Two studies were conducted to compare the performance instability of children (grades 3-9) labeled learning disabled/brain injured (LD/BI) to the performance instability of emotionally handicapped (EH) children. In the first study, 50 LD/BI and 37 EH students were measured on three third grade reading passages twice, once within one sitting and once with the three passages administered during three consecutive weeks. In the second study, a subsample of 43 students (24 LD/BI and 19 EH) were measured on instructional level reading passages 28 to 47 times within 18 school weeks. On the words correct scores on each administration of the third grade passage reading test and on the word correct and errors scores on the instructional level reading passages, standard errors of measurement (SEMs) were calculated and analyses of covariance were run on the SEMs. Results indicated no difference between the performance instability of the two groups of children. Implications for identifying and treating LD/BI and EH students are discussed.

(Author/CL)
University of Minnesota

Research Report No. 127

VARIABILITY OF PERFORMANCE: A "SIGNATURE" CHARACTERISTIC OF LEARNING DISABLED CHILDREN?

Douglas Fuchs, Lynn S. Fuchs, Gerald Tindal, and Stanley L. Deno

Institute for Research on Learning Disabilities
The Institute for Research on Learning Disabilities is supported by a contract (300-80-0622) with the Office of Special Education, Department of Education, through Title VI-G of Public Law 91-230. Institute investigators are conducting research on the assessment/decision-making/intervention processes as it relates to learning disabled students.

During 1980-1983, Institute research focuses on four major areas:

- Referral
- Identification/Classification
- Intervention Planning and Progress Evaluation
- Outcome Evaluation

Additional information on the Institute's research objectives and activities may be obtained by writing to the Editor at the Institute (see Publications list for address).

The research reported herein was conducted under government sponsorship. Contractors are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent the official position of the Office of Special Education.
Research Report No. 127

VARIABILITY OF PERFORMANCE: A "SIGNATURE" CHARACTERISTIC OF LEARNING DISABLED CHILDREN?

Douglas Fuchs
Clark University

Lynn S. Fuchs
Wheelock College

Gerald Tindal
Pine County Special Education Cooperative

Stanley L. Deno
University of Minnesota
Institute for Research on Learning Disabilities

July, 1983
Abstract

Two studies were conducted to compare the performance instability of children labeled learning disabled/brain injured (LD/BI) to the performance instability of emotionally handicapped (EH) children. In the first study, 50 LD/BI and 37 EH students were measured on three third grade reading passages twice, once within one sitting and once with the three passages administered during three consecutive weeks. In the second study, a subsample of 43 students (24 LD/BI and 19 EH) were measured on instructional level reading passages 28 to 47 times within 18 school weeks. On the words correct scores on each administration of the third grade passage reading test and on the words correct and errors scores on the instructional level reading passages, standard errors of measurement (SEMs) were calculated and analyses of covariance were run on the SEMs. Results indicated no difference between the performance instability of the two groups of children. Implications for identifying and treating LD/BI and EH students are discussed.
Variability of Performance: A "Signature" Characteristic of Learning Disabled Children

The work of Alfred Strauss and his associates was a major catalyst for the LD field (cf. Lerner, 1981). Strauss suggested that a subgroup of EMR children deserved to be viewed as qualitatively different from other EMR children on the basis of their unique perceptual, cognitive, and social behaviors, which, according to Strauss, was caused by brain injury (Strauss & Lehtinen, 1947). Despite the continued influence of the brain injury (BI) perspective, as reflected in both current educational practice and research, it has generated widespread dissatisfaction. During the past two decades, numerous and well-documented alternative conceptualizations and descriptions of learning disabilities have been generated (cf. Smith, 1983). The large number and diversity of these competing definitions have led some to suggest that there are few, if any, salient behavioral characteristics to distinguish LD children from other mildly handicapped children, such as those who are labeled mentally retarded and emotionally disturbed (Hallahan & Kauffman, 1977).

Nevertheless, there are certain descriptions of LD/BI children that have endured from Strauss to the present. One such descriptor is instability of functioning, or a tendency to demonstrate variable performance on the same task from one day to the next. Strauss and Lehtinen (1947), for example, reported dramatic performance instability among students. Similarly, Ebersole, Kephart, and Ebersole (1968) indicated that LD/BI children inconsistently retained material that previously was learned. Recently, Swanson (1982) described this population of pupils as performing in a fragmented,
inconsistent, and variable manner. He proposed that such behavior is acknowledged implicitly in current research into selective attention.

Despite the seeming consistency of these observations and the agreement among teachers about the extent and importance of performance instability (cf. Aviezer & Simpson, 1980), there has been scant empirical investigation reported. In the only identified data-based study, Aviezer and Simpson (1980) tested 40 non-handicapped children and 34 LD/BI children on perceptual, cognitive, and reading tasks three times, with a two-week interval between the first and second testings and with a two-month interval between the second and third administrations. Instability was operationalized as a decrease in score with task repetition; stability was defined as performance constancy or improvement. Young children, regardless of their subgroup identification, tended to be unstable in performance, whereas differences were found in stability as a function of subgroup for older students (LD/BI 10 and 11-year olds displayed more unstable performance than their non-handicapped peers).

Thus, Aviezer and Simpson (1980) provided some support for the notion that at least older LD/BI children may be characterized by performance instability. However, two limitations of this study are noteworthy: first, the questionable meaningfulness of the definitions of instability/stability; second, the absence of a comparison between the instability of LD/BI children and the instability of children constituting different categories of exceptionality. Because of this second study limitation, it remains unclear whether performance instability characterizes only LD/BI children or is associated with
additional groups of children experiencing serious learning problems and academic retardation. Such ambiguity necessarily precludes the usefulness of performance instability in certain clinical endeavors, such as in conducting differential diagnoses, and in theoretical conceptualizations and research.

The purpose of the present series of interrelated investigations was to explore the legitimacy of the belief that performance instability is a distinguishing feature of LD/BI pupils. Toward this end, these studies employed (a) two groups of older handicapped children, LD/BI and emotionally handicapped (EH), and (b) a more appropriate index of stability, the standard error of measurement across repeated administrations of similar tasks.

**Study 1**

In Study 1, reading performance instability was compared for EH and LD/BI pupils under two conditions. In the first condition, instability across three reading measurements, which occurred at one-week intervals, was examined. In the second condition, instability across three reading measurements occurring within one 8-minute session was explored.

**Method**

**Subjects.** The sample included 87 pupils from seven New York City public schools. Fifty subjects (57%) were classified EH, and 37 students (43%) were labeled LD/BI. EH students (37 M, 13 F) read an average 2.59 years below grade level (SD = 2.01), and had spent an average 2.89 years in special education classes (SD = 1.81). The numbers of EH students placed in grades 4-9, respectively, were 10, 9,
In the LD/BI group (24 M, 13 F), students read an average 3.51 years below grade level (SD = 1.69), and had spent an average 2.71 years in special education classes (SD = 1.35). The numbers of LD/BI children placed in grades 3-8, respectively were 1, 6, 10, 9, 6; and 5.

Statistical tests revealed that EH and LD/BI groups were similar with respect to sex, $\chi^2(1) = 1.40$, ns, grade level placement, $\chi^2(6) = 5.68$, ns, and years spent in special education, $t(85) = .51$, ns. However, the LD/BI students were farther below grade level in reading, $t(85) = 2.25$, p<.025.

Measures. The Passage Reading Test (PRT) was employed, which is comprised of three reading passages from a third grade book of the Ginn 720 series (1976). Two passages were sampled randomly from the text and one was chosen to represent the readability of the last 25% of the book. (See Fuchs, Fuchs, & Deno, 1982, for selection procedure.) The PRT requires students to read aloud from each passage for one minute, and student performance is reported in terms of the number of correct words read. Test-retest reliability coefficients ranged from .93 to .96 (Fuchs, Deno, & Marston, in press). Concurrent validity coefficients with respect to the Woodcock Reading Mastery Tests, Word Identification and Passage Comprehension Tests, ranged from .89 to .92 (Fuchs, 1981). Internal consistency reliability (Cronbach's alpha) for the three passage test was .79 (Fuchs, Deno, & Mirkin, 1982).

Procedure. The PRT was administered twice to each student, with four school months intervening between administrations. During the
first administration (Time 1), students read from a different passage each week, for three consecutive weeks. During the second administration (Time 2), students read from the three passages on one day. The tests were administered individually, with total testing time for each of the two administrations ranging from 4 to 8 minutes. All examiners were trained in standard administration procedures (see Mirkin, Deno, Fuchs, Wesson, Tindal, Marston, & Kuehne, 1981).

Data analysis. Because initial data indicated that the LD/BI students were poorer readers than the EH children, preliminary t tests were performed on the average number of correct words across the three passages on each administration of the PRT. On both administrations, there was a statistically significant difference favoring the reading performance of the EH group; \( t(85) = 2.04, p < .05 \) and \( t(85) = 3.65, p < .001 \), at Time 1 and Time 2, respectively. On the first administration, EH children read correctly a mean 67.12 words (SD = 56.42) and LD/BI pupils read correctly a mean 45.45 words (SD = 36.75); on the second administration, the means for the number of words read correctly were 67.17 (SD = 46.33) and 33.71 (SD = 23.95) for the EH and LD/BI groups, respectively.

Since the EH average performance on the PRT was higher than that of the LD/BI pupils, there was a greater range of behavior and more potential instability for the LD students. In order to control for this source of error in the comparison of performance instability, two-way analyses of covariance (ANCOVAs) were run. In each analysis, the standard error of measurement across the three-passage test, an index of intra-individual variability, was the dependent variable.
Student classification was the independent variable, with sex employed as a blocking factor. The average number of words read correctly across the three-passage tests was entered as the covariate.

**Results and Discussion**

Table 1 shows the means, adjusted means, and the ANCOVA results for the standard error of estimate on the first and second administrations of the PRT. The two-way ANCOVAs revealed no statistically significant effect for the independent factor, Student Classification, at either administration. Additionally, there was no statistically significant effect for the blocking factor, Sex, $F(1.81) = 1.32, \text{ns}$ and $F(1.81) = .85, \text{ns}$, at Times 1 and 2, respectively, or for the interaction between Student Classification and Sex, $F(1.81) = .34, \text{ns}$ and $F(1.81) = .00, \text{ns}$, at Times 1 and 2, respectively.

Therefore, there was no reliable difference in the intra-individual reading performance variability between the EH and LD/BI children. This finding remained unchanged when instability across week-long intervals (first administration of the PRT) or within a relatively short time frame (second PRT administration) was examined. Additionally, there appeared to be little, if any, practical significance in the instability of these groups of children; adjusted mean standard error of estimates for the PRT Time 1 and Time 2 administrations, respectively, were 1.30 and 1.60 words correct.
Study 2

While Study 1 examined the intra-individual instability across three measurements on third grade passages administered at different time intervals (at one-week intervals or within 8 minutes), Study 2 explored intra-individual instability across many more measurement points (28 to 47) on an instructional level oral passage reading task administered at intervals of one day to several school days. The purpose of this operational replicate (Borg & Gall, 1979) of the first study was to examine whether potential artifacts of the experimental procedure, specifically either (a) the relatively few data points collected, or (b) the use of a third grade passage test rather than instructional level material, might have accounted for the findings of no difference.

Method

Subjects. A randomly selected subsample of 43 children from Study 1, representing the EH and LD/BI categories in nearly the same percentages (56% EH and 44% LD/BI), served as subjects in Study 2. EH students (18 M, 6 F) read an average 3.14 years below grade level (SD = 2.04) and had spent an average 2.61 years in special education classes (SD = 1.97). The numbers of EH students placed in grades 4-9, respectively, were 4, 4, 5, 5, 5, and 1. LD/BI (14 M, 5 F) students read an average 4.26 years below grade level (SD = 1.32), and had spent an average 2.66 years in special education (SD = 1.31). The numbers of LD/BI children placed in grades 3-8, respectively, were 1, 5, 6, 4, and 3.

Statistical tests revealed that EH and LD/BI groups were similar with respect to sex, $\chi^2(1) = .01$, ns, grade level placement, $\chi^2(5) =$
5.03, ns, and years spent in special education, $t(41) = .05$, ns. However, the reading achievement of the LD/BI students was farther below grade level than the reading achievement of the EH pupils, $t(41) = 2.07$, $p < .05$.

Procedure. For 18 school weeks, subjects' teachers implemented ongoing curriculum-based measurement. Teachers selected for each child an instructional level book and, at least twice weekly, randomly chose a passage from that book on which to measure the child's performance. Employing a standard administration procedure (see Mirkin et al., 1981), teachers required the student to read orally for one minute, and marked errors including omissions, substitutions, insertions, and mispronunciations. Teachers then graphed the numbers of words correct and errors read. Teachers were trained to design and implement this measurement system in weekly individual meetings with teacher trainers. Within the 18-week implementation, each teacher was observed three times while measuring a student. The observers rated the accuracy with which the teachers (a) administered the measurement task, and (b) graphed the data, each on a 5-point Likert-type scale. The ratings, averaged across teachers and across the three observations, for task administration and data graphing, respectively, were 4.80, and 4.21 (Fuchs, Deno, & Mirkin, 1982).

Data analysis. Because preliminary data suggested a discrepancy between the reading skills of EH and LD/BI children, preliminary t tests were conducted on the numbers of words correct and errors averaged across each student's graphed data. EH students read more words correctly, $t(41) = 2.58$, $p < .02$, with the average scores for...
the EH and LD/BI groups, respectively, 60.23 (SD = 44.49) and 30.78 (SD = 26.57). EH children also made fewer errors, \( t(41) = -4.06; p < .001 \), with the mean scores for the EH and LD/BI groups, respectively, 6.05 (SD = 4.43) and 11.65 (SD = 4.57).

Given the greater range of behavior and potential instability for the EH students on words correct and for the LD/BI pupils on errors, two two-way ANCOVAs were employed: one analysis for word correct scores and one for error scores. In each analysis, the standard error of measurement (on either words correct or errors) across each student's 28 to 47 graphed data points was the dependent variable. Student classification was the independent variable, with sex employed as a blocking factor. The average number of words correct across the graphed data was entered as the covariate when the standard error of measurement on words correct was employed as the dependent variable; the average number of errors was entered as the covariate when the standard error of measurement on errors was used as the dependent variable.

Results and Discussion

Table 2 displays the means, adjusted means, and the ANCOVA results for the differences in standard error of measurement on words correct and errors. As shown, there was no statistically significant difference in the standard error of measurement for the independent factor, Student Classification, on either words correct or errors. Additionally, there was no statistically significant effect for Sex as the blocking factors, \( F(1,37) = .41, \text{ ns} \) and \( F(1,37) = .18, \text{ ns} \), on words correct and errors, respectively, or for the interaction between
Student Classification and Sex, $F(1,37) = 1.60, \text{ns}$ and $F(1,37) = .00, \text{ns}$, on words correct and errors, respectively.

Consequently, as in Study 1, there were no reliable differences in the intra-individual instability between EH and LD/BI children. Additionally, there appeared to be no practical difference in the instability of these two groups of children; adjusted mean standard error of estimates for the two classes differed by only .60 words correct and .02 errors. Further, this operational replicate of Study 1 indicates that findings of the two studies are not artifacts of specific methodological procedures employed, but rather appear to generalize across the number of data points collected as well as the level of reading material employed.

**General Discussion**

The purpose of the present studies was to compare the performance instability of two handicapped groups of children—LD/BI and EH students. Results indicated that there were no reliable differences between the two groups. Further, there appeared to be little practical difference in their performance instability. This finding was robust; the same question was explored in three different ways: (a) across three measurements on third grade reading material over three consecutive weeks, (b) across three measurements on third grade reading material within one session, and (c) across 28 to 47 measurements on instructional level reading material over 18 school
weeks. In each case, results indicated no difference in the performance instability of these two groups of handicapped students.

Unlike the investigation by Aviezer and Simpson (1980), the present studies did not compare the performance instability of LD/BI children to non-handicapped children; it remains quite possible that, in relation to their non-handicapped peers, LD/BI students exhibit greater variability in their school functioning. However, findings from the present studies do not support the use of performance instability as a "signature" or distinguishing characteristic of LD/BI children. Results indicate that EH children display this pattern of behavior to a similar degree. Moreover, there appears to be no reason to preclude the possibility that children in other categories of exceptionality also may manifest similar variability of performance.

There appear to be at least two practical implications from these findings. First, although LD/BI children historically have been described partly in terms of the variability of their performance, diagnosticians should not presume that this characteristic may serve to differentiate reliably an LD/BI subgroup from other groups of handicapped children. Rather, findings are consonant with (a) Hallahan and Kauffman's (1977) view that there is an absence of characteristics with which to make differential diagnoses among the mildly-to-moderately-handicapping conditions, and (b) the notion that special educators should identify handicapped children within a non-categorical framework. Second, results indicate that recent and promising instructional strategies, such as metacognitive training
(cf. Loper, 1982), that have been used with LD/BI children also might prove effective with EH and perhaps other handicapped pupils.
References


Ebersole, M., Kephart, N. C., & Ebersole, J. B. Steps to achievement for the slow learner. Columbus, OH: Charles E. Merrill, 1968.


Fuchs, L. S., Deno, S. L., & Marston, D. Improving the reliability of curriculum-based measures of academic achievement for psychoeducational decision making. *Diagnostique*, in press.


Swanson, H. L. In the beginning was a strategy: Or was it a constraint? *Topics in Learning and Learning Disabilities*, 1982, 2(2), x-xiv.
Table 1

Standard Error of Measurement on Two Administrations of the PRT

<table>
<thead>
<tr>
<th>Student Classification</th>
<th>EH (N=50)</th>
<th>BI (N=37)</th>
<th>F test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measure</td>
<td>X</td>
<td>Adjusted X ≡</td>
<td>X</td>
</tr>
<tr>
<td>PRT, Time 1</td>
<td>3.80</td>
<td>3.65</td>
<td>4.75</td>
</tr>
<tr>
<td>PRT, Time 2</td>
<td>2.14</td>
<td>1.92</td>
<td>1.68</td>
</tr>
</tbody>
</table>

a Adjusted for covariate and blocking factor.
Table 2

Standard Error of Measurement on Students' Graphed Data

<table>
<thead>
<tr>
<th>Student Classification</th>
<th>EH(N=24)</th>
<th>BI(N=19)</th>
<th>F test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measure</td>
<td>X</td>
<td>Adjusted X^a</td>
<td>X</td>
</tr>
<tr>
<td>Words correct</td>
<td>10.95</td>
<td>10.66</td>
<td>9.61</td>
</tr>
<tr>
<td>Errors</td>
<td>2.01</td>
<td>2.25</td>
<td>2.53</td>
</tr>
</tbody>
</table>

^a Adjusted for covariate and blocking factor
PUBLICATIONS

Institute for Research on Learning Disabilities
University of Minnesota

The Institute is not funded for the distribution of its publications. Publications may be obtained for $4.00 each; a fee designed to cover printing and postage costs. Only checks and money orders payable to the University of Minnesota can be accepted. All orders must be pre-paid. Requests should be directed to: Editor, IRLD, 350 Elliott Hall; 75 East River Road, University of Minnesota, Minneapolis, MN 55455.

The publications listed here are only those that have been prepared since 1982. For a complete, annotated list of all IRLD publications, write to the Editor.


Graden, J.; Thurlow, M., & Ysseldyke, J. Instructional ecology and academic responding time for students at three levels of teacher-perceived behavioral competence (Research Report No. 73). April, 1982.


Thurlow, M. L., Ysseldyke, J. E., Graden, J., Greener, J. W., & Mecklenberg, C. Academic responding time for LD students receiving different levels of special education services (Research Report No. 78). June, 1982.


Fuchs, D., Dailey, Ann Madsen, & Fuchs, L. S. Examiner familiarity and the relation between qualitative and quantitative indices of expressive language (Research Report No. 83). (July, 1982.)


