational Testing Service. The analysis indicates that there are only slight differences in latent variable achievement means on verbal measures, but widening differences on quantitative measures. The results do not support the contention that sex differences exist with respect to verbal achievement, but do support the notions that sex differences exist in the mathematics domain and that these differences are increasingly divergent across time.
Sex Differences 3

Introduction

Gender differences have often been reported on tests of general intelligence and specific intellectual abilities (Rohrbaugh, 1979, p. 63). In particular, some studies have shown that girls generally outperform boys on verbal and linguistic ability measures, while boys do better on visual-spatial, arithmetical reasoning, and numerical ability measures (Maccoby & Jacklin, 1974; Rohrbaugh, 1979). Such sex differences are not usually apparent at the grade school level but first appear in junior high school and become increasingly more pronounced in the following years (Armstrong, 1981; Fennema, 1974, 1980, 1983; Fennema & Sherman, 1978; Hilton & Berglund, 1974; Maccoby, 1966; Maccoby & Jacklin, 1974; Mednick, Harway, & Finello, 1984).

Several explanations have been made regarding possible causes of these differences (e.g., Hilton and Berglund, 1974). The first concerns the possibility that males adopt a more precise and analytical approach to problem solving which enables them to handle mathematical tasks more efficiently than females (Witkin, Dyk, Faterson, Goodenough, & Karp, 1962). Thus, some argue that it is simply superior male mathematical ability that accounts for the sex differences in achievement (Benbow & Stanley, 1980). This is in contrast to the position of other researchers who explain the better performance of boys by suggesting that differential mathematics achievement results from differential patterns of quantitative course work taken by males and females (Ethington & Wolfe, 1984; Pallas & Alexander, 1983; Wise, Steel, & MacDonald, 1979, as cited in Ethington & Wolfe, 1984, 1986).
It may also be that sex differences in problem solving result from sex differences in attitudes toward problem solving (Alpert, Stellwagen, & Becker, 1963; Carey, 1958). Significant positive correlations have been found between favorable attitudes toward problem solving and achievement in mathematics (Lindgren, Silva, Faraco, & Da Rocha, 1964). Children's attitudes toward mathematics are also related to the extent to which parents support and encourage mathematics education for their children and to the parent's understanding of the educational goals of school mathematics courses (Alpert et al., 1963). A sex-role identification hypothesis has also been proposed. In this view, masculinity is associated with better problem-solving skills (Milton, 1957). Since children tend to imitate the same-sex parent, differential achievement patterns result when girls more strongly associate with feminine roles in verbal and non-quantitatively oriented tasks, while boys associate with quantitative tasks which are seen as more masculine (Milton, 1957). Maccoby (1966) and Fennema and Sherman (1977) suggest that sex differences in achievement are directly related to sex-typed interests. Boys and girls become interested in, and proficient at, the kinds of tasks in which they receive greater encouragement and which most often coincide with the roles they are currently in or are expected to fulfill in the future. This position is supported by Sherman (1980) who used attitudinal variables to show the importance of sex-role factors in the development of sex differences in mathematics achievement.

These hypotheses, while plausible to some degree, are generally based on conclusions obtained from the analysis of measured variables in studies employing single tests, cross-sectional designs, and/or small samples. Although this approach is valuable, a weakness of such analyses is that
scores in the behavioral sciences often contain sizeable measurement error. This error attenuates relationships among true scores and may mask important differences which the latent variables display. Consequently, the performance of boys and girls on latent variables may not be the same as performance based on measured variables. Because of the effects of error, significant differences in measured variables may not show up in the associated latent variables, while the converse could also occur.

Another criticism of much of the literature on this topic is that of sample selectivity. Larger differences are often reported in studies with more homogeneous samples (Fennema, 1974; Hyde, 1981; Hyde, Fennema & Lamon, 1990; Kimball, 1989). These results are often widely publicized but typically fail to acknowledge the somewhat selective nature of the study. Obviously, generalization from a study of this kind is very limited.

Still another concern with most studies of gender differences in verbal and quantitative achievement is that they are often based on cross-sectional research. Consequently, with a different group selected at each testing occasion individual growth is not measured as such. In interpreting results of this kind, researchers "cannot be certain whether the observed differences among age groups are produced by the aging process itself, by generational or cultural differences (cohort differences), or by time-related changes in the attitudes and values of society" (Aiken, 1989, p. 23). Rogosa (1979, p. 265) has concluded that "(r)epeated measurements on the same individuals are essential for assessment of individual growth and change."
Methodology

This paper examines latent growth in verbal and mathematical achievement from grades five to eleven. It uses a longitudinal model with structured means and invariant factors (Cattell, 1944; McDonald, 1984) to study the change in two latent variables—verbal and quantitative—across four time periods.

The data were selected from the Growth Study (Hilton et al., 1971), which embodied both cross-sectional and longitudinal components. It used a large, national, representative sample of four cohorts from 17 communities and 27 public elementary, junior-high, and senior-high schools. The Growth Study's initial focus (Hilton et al., 1971, p.11) was on describing and explaining the development of the "student's acquisition of knowledge, understanding and intellectual skills."

Academic performance was assessed at each grade level by several tests of ability and achievement, specifically the School and College Ability Test (SCAT) and the Sequential Tests of Educational Progress (STEP). The former battery measures general verbal and quantitative achievement while the later yields scores reflecting the student's problem-solving abilities in reading, writing, listening, social studies, science, and mathematics.

We examined four measures of verbal achievement and two measures of quantitative achievement at four time periods in consecutive two-year intervals for boys (n = 1433) and for girls (n = 1594) from grades five to eleven. Reading, writing, listening, and verbal tasks were taken as measures of latent verbal achievement while latent quantitative achievement was measured by math and quantitative tasks. Following the work of both
Cattell (1944) and McDonald (1980), we assumed that the regressions of the measured variables on the factors are invariant across time (i.e., across the four grades), that the correlation between the two factors is also invariant, but that the variances and means of the factor scores (latent traits) change from one grade level to another (Lord & Novick, 1968; McDonald, 1982, 1984). More specifically, we assumed an auto-regressive relationship between the latent factors at Grades 5, 7, 9, and 11. Details of the model are diagrammed in Figure 1.

The development of the latent variable and measurement models with structured means follows that described by McDonald (1984). The model for the covariance structure is essentially derived from Cattell’s (1944) work on parallel proportional profiles.

The general form of the invariant factors multimode model (McDonald, 1984) is given by:

\[
x_t^g = \nu + \Lambda \xi_t^g + \delta_t^g \quad \{ g = 1, 2 \; ; \; t = 1, 2, 3, 4 \}
\]

where there are q tests, p occasions, and r latent factors.

- \( g \) denotes boys and girls.
- \( t \) is an index that varies between 1 and p.
- \( \Lambda \) is a \((q \times r)\) matrix of factor loadings.
- \( \xi \) is a \((r \times 1)\) vector of factor scores (latent variables).
- \( \nu \) is a \((q \times 1)\) vector of intercepts.
- \( \delta \) is a \((q \times 1)\) vector of residual errors.
Then the measurement model in the present case can be written, at each occasion, as:

\[
\begin{bmatrix}
\text{STEP-reading} \\
\text{STEP-writing} \\
\text{STEP-listening} \\
\text{SCAT-verbal} \\
\text{STEP-math} \\
\text{SCAT-quantitative}
\end{bmatrix}
= \begin{bmatrix}
\nu_1 \\
\nu_2 \\
\nu_3 \\
\nu_4 \\
\nu_5 \\
\nu_6
\end{bmatrix}
+ \begin{bmatrix}
\lambda_1 & 0 \\
\lambda_2 & 0 \\
\lambda_3 & 0 \\
\lambda_4 & 0 \\
0 & \lambda_5 \\
0 & \lambda_6
\end{bmatrix}
\begin{bmatrix}
\xi \\
Q
\end{bmatrix}
+ \begin{bmatrix}
\delta_1 \\
\delta_2 \\
\delta_3 \\
\delta_4 \\
\delta_5 \\
\delta_6
\end{bmatrix}
\]

The autoregressive model for the change in the latent variables is

\[
\xi_{t+1} = \Gamma \xi_t + \epsilon_t
\]

where \( \Gamma \) is a \((r \times r)\) diagonal scaling matrix.

The means of the observed and latent variables are denoted, respectively, as

\[
E(x_t) = \mu_t \\
E(\xi_t) = K_t
\]

The mean structure is

\[
\mu_t = \nu + \Lambda K_t
\]

To determine the general covariance structure of the model we start by describing the covariance matrices of the latent variables across time.

At \( t = 1 \), the \((r \times r)\) matrix is

\[
\text{cov}(\xi_1) = \Phi \quad \text{such that } \text{diag}(\Phi) = I.
\]

The covariance matrix of these factor scores for any pair of occasions is

\[
\text{cov}(\xi_j, \xi_k) = \Gamma_j \Phi \Gamma_k \quad \text{where } \Gamma_1 = I.
\]

The covariance structure between the observed variables at occasions \( j \) and \( k \) is then assumed to be:

\[
\text{cov}(x_j, x_k) = \begin{cases}
\Lambda_j \Phi_j \Lambda_k, & j = k \\
\Lambda_j \Phi_j \Gamma_k \Lambda_k, & j \neq k, \quad j = 1, \ldots, t; k = 1, \ldots, t
\end{cases}
\]
The covariance matrix among residuals at any occasion, \( \delta_t \), was assumed to be diagonal, \((q \times q)\), with zero covariance between occasions. So,

\[
\text{cov}(\delta_t) = \Theta_t = \text{diag}(\theta^2_{1t}, \ldots, \theta^2_{qt})
\]

and \( \text{cov}(\delta_t, \delta_{t'}) = 0 \) where \( t \neq t' \).

Finally, the covariance structure of the invariant factors multimode model is given by:

\[
\Sigma \sim \Lambda_x \Gamma \Phi \Lambda_x' + \Theta_\delta;
\]

or, more completely (in the form of the present case),

\[
\Sigma = \begin{bmatrix}
\Lambda & & & \\
& \Lambda & & \\
& & \Lambda & \\
& & & \Lambda
\end{bmatrix} \Phi \begin{bmatrix}
\Gamma_1 \\
\Gamma_2 \\
\Gamma_3 \\
\Gamma_4
\end{bmatrix} \begin{bmatrix}
\Lambda' \\
\Lambda' \\
\Lambda' \\
\Lambda'
\end{bmatrix} + \begin{bmatrix}
\Theta_1 \\
\Theta_2 \\
\Theta_3 \\
\Theta_4
\end{bmatrix}
\]

such that \( \Lambda_x \) is a \((tq \times tr)\) matrix of factor loadings

\( \Gamma \) is a \((tr \times r)\) matrix of scaling terms

\( \Theta_\delta \) is a \((tq \times tq)\) diagonal matrix of residual errors.

A series of simultaneous two-group analyses were performed to investigate the differences between the two groups. The model presented here is, in the opinion of the authors, the best of the seven models examined. It was judged to be the most efficient based on Steiger and Lind's (1980) Root Mean Square Error of Approximation (RMSEA) criteria.
Results

The parameter estimates of the complete model are presented in Table 1. No numerical problems are evident, and the estimates seem reasonable. Variability in verbal and mathematical achievement increases across time. There was a relatively high correlation, 0.87, between the two latent factors.

The focus of our investigation, however, was on the mean structure incorporated in the model. Only a slight difference in latent means on the verbal factor between sexes exists, as seen in Figure 2, but this small performance difference of boys compared to girls does not support the result so often reported in the literature that girls have greater verbal ability than boys and that gender differences on the verbal factor increase with age (Maccoby & Jacklin, 1974).

On the quantitative factor, however, the results do indicate a gender difference in latent achievement. The means indicate that boys' mathematical skills increased faster than girls from grades five to eleven. The results also supported the contention that this difference in mathematical achievement between boys and girls widens with time, as the difference between the latent variable means at each successive grade level evaluated is greater than the difference at the previously evaluated time.
Conclusions

The study of possible sex differences in achievement has gone on for decades now, but the results of these investigations remain highly inconsistent. Particularly with respect to mathematics achievement, no consistent explanation has been established for the performance difference often found between boys and girls. Some argue that attitudes toward mathematics have strong effects on quantitative achievement (Ethington & Wolfle, 1986; Sherman, 1980), whereas others believe that sex differences in participation in mathematics courses cause the difference in mathematics achievement between males and females (Ethington & Wolfle, 1984; Pallas & Alexander, 1983; Wise, et al., 1979, as cited in Ethington & Wolfle, 1984, 1986). Two additional hypotheses consider the notions that quantitative achievement is due to differential socialization processes (Fennema & Sherman, 1977; Maccoby, 1966; Sherman, 1980), or simply, that superior male mathematical ability accounts for the sex differences in achievement on the quantitative factor (Benbow & Stanley, 1980).

We agree with others (Ethington & Wolfle, 1984, 1986; Hanna & Lei, 1985) who have argued that in order to address these issues, and the more elementary questions of whether or not sex differences in achievement exist and how these differences develop, one should perform a latent variable analysis on a large, heterogeneous, longitudinal sample. A latent variable model allows a synthesis of information obtained from the multiple measures of the theoretical variables of interest. It also permits the assessment of achievement with minimal influence of measurement error.
The present secondary-analysis of the ETS growth data supports the difference between boys and girls in quantitative achievement. Evaluation of the latent variable means indicates that boys mathematical achievement follows a faster developmental pattern than girls, and the differences which exist in quantitative achievement do become increasingly more divergent as boys and girls age. The latent variable analysis indicates only a minor sex difference on the latent verbal factor, however. These results do not support previous findings showing a significant difference between boys and girls on verbal achievement.
Sex Differences 13

References


Figure 1

The Invariant Factors Multimode Model (McDonald, 1984)
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Boys</th>
<th>Girls</th>
<th></th>
<th>Boys</th>
<th>Girls</th>
</tr>
</thead>
<tbody>
<tr>
<td>$v_1$</td>
<td>255.2</td>
<td>260.6</td>
<td></td>
<td>$\lambda_1$</td>
<td>12.4</td>
</tr>
<tr>
<td>$v_2$</td>
<td>267.3</td>
<td>267.4</td>
<td></td>
<td>$\lambda_2$</td>
<td>9.3</td>
</tr>
<tr>
<td>$v_3$</td>
<td>251.3</td>
<td>258.7</td>
<td></td>
<td>$\lambda_3$</td>
<td>11.6</td>
</tr>
<tr>
<td>$v_4$</td>
<td>251.3</td>
<td>251.6</td>
<td></td>
<td>$\lambda_4$</td>
<td>11.6</td>
</tr>
<tr>
<td>$v_5$</td>
<td>245.9</td>
<td>245.4</td>
<td></td>
<td>$\lambda_5$</td>
<td>6.5</td>
</tr>
<tr>
<td>$v_6$</td>
<td>257.4</td>
<td>257.9</td>
<td></td>
<td>$\lambda_6$</td>
<td>8.0</td>
</tr>
<tr>
<td>$\gamma_{V1}$</td>
<td>1.00</td>
<td></td>
<td>$\kappa_{V1}$</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>$\gamma_{Q1}$</td>
<td>1.00</td>
<td></td>
<td>$\kappa_{Q1}$</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>$\gamma_{V2}$</td>
<td>1.15</td>
<td></td>
<td>$\kappa_{V2}$</td>
<td>1.04</td>
<td>1.03</td>
</tr>
<tr>
<td>$\gamma_{Q2}$</td>
<td>1.53</td>
<td></td>
<td>$\kappa_{Q2}$</td>
<td>2.20</td>
<td>2.05</td>
</tr>
<tr>
<td>$\gamma_{V3}$</td>
<td>1.18</td>
<td></td>
<td>$\kappa_{V3}$</td>
<td>2.08</td>
<td>2.09</td>
</tr>
<tr>
<td>$\gamma_{Q3}$</td>
<td>1.85</td>
<td></td>
<td>$\kappa_{Q3}$</td>
<td>3.99</td>
<td>3.77</td>
</tr>
<tr>
<td>$\gamma_{V4}$</td>
<td>1.23</td>
<td></td>
<td>$\kappa_{V4}$</td>
<td>2.80</td>
<td>2.79</td>
</tr>
<tr>
<td>$\gamma_{Q4}$</td>
<td>1.98</td>
<td></td>
<td>$\kappa_{Q4}$</td>
<td>4.93</td>
<td>4.33</td>
</tr>
<tr>
<td>$\phi$</td>
<td>0.87</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta_{11}$</td>
<td>69.6</td>
<td>69.2</td>
<td></td>
<td>$\delta_{12}$</td>
<td>86.4</td>
</tr>
<tr>
<td>$\delta_{21}$</td>
<td>59.1</td>
<td>49.0</td>
<td></td>
<td>$\delta_{22}$</td>
<td>60.7</td>
</tr>
<tr>
<td>$\delta_{31}$</td>
<td>72.0</td>
<td>70.3</td>
<td></td>
<td>$\delta_{32}$</td>
<td>80.5</td>
</tr>
<tr>
<td>$\delta_{41}$</td>
<td>42.4</td>
<td>33.6</td>
<td></td>
<td>$\delta_{42}$</td>
<td>29.2</td>
</tr>
<tr>
<td>$\delta_{51}$</td>
<td>52.6</td>
<td>43.1</td>
<td></td>
<td>$\delta_{52}$</td>
<td>52.8</td>
</tr>
<tr>
<td>$\delta_{61}$</td>
<td>28.5</td>
<td>25.5</td>
<td></td>
<td>$\delta_{62}$</td>
<td>54.6</td>
</tr>
<tr>
<td>$\delta_{13}$</td>
<td>75.0</td>
<td>63.5</td>
<td></td>
<td>$\delta_{14}$</td>
<td>99.8</td>
</tr>
<tr>
<td>$\delta_{23}$</td>
<td>82.2</td>
<td>78.7</td>
<td></td>
<td>$\delta_{24}$</td>
<td>111.9</td>
</tr>
<tr>
<td>$\delta_{33}$</td>
<td>98.3</td>
<td>83.4</td>
<td></td>
<td>$\delta_{34}$</td>
<td>105.8</td>
</tr>
<tr>
<td>$\delta_{43}$</td>
<td>33.6</td>
<td>25.9</td>
<td></td>
<td>$\delta_{44}$</td>
<td>40.9</td>
</tr>
<tr>
<td>$\delta_{53}$</td>
<td>45.7</td>
<td>44.7</td>
<td></td>
<td>$\delta_{54}$</td>
<td>91.1</td>
</tr>
<tr>
<td>$\delta_{63}$</td>
<td>67.9</td>
<td>65.7</td>
<td></td>
<td>$\delta_{64}$</td>
<td>77.1</td>
</tr>
</tbody>
</table>
Figure 2

Verbal and Quantitative Latent Means