INSTITUTE FOR RESEARCH IN LEARNING DISABILITIES

The University of Kansas
Lawrence, Kansas, 66045

Emphasis on Adolescents and Young Adults

A MULTI-TRAIT, MULTI-METHOD ANALYSIS OF

THE BAYESIAN SCREENING INSTRUMENT AND

TEST BATTERY FOR LD ADOLESCENTS

Gordon R. Alley, Donald D. Deshler,

Daryl F. Mellard, and Michael M. Warner

Co-Directors: Edward L. Meyen
               Richard L. Schiefelbusch

Research Coordinator: Donald D. Deshler

Associate Coordinator: Jean B. Schumaker

Institute for Research in Learning Disabilities
· The University of Kansas
313 Carruth-O'Leary Hall
Lawrence, Kansas 66045

Cooperating Agencies

Were it not for the cooperation of many agencies in the public
and private sector, the research efforts of The University of Kansas
Institute for Research in Learning Disabilities could not be con-
ducted. The Institute has maintained an on-going dialogue with
participating school districts and agencies to give focus to the
research questions and issues that we address as an Institute. We
see this dialogue as a means of reducing the gap between research
and practice. This communication also allows us to design procedures
that: (a) protect the LD adolescent or young adult, (b) disrupt the
on-going program as little as possible, and (c) provide appropriate
research data.

The majority of our research to this time has been conducted in
public school settings in both Kansas and Missouri. School districts
in Kansas which are participating in various studies include: United
School District (USD) 384, Blue Valley; USD 500, Kansas City; USD
469, Lansing; USD 497, Lawrence; USD 453, Leavenworth; USD 233, Olathe;
USD 305, Salina; USD 450, Shawnee Heights; USD 512, Shawnee Mission,
USD 464, Tonganoxie; USD 202, Turner; and USD 501, Topeka. Studies
are also being conducted in Center School District and the New School
for Human Education, Kansas City, Missouri;. the School District of St.
Joseph, St. Joseph, Missouri; Delta County, Colorado School District;
Montrose County, Colorado School District; Elkhart Community Schools,
Elkhart, Indiana; and Beaverton School District, Beaverton, Oregon.
Many Child Service Demonstration Centers throughout the country have
also contributed to our efforts.

Agencies currently participating in research in the juvenile
justice system are the Overland Park, Kansas Youth Diversion Project
and the Douglas, Johnson, and Leavenworth County, Kansas Juvenile
Courts. Other agencies have participated in out-of-school studies--
Achievement Place and Penn House of Lawrence, Kansas, Kansas State
Industrial Reformatory, Hutchinson, Kansas; the U.S. Military; and
the Job Corps. Numerous employers in the public and private sector
have also aided us with studies in employment.

While the agencies mentioned above allowed us to contact
individuals and supported our efforts, the cooperation of those
individuals--LD adolescents and young adults; parents; professionals
in education, the criminal justice system, the business community,
and the military--have provided the valuable data for our research.
This information will assist us in our research endeavors that have
the potential of yielding greatest payoff for interventions with the
LD adolescent and young adult.

Abstract

Three related studies were designed to address some key issues confronting the learning disability field concerning the identification of learning disabled adolescents. The first study (Research Report No. 9) addressed the question of which group(s) of professionals or parents make the most homogeneous identification decisions on learning disabilities' criteria. In the second study, (Research Report No. 10) the temporal and interscorer reliability as well as the construct and content validity of the Modified Component Disability Instrument was investigated. The reliability and validity of the Modified Component Disability Checklist and Secondary Test battery were investigated in the third study (Research Report No. 11).

The first study included a statewide random sampling of seven groups of professional educators and a group of parents of LD students. These eight groups were compared for their degree of agreement on the component disability survey instrument. The results indicated that no one group had greater consensus than any other. The conclusion was that LD teachers were an appropriate group from which to obtain likelihood ratios to be used in obtaining posterior probabilities for the LD population.

In the second study the professionals who had responded in study 1 re-estimated the probabilities they had provided 14 days earlier. This provided a measure of temporal reliability of the items. In addition, a new sample was drawn from two (Speech clinicians and LD teachers) of the seven professional groups to cross-validate the initial results. The temporal reliability coefficient obtained for individual items was sufficiently high to suggest the reliability of the judgments. Secondly, no differences were found among 41 component disability estimates between the two independent samples of professionals. As a part of this second study the survey was subjected to a factor analysis. The logical clusters of component disabilities were found to be substantiated as statistical factors.

As a part of Study 3, (Research Report No. 11) a group of seven professionals in LD found the behaviors associated with the component disabilities of the survey generally to be: (a) important, (b) grade appropriate, and (c) accessible to the teachers' observations in the classroom. The conclusion was that the Modified Component Disability Checklist is a reasonably reliable screening measure, especially at grades 8-12. In the third study a group of learning disabled adolescents and a group of low achieving peers were administered both the classroom screening measure and the battery of pre-selected tests. A multi-trait, multi-method analysis was completed. The results show a trend toward the reliable and valid nature of these two screening methods.

A MULTI-TRAIT, MULTI-METHOD ANALYSIS
OF THE BAYESIAN SCREENING INSTRUMENT AND TEST
BATTERY FOR LD ADOLESCENTS

Perhaps the most pressing need in the learning disability (LD) field is that of defining the population. The confusion that has existed in education as a result of poor definitional direction for the LD population is well documented (Wissink, Kass, & Ferrell, 1975; Chalfant & King, 1975; Larsen, 1978). Progress in educational programming, research, and intervention development is contingent upon resolution of the definitional issue. A major focus of this Institute is to address those concerns that relate to the identification of the LD adolescent population.

The research outlined here is a series of studies that were designed to address some of the major questions related to identifying characteristics of the population and to reliable, valid idenfification procedures. While several of the hypotheses and questions in these studies related to previous work done at The University of Kansas using Bayesian aggregate procedures (Alley, Deshler, & Warner 1979); these three studies addressed issues beyond that specific procedure.

1. Do members of a professional group agree on the identifying characteristics of the LD adolescent?

2. Is any one professional group more homogeneous than others and consequently more consistent in their identification?

3. Are the subjective judgment decisions of child care agents reliable and valid?

4. Can regular classroom teachers reliably observe content valid behaviors in students that are indicative of LD?

1

5.   Is a test battery additionally useful in making an identification of LD?

This study concerned the reliability and validity of the secondary LD checklist (Checklist) and the secondary school level battery of tests (Battery). Alley et al. task analyzed the four component disabilities of the Best Discriminator cluster and three component disabilities of the High Frequency cluster. The task analysis was completed using the regular language arts classroom as the setting from which to draw the target behaviors. Secondary LD teachers assisted in providing informal content validity of the behaviors. The development of the checklist is described in the Alley et al. article.

Each test in the Battery was selected by Alley and Deshler to correspond with one of the component disabilities of the Best Discriminator cluster and the nine target behaviors. These tests were selected based on the clinical experiences of the two investigators and the reliability and validity of the measure.

The Checklist and the Battery were field tested during the 1977-1978 school year by the staff of the Lawrence Child Service Demonstration Center (CSDC). This evaluation provided clinical validity to the Checklist and the Battery, but statistical study of the two assessments forms had not been completed.

It was the purpose of this study to statistically investigate the reliability and validity of the Checklist and Battery. Two interrelated phases were included in the study. First, the content validity of the checklist was systematically studied. Second, the temporal reliability of the Checklist, the intercorrelations among the component disabilities for both the Checklist and the Battery,

and the convergent and divergent validity of the two forms of assessment were studied.

Phase I

Methodology

Subjects. A group of five secondary LD teachers with at least two years of teaching experience with LD students participated in this study. All teachers were students in doctoral programs at The University of Kansas. These teachers had specialized in secondary LD at both the master's and doctoral levels. They volunteered to provide judgments for studying the content validity of the Checklist.

Setting. The LD teachers independently judged the content validity of the Checklist behaviors as they judged them to be measures of each of the four Best Discriminators and four High Frequency component disabilities. This activity took place in February, 1979. The Rating Scale (Appendix A) was collected from each teacher by one of the investigators. A second session, held one week later, was conducted to obtain consensus on the items during group discussion and to review the teachers' judgments of the Rating Scale.

Measurement. A simple matching task was constructed for the initial judgments of the LD teachers on condition 1. The target behaviors were listed on the left hand side of the page. On the right hand side of the page was a blank. The target behaviors were randomly ordered. The task of the teachers was to match the component disability to the target behavior.

The second measurement for condition 2 was the Rating Scale.

3

The Rating Scale was developed to measure the judgments of the five LD teachers. It was a modification of a procedure described by Meyen (1968) and Wimmer (1979). The teachers were told to judge each target behavior on the basis of five criteria. They were: (a) importance, (b) frequency of behavior for LD students, (c) frequency of the behavior for LD students, (d) accessibility of the behavior, and (e) grade level when the behavior would initially interfere with classroom performance (See Appendix A).

The "importance" criterion was measured on a seven-point Likert-type scale which incorporated two terms: "Not Important" ranging from 1 to 3 and "Important" ranging from 5 to 7. The medial position, 4, was used as an intermediate quality of importance. The "Frequency" criterion, also a seven-point scale, was similar to the "Importance" scale. The terms "low LD/low non-LD" and "high LD/high non-LD" were the frame of reference terms to be used by the five LD teachers for their judgments. The accessibility scale used seven points and the frame of reference terms were "easy" and "difficult". Finally, the teachers were instructed to circle the grade level from "Pre-K" to "Post-High". No decimal scores were used in the analysis. Scoring of the ratings was done independently by two of the investigators. There was 100% agreement between the investigators' scoring.

Procedures. The instructions to the LD teachers were verbal under both content validity conditions. Under the initial adminis-tration of the matching task (condition 1), the LD teachers were told to write in the Component Disability appropriate to the target behav-ior. The teachers were asked if they had any questions and were re-quested to contact the investigators if they had questions. No con-

4

tacts were made with the investigators after the initial instruction session. Each teacher completed the task independently. One of the investigators collected the worksheets from each of the LD teachers on the four days following the instructional session. A criterion of 80% consensus was applied to the data. If the criterion was met the target behavior was determined to be content related to the component disabilities.

Under condition 2, the LD teachers were told to judge each component disability item for its (a) importance, (b) frequency of occurrence in LD students, (c) frequency of occurrence in non-LD students, (d) accessibility to observation in a classroom setting, and (e) the grade level at which the behavior become important when differentiating LD from non-LD students. No criterion was set for any of the four variables. A description of the judgments of the LD teachers was the measure of interest.

Results

Condition 1. Of the 29 target behaviors, 19 reached the criterion. All 19 target behaviors were content related to the component disabilities as they were task analyzed by Alley and Deshler (Alley et al., 1979) (See Table 1).

The five teachers met together as a group to discuss and determine if they could reach criterion on the 10 remaining target behaviors. After changing the content of seven of the 10 target behaviors, the group met criterion on all 10 remaining behaviors (See Table 1). One target behavior, "unable to allot a reasonable amount of study time", i.e., does not use study time appropriately in class, was originally considered as a part of the Test-taking Skills component disability.

5

After modifying the content of this item, the group consensus was that the behavior was related to the Study Skills component disability.

Condition 2. The LD teachers judged the content of all 29 items to be "important". They judged all target behaviors to be "high frequency" disabilities among secondary LD students. The teachers reported that a high frequency of non-LD students do not "survey material before studying in depth" or "scan for major points before reading material intensively" (Item 10). Seven items, two of the Best Differentiating component disabilities and five of the High Frequency component disabilities were judged to be found in moderate frequency among non-LD adolescents. Twenty-one of the behaviors were judged to be found in low frequency among the non-LD students.

The judges felt that four behaviors were difficult to observe in a classroom setting. One item, "unable to recognize words that are common across content areas, i.e., 'knowledge', 'suggestion', selected', etc." (Item 9) was associated with the Best Differentiating component disabilities. Nine items were found to be in the medial position. The remaining 16 items were found to be easily accessible to observation in the classroom.

Finally, the grade at which the target behavior would seriously affect school performance provided the widest variance on the measure (See Table 2). Nine of the 29 items were judged to initially create severe difficulties at the first grade level by at least one teacher. Conversely, at least one teacher judged 11 other items to initially create severe difficulties at the ninth grade level. The range of judgments was found to be from three to eight grade levels. Four items, one item on the Test-Taking Skills component

6

disability and three items on the Study Skills component disability
were judged by a majority of teachers to initially cause problems
for secondary students in the junior high school (grades seven to
nine).

Discussion

Condition 1. The five secondary LD teachers judged the 29
target behaviors on the Checklist to be subcomponents of the eight
component disabilities (listed in Table 1). The judges agreed with
the investigators' task analysis on 19 of the items. These 19 be-
haviors remained unchanged in content or structure by the five sec-
ondary LD teachers. The judges met for a second session to discuss
and obtain consensus on the remaining 10 items. Agreement was reach-
ed which resulted in changing the wording of the target behavior item
or a change of judgment. One item was changed in its content and
reassigned by the teachers from the component to which it was assigned
by the investigators to another component disability.

On the basis of these findings, one-half of the original Check-
list items obtained content validity. The checklist was then modified
(Modified Checklist) to account for the teachers' judgments (See
Appendix B). The Modified Checklist was used in Condition 2 of
Phase 1 as well as Phase 2 of this study.

Condition 2. The Modified Checklist contains behaviors that
are judged to be important to teachers and found in high frequency
among secondary LD students. The behaviors are generally found in
low frequency among non-LD students. On the whole the behaviors
are accessible to observation in a classroom and can seriously affect
a student's performance before entering seventh grade. It appears to

be an adequate screening measure in grades eight through twelve.

However, the Modified Checklist should be further modified to remediate its content validity limitations. First, item 10 was judged to occur with approximately equal frequency in the LD and non-LD populations. This behavior should be deleted from the Study Skills component disabilities especially when one considers that it was judged to be difficult to observe in the classroom and the majority of judges felt the behavior initially impeded performance after the student reached the seventh grade. Second, items on the Test Taking component disability should be studied in more detail because of either the difficulty a teacher may have in observing the behavior in the classroom or the behavior's effect on classroom performance may not seriously affect classroom performance in seventh grade. Two behaviors on the Study Skills component disability, items 1 and 20, may not seriously affect a student's performance in seventh grade.

Based on the findings and conclusion of this phase of the study, question 3 can be answered affirmatively. Teachers report that they can observe target behaviors in their classroom that are indicative of learning disabilities.

This phase of the study is limited by the small number of judges. It should be cross-validated with a larger sample. However, the sample size used in Phase 1 of this study is of sufficient size to provide meaningful results (Angoff, 1971).

Phase 2

Methodology

Subjects. The subjects for Phase 2 were 21 low-achieving students and 21 LD students in Grades 7, 8, and 9. The low-achieving students

achieved below the 25 percentile on the SRA Achievement Test Composite
score, received at least one failing grade in the most recent grading
period, and had never been staffed by special education. The LD group
was classified as learning disabled by a school-based team and evidenced
behaviors on three or four of the four Best Discriminating component
disabilities.

Setting. Three junior high schools from Lawrence School District
were selected for this phase. It was planned to include students from
the senior high school of the same district but this population could
not be included in the study because of scheduling conflicts with the
school calendar.

Measurement. Two measurement methods were used in the analysis.
The first method was the Modified Checklist of Learning Difficulties
(Secondary). The checklist was described in Phase 1 of this study.
The Modified Checklist used here was the version modified from the
results of Phase 1 of this study. It contains 29 items; each item
is a behavior related to one of the Best Discriminating or to the four
High Frequency component disabilities. A Modified Checklist was complet-
ed by the regular class language arts teacher for each student he/she
believed had a learning problem. The teacher compared the student's
behavior to each behavior on the Modified Checklist and "checked"
if it was present, absent or unknown/no information.

The Modified Checklist is a screening measure which has been
used to refer secondary students with suspected learning disabilities.
The developers of both the Modified Checklist and its preceding edi-
tion have required that the student must be judged to evidence behav-
iors associated with three of the four Best Discriminators and all

9

four of the High Frequency component disabilities.

A posterior probability equal to or greater than .86 has been used as the cut off criterion for referring students for learning disabilities. The probability of evidencing behaviors associated with three of the four Best Discriminators is affected by two conditions. First, depending upon which set of the three component disabilities is identified with the student, a posterior probability will range from .77 to .86. However, if the High Frequency component disability likelihood ratios are included when computing the posterior probability, then the posterior probability is .99. This posterior probability of .99 is met regardless of which three Best Discriminator component disabilities are observed.

The regular classroom teacher judges: (a) whether or not the student displays the behaviors identified as being the Best Discriminators of LD and non-LD or (b) that he/she does not know if the student displays these behaviors. This second option (b) was given because previous experiences have shown that some behaviors, such as a student's math skills, are not observed in all classes. Therefore the teacher lacks the information for making such a judgment. In such cases, the probability estimates of being learning disabled need to be revised.

As a simple example, consider a case in which three of the four Best Discriminating components are checked as being observed. This student has the assigned probability of .86 of being learning disabled and is referred for further evaluation.

The investigators chose to use the first and simpler scoring system to refer a student if he/she evidenced three of the four Best Discriminating component disabilities and all four of the High

Frequency component disabilities as judged by the regular teacher. This system was selected for its ease of calculation and because the Modified Checklist is used only for screening purposes and not identification of LD. However, the consumer should be aware of this departure from Bayesian statistical procedure.

The Test Battery for Learning Problems (secondary) (Battery) (Alley et al., 1979) was used as the refinement phase when screening for LD secondary students. The Battery is composed of subtests from five achievement measures. These subtests were selected by Alley and Deshler to measure the four Best Discriminating component disabilities with emphasis on the original Checklist behaviors associated with these components. One subtest was included for the mathematics and spelling component disabilities. Because of the nature of the reading tasks, two subtests were included to measure reading recognition and also decoding words. For purposes of this study, one subtest was used to measure each of the four component disabilities. The tests and selected subtests were:

1.   Woodcock Reading Mastery Test, Word Attack--to measure decoding skills

2.   Stanford Achievement Test, Vocabulary--to measure word recognition skills (Advanced Level, Form A, odd numbered items)

3.   Ross Tests of Higher Cognitive Processes, Relevant and Irrelevant Information--to measure knowledge of mathematical algorithms (odd numbered items)

4.   Stanford Achievement Test, Spelling--to measure monitoring of errors

All subtests were scored using raw scores obtained by the student.

Procedures. All students of three junior high schools in one school district were screened using the Modified Checklist. The LD teachers in each junior high school agreed to coordinate the screening.

The language arts teacher of the junior high schools completed the Modified Checklist. This screening served two purposes: (a) gathering data for this investigation and (b) to refer students as part of the school district's screening program.

The teachers were told of the joint effort and permitted to take part in either the investigation or to only comply with the district's programming needs. One LD teacher chose not to volunteer for selected parts of the investigation. This teacher was the only LD teacher of one junior high school. The investigators chose to include only the two junior high schools from which complete temporal reliability data was available on the Modified Checklist.

The teachers were provided specific instructions on the administration of the Modified Checklist. The Modified Checklist was completed by the language arts teachers during a half-day in-service session. The completed checklists were collected by the building LD teacher(s). These LD teachers notified one of the investigators when all data was completed. The elapsed time period ranged from two to four weeks between the time the checklists were distributed to the LD teachers and the date the investigator was notified that all of the data had been collected.

The checklists were scored and achievement information was obtained from student folders by the LD teacher(s). A second investigator

scored the checklists to insure that no calculation error(s) had oc-
curred. The students were then classified into two groups. One group
(the low achievers) manifested problems in no more than two Best
Discriminator component disabilities and performed within the first
quartile of the SRA Achievement Test. The second group (LD) which
included those who had manifested problems in three or four Best
Discriminating component disabilities, had been identified as LD by
a multidisciplinary team in the school and were currently receiving
intervention services in a resource room for LD students. The names
were coded by number at the school so that no member outside the
school district could identify the student subjects. There were 23
students comprising the low-achieving group and 21 students in the
LD group.

Two weeks after the in-service day, a new package of checklists
was distributed to the LD teachers of the two schools. The LD teachers
were told of the need for temporal reliability of the Modified Check-
list. The same directions were given to the language arts teachers,
except that the teachers were to judge only the behaviors of the low-
achieving and LD group members in their class(es). The regular class
teachers completed the checklist in two weeks. One of the investi-
gators collected the checklists from the LD teachers at the end of the
two-week period.

Concurrent to the temporal reliability data gathering period, the
LD teachers administered the Battery to the students identified in the
two groups. The students were told that their performance was important
to a new program for students in junior high schools the next academic
year. All names were coded to coincide with the checklist data.

13

It was planned to obtain temporal reliability on the Battery with these groups. However, the school year was closing and it was the consensus of the school district personnel and the investigators to delete this portion of the investigation. The data gathering was completed by May 23, 1979.

Research Design. A multi-trait multi-method design (Campbell & Fiske, 1959) was used to analyze the data. Using this design, the component disabilities were considered as four traits and the Modified Checklist and the Battery were considered as methods. This design permitted the investigators to effectively study the two- to four-week temporal reliability and intercorrelations among the component disabilities on the checklist. The missing data did not permit study of the temporal reliability of the Battery but the intercorrelations among the component disabilities was analyzed. In addition, the convergent and divergent validity of the methods was studied.

Results

Forty-two low-achieving and LD junior high school students were used to study the reliability and validity of the Modified Checklist and the Battery. Table 3 displays the data using the multi-trait, multi-method design.

Three steps were involved in the analysis, each one becoming more stringent. The first step viewed the correlation coefficients along the diagonal of the heavy line triangles. These correlations are the two- to four-week temporal reliabilities of the component disabilities as measured by the Modified Checklist. The general results are satisfactory, but a notably low correlation exists between the two administrations of the Checklist as it measures "difficulty with mathe-

14

matics algorithms." As described, the temporal reliabilities on the Battery were not obtained. This missing data is denoted by dashes (-).

The first test in the convergent validity of the components is on each of the methods, i.e., hetero-trait, mono-method. The inter-correlations among the component disabilities on the Modified Check-list are in the expected direction and of the magnitude expected (Box A). A notable exception is the moderate relationship between the two component disabilities--Decoding Words and Reading Recognition ($r = .51$). The inter-correlations among the component disabilities on the Battery are generally of the expected magnitude and in the expected direction (Box B). However, the moderate intercorrelation between the component disabilities (Difficulty with Mathematical Algorithms and Monitoring Spelling Errors) was not expected ($r = .43$).

The second test is the relationship of the same component disabilities on the two methods, i.e., mono-trait, hetero-method. This is a test of convergent validity, i.e., do the two methods measure the same trait? This data is presented on the diagonal between the two broken line triangles. These correlations are of lesser magnitude than those of the first test. They should be of moderate magnitude and in a positive direction. The results are of lesser magnitude than would be expected. The correlation between the two methods as they measure the component disability, Decoding Words, is very low ($r = -.03$) and suggests that only minimal relationship exists between the two methods.

The third and final test is of the validity of the component disabilities. These correlation coefficients are hetero-traits on hetero-methods. The broken line triangles provide the divergent validity of the components. The expected magnitude among the com-

ponents should be small and near zero. The results of this test
are mixed. The same type of finding occurred between Difficulty in
Decoding Words and Difficulty with Mathematical Algorithms and be-
tween Reading Recognition Problems and Detecting Spelling Errors as
was noted on less stringent analyses. The remaining correlations
are in the expected direction but several possess a greater negative
magnitude then expected.

Discussion

The results of the multi-trait, multi-method analysis lead us to
three conclusions: (a) the Modified Checklist is generally a stable
measure, (b) the component disabilities are relatively independent,
and (c) the two measures are moderately related but appear to measure
different aspects of the component disabilities.

Several findings seem to demonstrate some measurement problems
inherent in the two measures. First, the language arts teachers are
not reliable judges of the students' disability in using mathematical
algorithms. Second, decoding words and reading recognition appear to
be moderately interrelated as they are measured on the Modified Check-
list. This interrelationship was not found on the Battery but rather
a moderate interrelationship was found on the Battery between Diffi-
culty with Mathematical Algorthims and Detecting Spelling Errors.

These two measures, the Checklist and the Battery, are not highly
related. They seem to provide two sets of data. This is particularly
true when measuring decoding skills. Finally, the component disabil-
ities were not validated to the extent that would be expected by the
investigators.

It is concluded that the Modified Checklist is a viable measure

for initial screening of learning disabled secondary students. This finding is supported by Bronoski (1977) who found that the elementary school checklist had high temporal reliability. However, Bronoski studied the elementary school checklist without a study of the behaviors associated with each of the five component disabilities. This may be the reason for the lower correlation coefficients of the present study. It must be cautioned that the behaviors associated with the Mathematical Algorithms component disability are not stable. This may be the result of the language arts teachers unfamiliarity with the behaviors or because they do not observe these behaviors in their classroom evaluations of the student.

The Battery appears to have possibilities of use by secondary LD teachers when formally measuring the component disabilities. The data on temporal reliability is not available at this time and the moderate relationship between the Reading Recognition and Detecting Spelling Errors component disabilities is presently unclear to the investigators. The Battery requires further study before its use can be strongly advocated for more than screening purposes.

In answer to the fourth and final research question, the two measures appear somewhat convergent but not to the extent that one is simply more precise than the other. Each instrument measures specific but somewhat different concepts of the component disabilities in a slightly different manner. This conclusion partially supports the field study data obtained by Carlson (Alley et al., 1979). He found that the Battery further delineated the senior high school students after the checklist had been completed by the language arts teachers.

There are several limitations to this second phase of this study.

First, because of miscommunication with the senior high school LD teaching staff and the withdrawl of one junior high LD teacher from a portion of this phase of Study 3, the sample is small and may be unrepresentative.

Second, there was missing data on the temporal reliability of the Battery. This occurred because of time limits which required that this portion of the investigation be dropped. This information must be obtained without undue delay.

<div align="center">

Discussion Related to the Assumptions and Criteria
when using the Bayesian Screening Procedure

</div>

Alley et al. have made several important assumptions and criteria that consumers should carefully consider when using the Bayesian Screening Procedure. The three studies described have direct implications to these assumptions and criteria. Each assumption and criteria are stated below. Using the findings of the present studies, statements are made that provide support, question or refute the assumptions and/or criteria made by these investigators.

Assumptions

1. Some behavioral characteristics differentiate LD from non-LD secondary students better than others. This assumption has been supported by the original investigation of Alley et al. It is also supported by the findings of Study 1 and Study 2. Study 1 findings showed that the likelihood ratios for some of the component disabilities were different from others. In addition, Study 2 findings obtained using the Principal Components analysis suggested that the Best Discriminating component disabilities were viewed as a cluster by the combined sample of professionals and parents.

<div align="center">

18

</div>

2. <u>LD specialists are an appropriate source of knowledge and experience to provide probabilities for component disabilities of LD secondary students</u>. The findings of the field study conducted by Carlson support this assumption. He did find that there were more identified LD students than the LD teachers might expect using the Checklist, 7.2 percent as compared to 5 percent. Conversely, he found that using both the Checklist and the Battery fewer students were identified than the LD teachers would expect, i.e., 3 percent as compared to 5 percent. These variations are not wide enough to refute this assumption but consumers will find that there will be variance in the prevalence when using the procedure. In addition, the findings of the cross validation phase and temporal reliability phase of Study 2 provided evidence that: (a) sub-groups of LD teachers provided similar probabilities and (b) the probabilities provided are stable.

3. <u>There is a unique cluster of behavioral manifestations that can be used when identifying LD adolescents</u>. Alley et al. found not one cluster but two clusters that can be used to identify LD adolescents. They labelled them as the Best Discriminating cluster and the High Frequency cluster of component disabilities. The principal components analyses of Study 2 yielded results which demonstrated that the two clusters of component disabilities are identified by the large combined group of professionals and parents. These two clusters are more discrete than any other of the eight clusters delineated by the factor analysis. The assumption as stated must be revised to account for these findings. It should read: <u>There are unique clusters of behavioral manifestations that can be used when identifying LD adolescents</u>.

4.   One behavioral characteristic is not sufficient to identify LD adolescents. The original investigation by Alley, et al. supported this assumption. Of the 71 component disabilities, no characteristic yielded a probability of 1.00 for either the LD or the non-LD population. No component disability was found to be exclusive to either the LD or the non-LD population--i.e., LD 40%, non-LD 0%; or LD 0%, non-LD 20%. The same findings were obtained in Study 1 among all professional and parent groups. This assumption is strongly supported by empirical evidence.

5.   The use of probability statements and their Bayesian derivatives is appropriate for describing LD among secondary students. This assumption is supported by Carlson's (Alley et al., 1979) field test, the cross validation, and temporal reliability phases of Study 2 and the findings based upon the multi-trait, multi-method analysis. The LD teachers can give reliable probabilities among subgroups of specialized teachers. These probabilities are also temporally stable. Regular class language arts teachers can provide stable judgments of behaviors based on probability statements. Finally, there is a common variance between the probability-based Modified Checklist and the probability-based Battery used in the Bayesian Screening Procedure. Strong empirical support is given to this assumption.

Criteria of the Procedure

1.   Weight the component disabilities for best differentiating results. There are two pieces of information that must be considered when using the Bayesian Screening Procedure. As found in Study 1, the weighting of the component disabilities is different among the professional or parent groups. Using only this piece of information, the

20

posterior probabilities have a great deal of variance. However if the prior probability (prevalence figure) is used for each group of judges and the posterior probability(ies) are computed based on this specific group's likelihood ratio for each component disability and/or cluster, the posterior probabilities for seven groups differ only slightly from the original group, i.e. P_{LD} = .87 to .98 as compared to P_{LD} = .96. The school psychologists' group P_{LD} = .78 is felt to be markedly different from the original P_{LD} = .96. Based on these findings this first criteria statement should be revised to state: Use the highest weighted component disabilities (likelihood ratios) and also use the prior probability (prevalence figure) of the experts when providing probabilities for best differentating results.

2. Use both the Modified Checklist and the Test Battery for the most cost efficient method. The findings of Carlson appeared to support this increased cost effectiveness (when using the Bayesian procedure). He found that the Battery further delineated the students suspected of LD from 7.2% of the general ninth grade population as determined by the checklist to 3% of this same population. The results of the multi-trait, multi-method analysis did not provide the investigators with the same type of data as Carlson's. Rather, the present data suggested that the Modified Checklist and Battery are more divergent then convergent. That is, the Battery provides more than simply a refinement of the Modified Checklist data. The Battery is measuring some different behaviors than the Checklist. Both measures should be used to refer a student as having a high probability of LD. Consumers are cautioned that the increased cost effectiveness of the Battery when used as a part of Bayesian procedure has not yet been established.

21

3.   The regular class language arts teacher should use the
Modified Checklist to determine whether a component disability is
manifested by a secondary student in the classroom.   This criterion
is supported by data from Study 3.   The regular classroom language
arts teachers' judgments demonstrated moderately high temporal re-
liability covering a two- to four-week period.   In addition, a common
variance between the teachers' judgments of behaviors associated with
the component disabilities on the Modified Checklist and the student's
performance on the test battery was found.   One particular area of
concern when applying this criterion is that language arts teachers
are an unreliable source when using their judgment to measure a
student's disability in applying mathematical algorithms.   Based on
these findings this criteria should be revised to state: The regular
class language arts teachers should use the Modified Checklist to
determine whether language arts component disabilities are manifested
by a secondary student in a classroom.   This revision implies that the
mathematical algorithm component disability be deleted from the Modified
Checklist.

4.   The behaviors associated with the Best Discriminating com-
ponent disabilities can be observed in the language arts classroom.
This criterion is related to the third criterion and must be rejected
based on the findings of Study 3.   It is true that the mathematical
algorithm behaviors appear to have content validity to LD teachers
but they are not reliably judged as present or absent by Language
Arts teachers.   Therefore, the criterion must be deleted.   The re-
vision of criterion 3 subsumes this information.   Two new criterion
should be added to the four provided by Alley et al. based on

Carlson's results and the present results. It can replace criterion 4. This criterion should read: The Battery should be used to de- lineate the severity of the secondary LD students' component disabil- ities. This criterion must be evaluated by future investigations. It seems evident that a first step is to modify the Battery. Most notably a decoding measure must be identified that more adequately reflects the behaviors included on the Modified Checklist.

5. The language arts teacher should refer a secondary student for the Battery portion of the procedure when the student is judged to have disabilities on the components of decoding words, word recognition, and recognizing and monitoring spelling errors, and, in addition, the student manifests concommitant disabilities of study skills, test taking skills, organization of written material, and knowledge of abstract concepts. This criterion must be evaluated by future investigators after the content revision of Modified Checklist has been made, i.e., deletion of the mathematical algorithm component disability. At this time the Bayesian procedure can be judged to be useful in screening for LD students. It will require future modifications based on its wider use. Consumers are encouraged to provide data to the developers making these revisions.

References

Alley, G.R., & Deshler, D.D. Checklist of learning difficulties (Secondary). Library of Congress, 1978.

Alley, G.R., & Deshler, D.D., & Warner, M.M. Identification of learning disabled adolescents: A Bayesian approach. Learning Disability Quarterly, 1979, 2, 76-83.

Angoff, W.H. Scales, norms, and equivalent scores. In R.L. Thorndike (Ed.), Educational measurement (2nd ed.). Washington: American Council on Education, 1971.

Bronoski, J.H. A component disability screening instrument to be used by regular class teachers to screen learning disabled children. Unpublished master's thesis, The University of Kansas, 1977.

Campbell, D., & Fiske, D. Convergent and discriminant validation by multi-trait multi-method matrix. Psychological Bulletin, 1959, 56, 81-105.

Chalfant, J.C., & King, F.S. An approach to operationalizing the definition of learning disabilities. Journal of Learning Disabilities, 1976, 9, 228-243.

DeRuiter, J.A., Ferrell, W.R., & Kass, C.E. Learning disability classification by Bayesian aggregation of test results. Journal of Learning Disabilities, 1975, 8, 365-372.

Harber, P.I., & Harber, J.E. Note to Editor. Learning Disability Quarterly, August 10, 1979.

Larsen, S. Learning disabilities and the professional educator. Learning Disability Quarterly, 1978, 1, 5-12.

Meyen, E.L. An investigation of age-placement, difficulty, and importance of basic skills in the curriculum for educable mentally retarded students. Unpublished doctoral dissertation, The University of Iowa, 1968.

Wimmer, D. A validation of the cognitive content of career education and investigation of age placement on cognitive career education skills of mildly mentally retarded. Unpublished doctoral dissertation, The University of Kansas, 1979.

Wissink, J.F., Kass, C.E., & Farrell, W.R. A Bayesian approach to the identification of children with learning disabilities. Journal of Learning Disabilities, 1975, 8, 158-166.

Table 1

Results of Content Validity Judgments
on Secondary Learning Problem Checklist

Disability in:	Percentage Agreement	FEBRUARY 12					FEBRUARY 19	
		40%	60%	80%	100%		80%	100%
A. decoding words		29		7	22,28			(29)
B. recognizing sight words				9,17				
C. use of mathematical algorithms			16		24,25			(16)
D. production of themes of adequate length				21	8			
E. the organization and arrangement of written materials			5,19	6,15 20				5,(19)
F. using study skills		2	1,27	4,11	10,18,26		2	(1,(23),27
G. test taking skills			3,23		14			(3)
H. detecting errors			12		13		(12)	
Total		2	8	9	10	Totals following consensus meeting & item revision	11	18

(N) Change in item content. Item content was altered for seven items.

Item #23: Was originally considered as part of the "test taking skills" component. After modifying the item's content, the group consensus was that the item related to the "study skills" component.
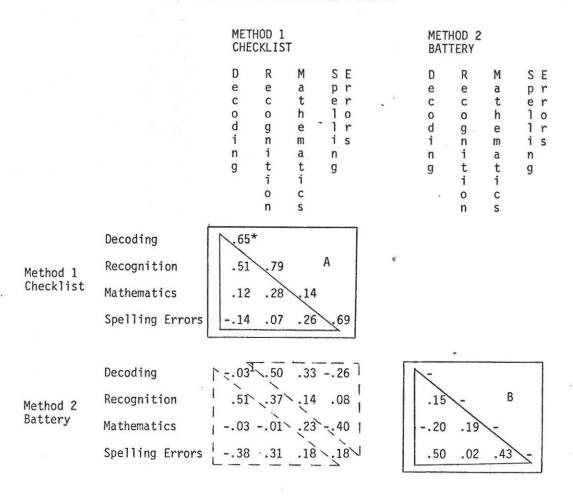
Table 2

LD Teachers' Judgment data on 29 Target Behaviors
of the Modified Checklist of Learning Problems (Secondary)

Target Behavior by Component Disability	Importance x̄	SD	Frequency LD x̄	SD	Frequency Non-LD x̄	SD	Accessibility x̄	SD	Pre-K	K	1	2	3	4	5	6	7	8	9	10	11	12	Post HS
Decoding Words																							
7 Context clues	5.40	0.89	5.40	0.55	3.20	1.30	4.20	3.35	-	-	60	-	40	-	-	-	-	-	-	-	-	-	-
22 Structural Skills	5.20	1.30	5.20	0.84	2.60	1.34	2.80	1.86	-	-	20	40	40	-	-	-	-	-	-	-	-	-	-
28 Sound out words	5.80	1.30	5.60	0.97	3.00	1.00	3.00	1.58	-	-	40	-	60	-	-	-	-	-	-	-	-	-	-
29 Word attack **Word Recognition**	5.80	0.84	6.20	0.84	3.00	1.41	2.80	1.92	-	-	20	20	40	20	-	-	-	-	-	-	-	-	-
9 Common Words	5.40	0.55	6.00	0.00	3.20	1.30	5.00	0.71	-	-	-	-	40	20	-	40	-	-	-	-	-	-	-
17 Technical Words **Mathematical Algorithms**	5.60	1.24	6.00	0.71	3.60	1.34	3.80	1.79	-	-	-	-	-	40	20	20	20	-	-	-	-	-	-
16 Reading graphs	5.40	0.89	5.60	1.02	3.20	0.84	4.20	1.03	-	-	-	20	-	20	20	20	20	-	-	-	-	-	-
24 Changing units	6.20	0.84	5.00	1.87	3.00	0.71	3.00	1.22	-	-	20	20	-	20	-	20	20	-	-	-	-	-	-
25 Calculating Percent **Monitoring Errors**	5.00	1.22	6.00	1.00	3.00	1.00	3.40	1.14	-	-	-	-	-	20	-	40	40	-	-	-	-	-	-
12 Correct spelling	5.40	0.55	6.00	0.71	3.00	0.71	3.40	2.51	-	-	-	20	20	-	40	-	20	-	-	-	-	-	-
13 Detect Comp. Errors **Writing Themes**	6.00	0.70	6.40	0.54	3.40	0.54	3.20	1.56	-	-	-	-	40	20	-	20	20	-	-	-	-	-	-
8 Adequate length Themes	5.40	1.34	6.20	1.30	3.40	0.54	2.80	1.92	-	-	20	-	-	40	-	-	40	-	-	-	-	-	-
21 Adequate length research	5.20	1.30	6.40	0.89	3.80	0.83	3.40	1.51	-	-	-	-	-	-	40	20	-	20	20	-	-	-	-
Test Taking																							
3 Estimate answers	6.00	0.70	6.40	0.54	3.60	0.54	5.20	1.09	-	-	-	20	-	-	-	40	20	20	-	-	-	-	-
4 Review sessions	5.60	0.74	5.80	1.30	3.20	0.83	4.40	1.81	-	-	-	-	20	-	-	20	40	-	20	-	-	-	-
4 General test taking skills	6.20	0.44	6.80	0.44	3.20	0.44	5.20	1.48	-	-	20	-	20	-	-	20	20	-	20	-	-	-	-

Table 2 (cont.)

Target Behavior by Component Disability	Importance $\bar{x}$	SD	Frequency LD $\bar{x}$	SD	Frequency Non-LD $\bar{x}$	SD	Accessibility $\bar{x}$	SD	Pre-K	K	1	2	3	4	5	6	7	8	9	10	11	12	Post HS
Study Skills																							
23 Study Time	6.40	0.54	6.80	0.44	4.00	1.00	5.20	1.48	-	-	-	-	20	-	20	20	20	-	20	-	-	-	-
1 Notetaking	6.00	0.70	6.80	0.44	3.60	0.54	3.80	2.58	-	-	-	-	-	20	-	-	60	-	20	-	-	-	-
2 Read Flexibility	6.40	0.54	6.60	0.89	3.40	0.54	4.00	2.17	-	-	-	-	40	-	40	20	-	-	-	-	-	-	-
10 Survey	5.40	1.14	6.60	0.54	4.80	1.30	4.80	2.16	-	-	-	-	-	20	-	20	40	-	20	-	-	-	-
11 Reference	6.20	1.09	6.00	0.70	2.80	1.48	3.80	1.78	-	-	-	-	20	20	-	20	20	-	20	-	-	-	-
18 Review	5.80	0.44	6.60	0.54	4.00	1.58	4.20	1.64	-	-	-	-	20	-	-	40	20	-	20	-	-	-	-
19 Summarize	6.40	0.54	6.00	1.22	2.80	0.54	3.20	1.64	-	-	-	-	25	25	-	-	25	-	25	-	-	-	-
20 Classify	5.80	1.30	6.60	0.54	3.40	0.89	3.00	1.73	-	-	-	-	20	20	-	-	40	-	20	-	-	-	-
26 Question	5.40	0.79	6.20	0.45	3.00	1.00	3.60	1.82	-	-	20	-	-	-	20	20	20	-	20	-	-	-	-
27 Sequence of directions	6.60	0.55	6.20	0.45	3.20	0.45	2.60	1.52	-	-	40	-	40	-	-	-	20	-	-	-	-	-	-
Organization Skills																							
5 Organizing ideas	5.60	0.89	6.00	1.00	3.40	1.71	3.40	1.71	-	-	-	20	-	20	20	-	40	-	-	-	-	-	-
6 Differentiate Paragraphs	5.60	0.89	6.20	0.84	3.20	0.45	3.20	1.30	-	-	-	-	20	20	-	40	20	-	-	-	-	-	-
15 Paragraph elements	6.40	0.89	6.60	0.89	3.60	1.14	2.60	1.34	-	-	-	-	20	20	-	40	20	-	-	-	-	-	-

Table 3

The Multi-trait Multi-method Matrix

		METHOD 1 CHECKLIST				METHOD 2 BATTERY			
		Decoding	Recognition	Mathematics	Spelling Errors	Decoding	Recognition	Mathematics	Spelling Errors
Method 1 Checklist	Decoding	.65*							
	Recognition	.51	.79		A				
	Mathematics	.12	.28	.14					
	Spelling Errors	-.14	.07	.26	.69				
Method 2 Battery	Decoding	-.03¹	.50	.33	-.26	-		B	
	Recognition	.51	.37	.14	.08	.15	-		
	Mathematics	-.03	-.01	.23	-.40	-.20	.19	-	
	Spelling Errors	-.38	.31	.18	.18	.50	.02	.43	-

* 14 day temporal reliability

[1] Convergent validity index

APPENDIX A

Directions for Rating Scale

This rating scale asks you to make judgments (yes or no) concerning several behaviors you have possibly observed in your students.

The behaviors have been grouped into eight component disabilities (A-H), the first of which is Decoding words. For each component disability there are several related behaviors. For example, under Decoding words there are four related behaviors listed.

FOR EXAMPLE:

A)   Component Disability: <u>Decoding words</u>

<u>Item 7</u>   Unable to use Context clues as aids in unlocking words.

<u>Item 22</u>   Unable to use structural skills to unlock words, i.e., does not divide words into smaller units - "candidate" into can-di-date, root words and ending - "playing" into play and ing, and prefixes, roots and endings - "reporter" into re-port-er.

<u>Item 28</u>   Unable to sound out words.

<u>Item 29</u>   Unable to read fluently, i.e., reads word-by-word, incorrect pauses, wrong inflections, etc.

For each behavior listed you are asked to answer the following three questions:

The content of this item is clear to me?

(yes)   no

The behavior described in this item is relevant to successful performance in my class?

yes   (no)

The behavior described in this item can be observed in students in my classroom?

(yes)   no

Please circle your response to these questions as listed for each of the 29 items in this rating scale.

APPENDIX B

STARTING TIME: _____

ENDING TIME: _____

TOTAL TIME: _____

_____
(Student's Name)

_____
(Teacher's Name)

_____
DATE

CHECKLIST OF ACADEMIC PROBLEMS
Gordon R. Alley and Donald D. Deshler, 1977 ©
(revised 3-12-79)

Based on your observations of the student, please check behaviors which are so severe that they seriously affect this student's school program.

		YES	NO	NO INFOR-MATION
1.	Unable to outline/take notes on material. In-class notes are sketchy and much doodling evident.	___	___	___
2.	Unable to read with flexibility, i.e., does take an equal amount of time to read easy and difficult passages.	___	___	___
3.	Unable to apply appropriate test taking skills within specific subject areas., i.e., does not estimate answers in math.	___	___	___
4.	Unable to consult with teacher a few days before the test concerning the major emphasis of the test, does not attend review sessions that meet before or after school.	___	___	___
5.	Unable to structure materials into a logical sequence. Cannot organize two or more sources of information into one idea or topic.	___	___	___
6.	Unable to differentiate one paragraph from another. Does not recognize when one major point should be separated from a second major point, or fails to use transition sentences.	___	___	___
7.	Unable to use context clues as aids in unlocking words.	___	___	___
8.	Unable to produce themes of adequate length for the student's given grade level in English.	___	___	___
9.	Unable to recognize words that are common across content areas, i.e., knowledge, suggestion, selected, etc.	___	___	___

10. Unable to survey material before studying in depth. Does not scan for major points before reading material intensively.   ___ ___ ___

11. Unable to use reference materials or resources, i.e., does not use dictionary, encyclopedia, or the library.   ___ ___ ___

12. Unable to choose correct spelling of words in multiple choice format, by recognition of correct word among incorrect words.   ___ ___ ___

13. Unable to detect errors that he/she makes in composition writing.   ___ ___ ___

14. Unable to apply appropriate test taking skills across subject areas, i.e., does not thoroughly read instructions, preview the entire test before responding, checking answers, does not recognize "give away" questions, or obvious answers, or chooses an answer that is obviously wrong.   ___ ___ ___

15. Unable to construct a logical paragraph which includes a topic sentence, supporting facts, and a concluding sentence.   ___ ___ ___

16. Unable to use rules to solve mathematics problems, i.e., cannot read graphs or has difficulty constructing a simple graph.   ___ ___ ___

17. Unable to recognize words that are specific to a given content area. i.e., composition, author, irony, poetry, perpendicular, radius, constitution, meteors, photosynthesis, etc.   ___ ___ ___

18. Unable to review materials, i.e., does not reread assigned materials.   ___ ___ ___

19. Unable to summarize information, i.e., cannot identify main points and important facts of a selection in two or three sentences.   ___ ___ ___

20. Unable to classify/organize materials, i.e., notebook is disorganized.   ___ ___ ___

21. Unable to produce research papers of adequate length for the student's given grade level specific to a content area, i.e., Science, History, Health, etc.   ___ ___ ___

22. Unable to use structural skills to unlock words, i.e., does not divide word into smaller units - "candidate" into can-did-ate, root words and ending - "playing" into play and ing, and prefixes, roots and endings - "reporter" into re-port-er.   ___ ___ ___

23. Unable to allot a reasonable amount of study time
    i.e., does not use study time appropriately in class. ____ ____ ____

24. Unable to use mathematics rules which are specific
    to application type problems, i.e., has difficulty
    changing units of measure as dollars to cents, inches
    to yards, and minutes to hours. ____ ____ ____

25. Unable to use rules specific to mathematics con-
    tent, i.e., difficulty counting number of errors
    or calculating percent correctly on tests, etc. ____ ____ ____

26. Unable to apply questioning skills. Ask inappropriate
    questions to the discussion or content topic, or does
    not ask questions during the class period. ____ ____ ____

27. Unable to follow a sequence of directions, verbal
    or written. ____ ____ ____

28. Unable to sound out words. ____ ____ ____

29. Unable to use word attack skills. ____ ____ ____

30. Please add any other specific problems of which
    you may be aware. ____ ____ ____