INSTITUTE FOR RESEARCH IN LEARNING DISABILITIES
The University of Kansas
Lawrence, Kansas, 66045
Emphasis on Adolescents and Young Adults

PRODUCTION DEFICIENCY VS. PROCESSING DYSFUNCTION: AN EXPERIMENTAL ASSESSMENT OF LD ADOLESCENTS

Daryl F. Mellard and Gordon R. Alley

Research Report No. 40

April, 1981

The University of Kansas Institute for Research in Learning Disabilities is supported by a contract (#300-77-0494) with the Bureau of Education for the Handicapped, Department of Health, Education, and Welfare, U. S. Office of Education, through Title VI-G of Public Law 91-230. The University of Kansas Institute, a joint research effort involving the Department of Special Education and the Bureau of Child Research, has specified the learning disabled adolescent and young adult as the target population. The major responsibility of the Institute is to develop effective means of identifying learning disabled populations at the secondary level and to construct interventions that will have an effect upon school performance and life adjustment. Many areas of research have been designed to study the problems of LD adolescents and young adults in both school and non-school settings (e.g., employment, juvenile justice, military, etc.)

Director: Donald D. Deshler

Research Coordinator: Jean B. Schumaker

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

The preparation of this document was supported by a government contract. The views expressed here are those of the Institute, and do not necessarily reflect official positions of the Bureau of Education for the Handicapped, DHEW, USOE.

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 conducted. 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 have or currently are participating in various studies include: Unified School District USD 384, Blue Valley; USD 500, Kansas City, Kansas; 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 several school districts in Missouri, including Center School District, Kansas City, Missouri; the New School for Human Education, Kansas City, Missouri; the Kansas City, Missouri School District; the Raytown, Missouri School District; and the School District of St. Joseph, St. Joseph, Missouri. Other participating districts include: 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, Leavenworth, and Sedgwick County, Kansas Juvenile Courts. Other agencies which have participated in out-of-school studies are: Penn House and Achievement Place of Lawrence, Kansas; Kansas State Industrial Reformatory, Hutchinson, Kansas; the U. S. Military; and 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 support 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

According to Torgesen (1977), LD students' deficient performance is not reflective of cognitive processing deficits but of a production deficiency. The student is capable of satisfactory achievement, but does not achieve at such a level due to a passive approach to learning. This study evaluated these hypotheses using a discrimination learning task and varied reinforcement. LD students were matched with nonhandicapped peers and administered discrimination learning problems with treatment (reinforcement, response cost) and control conditions. Torgesen's hypothesis was not supported. Processing deficits were identified in the LD students' ability to code, recode, and recall information compared to regular class students. They also were deficient in benefiting from explicit feedback. All students in the control group demonstrated overall superior performance to those in the reinforcement, response cost condition. The findings were related to influencing behavioral traits and cognitive deficits.

PRODUCTION DEFICIENCY VS. PROCESSING DYSFUNCTION: AN EXPERIMENTAL ASSESSMENT OF LD ADOLESCENTS

Torgesen (1977) theorized that learning disabled (LD) children's performance deficits occur as a result of a "passive" approach to tasks rather than a cognitive processing deficit as assumed by other existing definitions. Through extensive research investigating observed performance of LD students, "a whole catalogue of discrete disabilities" have been identified, e.g., attentional, perceptual, short-term memory, retrieval, associational, and processing deficits (Torgesen, 1977, p. 27). Too often, the failure to establish operational definitions and reliable and valid measurement procedures has led to confusion and division within the field of learning disabilities and the professional groups desiring to serve them. Torgesen suggested that the field must distinguish between ability deficits due to a hypothesized process delay or dysfunction and performance deficits. "Performance deficits can be caused either by a lack of ability related to a task or by failure to apply efficiently those abilities or capacities which are present" (Torgesen, 1977, p. 28). The failure to employ those available abilities has been described by others as a "production deficiency" (Flavell, 1970, p. 182).

Consistent with Torgesen's view of performance deficits, the student's motivation and level of involvement with a task must be considered in attempts to try to understand why he/she failed. To conclude that task failure reflects a cognitive process deficit would be premature. The alternative hypothesis, that the child was only passively involved in the task and that, consequently, task failure reflects a production deficiency, should also be considered.

> It has been shown that failure in a particular task setting or in certain school subjects does not necessarily indicate a defect in some specific psychological process such as attention, short-term memory, or perception. Rather, poor performance in many different task settings may be due to the child's failure to actively agree the task through the use of efficient strategies and other techniques of intelligence. (Torgensen, 1977, p. 33)

Support for Torgesen's view was reported in a sequence of studies that initially investigated the development of visual selective attention and verbal rehearsal in LD boys (Tarver, Hallahan, Cohen, & Kauffman, 1977). The focus of later studies by Hallahan and his colleagues included: (a) the influence of locus of control (Hallahan, Gajar, Cohen, & Tarver, 1978), and (b) reinforcement and response cost (Hallahan, Tarver, Kauffman, & Graybeal, 1978). These researchers have concluded that their LD population was not deficient in selective attention, but rather that "normal children actively engage in using learning strategies whereas learning disabled children do not" (Hallahan, Tarver, Kauffman, & Graybeal, 1978, p. 438). Consistent with Torgesen's conceptualization, the above authors viewed the LD child as a passive rather than an active learner and attributed the deficient performance to LD students' failure to be actively involved in their tasks as evidenced by their limited use of facilitative problem-solving strategies (e.g., cumulative verbal rehearsal). A second conclusion was that the LD student is capable of actively adopting a problem-solving strategy under controlled conditions (e.g., reinforcement and punishment/response cost) for correct/incorrect responses.

The results of the Hallahan, Tarver, Kauffman and Graybeal (1978) study included significantly improved performances by the LD students in a reinforcement (e.g., pennies) condition. Selective attention and rehearsal deficits were removed by modifying the task (i.e., providing monetary reinforcements). Thus, the productionally deficient, passive learner improved his/her performance through the use of a reinforcer.

Haines and Torgesen (1979) investigated the interaction between the efficient use of problem-solving strategies (e.g., verbal rehearsal as a mnemonic strategy) and incentives with groups of second graders designated as good and poor readers. The students were told to recall (i.e., point to) a sequence of pictures in the same order as presented by the examiner. The task was completed under two conditions--with and without rein-forcement (e.g., pennies and social reinforcement). Students were rewarded for accurate, sequential recall. Haines and Torgesen reported statistically significant differences between the good and poor readers' use of rehearsal and recall accuracy on the first session (no incentive). Also, 21 good readers clearly explained their rehearsal strategy after the trials, while only 10 of the poor readers did; the difference was found to be statistically significant. This finding was thought to support Hallahan, Tarver, Kauffman and Graybeal's (1978) findings "that many of the performance problems of reading disabled children may not be due to a limited memory or learning capacity per se, but rather to a failure to apply efficient task strategies" (Haines & Torgesen, 1979, p. 53).

Haines and Torgesen reached a major conclusion in examining the results of the incentive on the performance of the poor readers in the second session: incentives influenced recall and rehearsal scores resulting in equivalent performance improvement (recall and rehearsal) for both groups. In relation to the poor readers, "not only did the overall rate of rehearsal increase, but also the number of children who reported the conscious application of the strategy increased" (Haines & Torgesen, 1979, p. 53). Incentives raised the poor readers' recall and rehearsal to a level equivalent to that of the good readers in the first session.

However, since the incentives did not eliminate rehearsal differences between good and poor readers, "the results do not support the conclusion that the only difference between children in the two reading groups in Session I was a motivational one" (p. 53). The authors speculated that in spite of incentives or because of a differential effect of the incentive, the two groups' motivation may not have been equivalent. Also, poor readers may need instruction and practice in the use of rehearsal while good readers use rehearsal more frequently and, thus, are more adept at it. For whatever reason, the poor readers "do not appear to bring the same level of intrinsic motivation for good performance to experimental tasks as do children who have learned to read normally" (p. 54).

These two studies (Hallahan, Tarver, Kauffman, & Graybeal, 1978; Haines & Torgesen, 1979) were reviewed as supporting evidence that performance differences between LD children or poor readers (Haines & Torgesen, 1979) and normal children, may reflect motivational characteristics (e.g., passive learner) and production deficiences.

The purpose of the study conducted by the University of Kansas Institute for Research in Learning Disabilities was to replicate these results in an LD defined, adolescent population with a higher level cognitive strategy (e.g., hypothesis testing) than verbal rehearsal in a research design lending itself to a causal analysis.

Baumeister (1967) and Stanovich (1978) proposed that a processing deficit or developmental delay could be evaluated through the interaction in a factorial design. For example, by manipulating a variety of experimental factors (e.g., level of motivation) within a given task (e.g., verbal rehearsal, visual imagery, hypothesis testing) one can examine the interaction between subject group (e.g., LD, normal, and low-achieving adolescents) and the manipulated experimental factor known to affect the specific cognitive process.

To resolve the issue of developmental delay vs. production deficiency, an experimental paradigm was needed that would control for these competing hypotheses of performance. Levine (1969, 1975) and his colleagues have developed a discrimination learning task which is sensitive to developmental differences of cognitive development (Gholson, 1977; Neimark, 1975; Phillips & Levine, 1975; Phillips, Levine, & O'Brien, 1978; Richman & Gholson, 1978) and to the effects of reinforcement conditions (Mims & Gholson, 1977).

Methodology

Subjects

The 85 students participating in this experiment were volunteers from the sixth-, seventh-, eighth-, and ninth-grade populations in six schools in four school districts in Northeast Kansas.

Learning disabled population. The 43 learning disabled students were identified as such by their respective schools and were currently receiving intervention services in the LD resource rooms for approximately one hour each day. Given the limited size of the LD population, the students were not randomly selected and may more be accurately considered as volunteers.

Non-learning disabled population. The non-LD comparison group (N=42) was randomly selected from the participating schools' regular class non-handicapped populations. Each student was matched to an LD student from the same school on the criteria of sex and age. None of the comparison group students had previously participated in any special education program (e.g., speech therapy, gifted program, visually impaired program, etc.).

Participating samples. After return of a signed parental consent form, background information was collected for each student, e.g., number of schools attended, school absences, achievement and ability test scores, and birth information.

Of the 85 experimental subjects, 84 (42 LD and 42 non-LD students) were included in the analyses. One LD student's scores were not analyzed with the group data. This student had not completed the preliminary training problems during the alotted time and consequently had not solved any of the experimental problems. This subject was replaced by another LD student matched to the non-LD student cohort.

The mean age of the 84 treatment subjects was 14.2 years with a standard deviation of .83 years. Mean grade level was: LD group, 8.5; NLD group, 8.6. Seventy-six percent (N=64) of the sample was male. Eighty-five percent (N=71) were Caucasian; 13 percent (N=11) Black Americans, while two percent (N=2) represented other minority populations.

For the LD students grade level scores in reading ranged from 2.4 to 7.1 ($\bar{x}$ = 4.33), while math grade level scores ranged from 2.5 to 7.4 ($\bar{x}$ = 5.06). For the NLD group, reading grade level scores ranged from 4.0 to 12.9 (x = 8.94) with math scores ranging from 4.4 to 12.9 ($\bar{x}$ = 8.5). Table 1 presents specific characteristics of the subjects.

Insert Table 1 about here

Materials

A series of bi-valued discrimination learning problems were designed according to the restrictions described by Levine (1969). The stimuli for these problems were prepared on 14 by 21 cm (5.5 x 8.5 inch) cards which were laminated in transparent plastic. Each card was used for one trial.

With the exception of the first training problem, consisting of four bi-valued dimensions, the 14 remaining problems had eight dimensions. The four dimensions and respective values in the first training problem were: letter (G vs. V), size (big vs. little), color (red vs. green), and box position (box above the letter vs. box below the letter) (see Figure 1). For all eight dimensional problems, the dimensions were: letter, size, color, box position, box contents (dots vs. empty), number of borders (one vs. two), border shape (square vs. circle), and border consistency (solid line vs. dashed line) (see Figure 2). Letter position (i.e., left vs. right side) was not a dimension in any of the problems.

Eleven colors were paired in eight combinations for the fifteen problems. Color pairings for each problem were made by selecting one color on a color spectrum wheel and pairing it with the color on the opposite side of the spectrum. This procedure ensured maximum recognitory distinction between color pairs. Within each problem, only one pair of colors was used.

Nineteen upper case consonants were randomly paired in fifteen different combinations for the letter stimuli. An S and Z pairing was not included. The letters were centered 9 cm apart. The large and small letters were 3.7 cm and 1.6 cm high, respectively. The outer border had an 8 cm diameter (circle) or a 7.8 cm side (square). The inner border was .3 cm to the inside of the outer border. The boxes were centered .8 cm inside of the border, were .19 cm long and .6 cm wide. The five dots in the dotted box were .1 cm in diameter and were spaced the length of the box.

The stimuli in the four-dimensional problem were arranged such that after any three consecutive trials, the student could logically define solution. In the eight-dimensional problems, after any four consecutive trials (stimuli) the student could logically define the solution (Levine, 1969, 1975).

The stimuli for each of the fifteen problems were placed in the aforementioned sequence, and held together at the top with two metal rings. The examiner turned the stimulus card for each trial. A blank 20 by 12.7cm white index card was placed between each stimulus card for those trials on which an introtact was requested of the student (examiner asked if student had noticed, "What was right all the time").

The first three training problems for Session 1 consisted of twenty trials. The fourth and fifth training problems for Session 1 included twenty-four trials. Only two training problems, each with maximum of 24 trials, were used in Session 2. All experimental problems had a maximum of twenty trials.

Measurement Systems

The individual sessions were recorded with cassette tape recorders to provide a procedure for gathering reliability measures and data verification. Information on the student (e.g., identification number, date, session number) and his/her responses were recorded on a DATAMYTE 900 series data collector (Electro General Corp., Minnetonka, MN). The data collected included the ordinal and cardinal numbers of the problem. Three student responses were also coded: (a) correct or incorrect choices on each trial, (b) the student's hypothesis expressed as the problem's solution, and (c) whether or not the student's response to a stimulus was consistent with the previous verbally stated hypothesis, i.e., response consistency. Each possible student response was assigned a code number. Immediately after a student made a verbal response, the researcher pushed the appropriate key on the DATAMYTE which had an internal clock calibrated to .01 minute. Thus, each time a key was pressed, the time of the response was also recorded. This permitted the measurement of response latencies and the total time required to complete the problem.

Correct responses regarding stimulus choices were also recorded with paper and pencil by the researcher. The paper- and- pencil recordings which were completed after the DATAMYTE recordings allowed the researcher to determine when five successive trials were correct.

Setting

Each student individually met with the researcher in a classroom or office area in his/her respective school. While the physical arrangement screened the students from visual distractions, auditory distractions (e.g., band practice, class bells, student discussions) were not always effectively screened.

The student faced the examiner across a desk containing the stimulus materials, the DATAMYTE and cassette recorder. During the second session a stack of 27 poker chips and an opaque container for collecting the chips were added.

Procedures

Training of data collectors. To model the correct procedures, the principal investigator administered the task to participating research assistants. All assistants administered the task to a minimum of three age-appropriate subjects to refine their skills.

A research assistant with two years' experience as a paraprofessional in a special education class for the moderately mentally retarded was employed. She administered the pilot procedures of Sessions One and Two to seven naive pilot subjects plus the principal investigator as a training exercise. Performance was evaluated and corrected based on visual observations and cassette recordings. No further training was required to administer the procedure during the experiment.

Session 1 (no monetary reinforcement). Each student individually met with a research assistant for the length of a class period which ranged from 42 to 55 minutes in the five schools. During this time, each student completed the five training problems and as many experimental problems as time allowed. Problem presentation stopped at the end of the class period; data from responses on incomplete problems were not analyzed.

Prior to presentation of the first training problem, a color blindness screening test was administered to the students. Results indicated that all students were able to identify, by name, the stimulus colors in the five training problems. Second, none of the students reported that they were color blind nor did contradictory information appear in any of the students' school health records. This screening was considered necessary since color was a possible solution in the discrimination problems. Five training problems were used to train the students on the requirements of the experiment. These procedures were adapted from the previous work of Byrd (1979) and Phillips and Levine (1975).

For each problem, the deck of stimulus cards was placed in front of the student who was instructed to point to the picture he/she thought was always right. The examiner indicated whether the choice was correct or wrong, recorded the response on the DATAMYTE, and then turned the stimulus card face down. The student responded to the first 10 cards in the same manner. After the student had responded to the tenth card, the examiner asked the student if he/she had noticed, "What was right all the time?" This was called an "introtact." If the student responded correctly, the examiner responded, "That's right, (attribute name) is the correct answer," and the trials continued until 10 consecutive trials were correct. The second problem was then presented.

If the student responded incorrectly to the introtact, or did not know the answer, the examiner gave him/her a "hint" (told him/her to try the relevant dimension (e.g., color, size, letter form, or line position) and resumed the problem. If the student did not verbalize the correct solution at the end of the 20 trials, the correct answer was given, the problem was repeated, and the student was instructed to "pick every picture with (attribute name) and see what happens." Students continued to a criterion of 10 correct responses in a row which ensured that they experienced a solution. The second problem, also consisting of 20 trials, was completed in a similar manner.

For the third, fourth, and fifth problems (20, 24, and 24 trials, respectively), introtacts were increasingly introduced. For Problem 3, the student was asked, "What is right all the time?" after trials 6, 10, 12, and every second trial thereafter. For Problem 4, the introtacts were requested after the first three trials, again after the fifth, and then every trial thereafter. On Problem 5, the student was asked to state his/her best hypothesis spontaneously after each trial, and to do so "from now on." The examiner also reminded the student if he/she forgot to verbalize the hypothesis each time. Statements of a compound, or an either/or hypothesis were corrected during pretraining, and queried in the subsequent experimental problems. By completing these five pretraining problems, the student learned the nature of the task and its mechanics, as well as to respond to the verbal hypothesis probes (introtacts) after each trial.

The experimental problems were then presented. These problems were arranged in ten different sequences, and students randomly received one of the ten problem sequences. After all the students had completed Session 1, they were matched on their performance within their LD and NLD groupings. Specifically, the mean trial of the last error for the student's experimental problems was calculated and used as the matching variable. For each problem, the trial on which the student made his/her last error was determined. These values were summed and divided by the number of problems administered to that student. These mean values were then ranked in order of increasing magnitude within the two subject groupings (LD and NLD). The rank ordered lists were blocked into pairs for the experimental and control groups, and one student from each pair was randomly assigned to either the experimental or the control group. The other student in the block was assigned to the remaining group. Following this random student assignment (42 LD and 42 NLD), the cell sizes in the 2 x 2 factorial design were: LD experimental group, 21; LD control group, 21; NLD experimental group, 21; and NLD control group, 21.

Session 2 (treatment with a symbolic reinforcement). The procedures for Session 2 were basically the same as for Session 1 with the exception of the consequences which followed student responses. For the experimental group, reinforcement and response cost in the form of tokens were contingent on correct and incorrect answers, respectively. The tokens were exchanged at the end of the session for a minimum of $3.00. Tokens were also given to and taken away from members of the control group contingent on their performance to control for the distracting nature of the poker chip delivery. The control students were told at the beginning of the session that they would receive $3.00 at the end of the session. The experimental students were told the amount earned after completion of each problem. This was done to heighten the student's motivation. For control group members, the next problem was presented without reference to poker chips or money.

During this session each student again met individually with the researcher for the length of a class period, or until all the problems had been completed, whichever was the shorter time period. Since the stimulus materials used in Session 1 were also used in Session 2, the color recognition test was omitted. Although the same materials were used, a different stimulus value (e.g., big, green, circle, etc.) than that assigned in Session 1 was selected as a solution to each problem in Session 2.

All students completed two preliminary problems, during which the students in the experimental group were told how the chips would be later exchanged for payment. If a student in the control group asked about the chips, the examiner responded that they were being used "to help keep track of your answers." Students were instructed not to discuss their earnings with their classmates for several days. These instructions were used to minimize possible contamination, expectancy, and demoralization effects (Cook & Campbell, 1979). While the order of administering the two preliminary problems was fixed, the experimental problems were administered randomly in one of ten possible sequences.

Research Design

Independent variables. The principal variables were two subject groupings (LD vs. NLD) and two treatment levels (experimental group vs. control group). These two variables were completely crossed and balanced in this design.

As indicated in the experimental procedures, students were grouped into blocks of two. This nesting factor, matched subjects, controlled (i.e., reduced) the error variance between the groups of students, LD and NLD. The blocking variable was calculated from each student's performance on the experimental problems administered in the first session. This variable was also used as a covariate in the statistical analysis. Marker variables including birthdate, age, sex, school attendance center, classification (LD or NLD), and intelligence test scores were used for matching subjects and in the analyses.

Dependent variables. Discrimination learning problems provide a variety of information concerning each student's performance. Most of the variables studied here have been discussed in the previous review of related research. Operational descriptions are provided for the sake of clarity, as well as to serve as representative citations of previous research into these variables.

1. Frequency of dimension sampling (Byrd, 1979; Phillips et al., 1978). Analysis of this variable indicated whether or not the student tested each of the eight dimensions with the same frequency. Byrd reported that students have shown response preferences and interproblem transfer. For each student, the number of different dimensions sampled in each problem are counted. These numbers are summed across all of the problems completed by a student.

2. Percentage of correct responses following the first four trials (Gholson & Danziger, 1975). Feedback on four trials is required before the student has enough information to solve a problem using the most efficient strategy behavior. Less efficient strategies would require additional trials. With these limitations, a measure of a student's learning is the percentage of his/her correct responses after the first four trials. For each student, the first four trials of the problems are omitted. For the remaining trials, the number of correct answers were tabulated. The ratio of the correct number of trials to the total number of trials is calculated. This quotient is then converted to a percentage which numerically indicates the effectiveness of the student's learning based on feedback from the first four trials. Students who are perfect focusers would be 100% effective.

3. Response consistency (Kemler, 1978; Mims & Gholson, 1977; Phillips & Levine, 1974). This dependent variable considered the correspondence between a student's verbally stated solution hypothesis and his/her response choice on the following trial. For example, if a student verbally states, "The big letter is correct all the time", one would predict that the student will point to the stimulus containing the big letter on the following trial. This consistency was calculated for each student by counting the number of trials on which the introtact and the subsequent response were consistent and dividing by the total number of trials. This quotient was then converted to a percentage.

4. Retention of confirmed hypothesis (Byrd, 1979; Gholson & Danziger, 1975; Kemler, 1978; Mims & Gholson, 1977). The theory of discrimination learning backed by previous research postulates that students will maintain an hypothesis when the intervening feedback indicates his/her choice is correct. For example, consider any two trials in a discrimination learning problem, and a situation in which the student selects the correct stimulus picture on the first trial. The student verbalizes his/her hypothesis of the correct solution (e.g., "the big letter"). This student has maintained a confirmed hypothesis over these trials.

   This dependent variable was calculated by counting the number of correct trials. From this set, the proportion of consecutive trials with the same stated hypothesis, i.e., introtact, was calculated. This quotient was converted to a percentage representing the percentage of retained confirmed hypotheses.

5. Retention of disconfirmed hypothesis (Byrd, 1979; Gholson & Danziger, 1975; Mims & Gholson, 1977). This measure indicated the effect of negative feedback, i.e., being told "wrong", on retaining disconfirmed hypotheses. This variable was calculated from those trials on which the student made an incorrect choice. From this set of trials, the proportion of two consecutive hypotheses being the same was calculated. These vaules were calculated for all the problems completed by each student.

6. Percentage of problems solved (Phillips, 1974). This dependent measure is a general indicator of performance effectiveness and has no relation to other efficiency measures. A problem was defined as being solved if the student demonstrated five consecutive correct responses and corresponding correct statements of the hypothesis. This percentage was calculated from the ratio of the number of problems solved to the total number attempted.

7. Trial of last error (Phillips, 1974). More efficient problem solvers will use more efficient solution strategies and be more efficient in their information processing. As a consequence, they will solve the problem in fewer trials. Overall efficiency was calculated by determining the mean trial of last error by considering all the student's problems.

8. Replacement of disconfirmed hypothesis (Kemler, 1978). The more efficient problem solvers will only test an hypothesis once, and if it is disconfirmed, reject it for the remainder of the problem. This measure assessed the student's memory for recalling which hypothesis had been disconfirmed. The variable was calculated by counting the number of different hypotheses stated in a problem. From this set, a count is made of the number of disconfirmed hypotheses which were repeated. The ratio is calculated by dividing the number of repeated disconfirmed hypotheses by the total number of disconfirmed hypotheses. For example, consider that a student was wrong on four of the five hypotheses he/she tested and that he/she later repeated three of the four disconfirmed hypotheses. The ratio is three-fourths (75%). The mean value was calculated from all the problems completed by a student.

9. Mean length of time for problem solution (Nuessle, 1972). Efficiency has been measured through various response patterns (e.g., maintaining confirmed and disconfirmed hypotheses, replacing disconfirmed hypotheses). Analysis of the length of time required to solve a problem indicated the relationship of this temporal measure to other performance indicators. The

total time for a given problem was the time of recording the student's response to the first trial to recording the student's last hypothesis statement in the last trial of the problem. The mean time was calculated from these intervals for all of the student's completed problems.

10. Latency period following feedback (Levine, 1969; Nuessle, 1972). This second rate measure was described by Levine as reflecting the information processing abilities of a student, especially as it relates to his/her retrieval and recoding of information. One would expect a shorter interval between trials after a hypothesis is confirmed and a longer interval between trials after the hypothesis was disconfirmed. The longer time interval reflects the time required for recoding the logical set of potentially correct hypotheses and selecting a single hypothesis.

Across all problems administered to a subject, the latency periods were split into two groups: (a) those following correct choices, and (b) those following incorrect choices. Within each grouping, the mean latency interval was calculated.

Experimental design. The nested design (2 x 2 x 21) had three factors: (a) student grouping--LD or NLD, (b) treatment--experimental group or control group, and (c) the matched subjects.

The experimental procedures were administered twice. During the first session, all students received the same treatment, and data collected during this session were used as covariates in the statistical analysis. One of the dependent variables (i.e., mean trial of last error) was used as a blocking variable for assignment to the treatment group for the second session.

Interrater reliability. Reliability measures were obtained by having the principal investigator listen to tape recordings of five sessions conducted and recorded by the research assistants. While listening to the tape, the investigator independently recorded the student's verbal responses (i.e., which stimulus he/she chose, the student's hypothesis statements, and the researcher's taped feedback to the student). The two records were compared item-by-item and an agreement was scored each time the same response was recorded on a given trial. A disagreement was scored each time different responses were recorded on a given trial. Next, the accuracy of the feedback to the student was determined by matching correct responses with reinforcing statements and incorrect responses with corrective statements. An error was scored if a reinforcing statement was made to an incorrect response or a corrective statement was made to a correct response.

Finally, the reliability of the measurement of each subject's response consistency was checked by having the investigator listen to and record whether or not each student's subsequent response was consistent with the previously stated hypothesis. The records were compared trial-by-trial; an agreement was scored if both records said the student's response was consistent or if both said the response was inconsistent. A disagreement was scored if one record said the response was consistent and the other said it was inconsistent. Table 2 shows the results of these reliability measures. The results indicate that the raters agreed on 99% of the trials regarding correct/ incorrect responses, on 99% of the trials regarding hypotheses, and on 99% of the trials regarding the recording of consistency of response. Researcher 1 was 99% accurate in providing feedback, while Researcher 2 was 100% accurate.

Statement of Hypotheses

Based on the review of the related literature and the methodological procedures described, the following null hypotheses were tested for each of the dependent variables.

Hypothesis 1: Subject group X treatment interaction

The learning disabled students in experimental treatment will demonstrate the same level of performance as subjects in the control group.

Hypothesis 2: Treatment groups

Students in the experimental treatment group will perform at the same level of proficiency as subjects in the control group.

Hypothesis 3: Student grouping

LD students will perform at the same level of proficiency as the non-learning disabled subjects.

Statistical analysis. The hypotheses were tested with multivariate analysis of covariance procedures (Finn, 1976). Wilks' likelihood ratio criterion (Λ) was calculated for the hypotheses tested. Rao's approximate multivariate *F*-test evaluated Wilks' criterion at the .05 probability level for statistical significance and for rejecting the null hypothesis. The scores from the first session on each dependent variable and the student's IQ scores were the covariates. Post hoc procedures evaluated which dependent variables contributed significantly to the interaction and main effects of the factors.

Results

The first and third null hypotheses were not rejected (Table 3). For the first hypothesis, after controlling for students' scores during the first session, the use of the reinforcement conditions did not differentially affect any group of students. The improved scores of the second session were not unique to any one group; all students demonstrated some improvement.

The third hypothesis tested differences between the LD students and their non-LD peers. The multivariate analysis of covariance procedures adjusted the second-session scores of each group based on their performances during the first session. Based on this adjustment, no statistically different scores existed between the LD students and their non-LD peers. The second session did not differentially affect either the LD students or the non-LD peers.

The second null hypothesis was rejected (Table 3). The students in the control group scored significantly better than those in the experimental group. To determine the dependent variables on which the two groups differed, further followup statistical tests were completed according to the procedures described by Finn (1974).

The results of the post hoc analysis indicated that there is a reliable difference between the experimental group and the control group. This difference is best evidenced in the two groups' differences on the response consistency dependent variable. The control group was more consistent in choosing responses identical to their hypothesized choice ($\bar{x}$ = 97.68) than was the experimental group ($\bar{x}$ = 95.31), who were reinforced for each correct answer and problem solved.

Validation of the Process Dysfunction

By analyzing the second-session data with the multivariate analysis of covariance procedures, no differences were calculated on the dependent variables between the LD and non-LD students. However, numerous differences on the dependent variables were measured on first-session scores. The lack of significant results with the covariance analysis indicated that the reinforcement conditions made no difference with the LD and non-LD groupings. It may be concluded that the LD students were not passive learners and did not exhibit a production deficiency. This finding is contrary to Torgesen's (1977) hypothesis. To determine the LD students' specific processing dysfunctions, a multivariate analysis of variance procedure was used. This analysis compared the LD students' second-session scores to those of their NLD peers. The results of the analysis indicated that the non-LD students obtained statistically better scores than the LD students. Specifically, the post hoc tests determined that nine variables reliably distinguished the two groups. The variable which best discriminated between the groups was mean trial of last error. With this variable, 68% of the LD and regular class students were correctly classified. Twenty-six of the 42 LD students were classified as LD (62%) and 31 of the 42 non-LD students were classified as non-LD (74%). The mean trial of last error for the LD students was 8.33 (SD = 2.09) and 6.38 (SD = 1.85) for the NLD students.

Discussion

Conclusions

The principal research question addressed in this study was whether the LD students' performance reflected a production deficiency or a process dysfunction. Further, if a process dysfunction was determined, the study was designed to ascertain which components of an information processing model were affected. Torgesen's (1977) hypothesis of production deficiency was not supported. The results of this study show that LD students have specific processing differences when compared to their non-LD peers. These findings were replicated twice by the data. First, in the initial session reliable differences were found between the LD and the non-LD students. Second, under conditions of reinforcement and response cost in the second session, processing differences were also found. If the LD students had indeed been passive learners as Torgesen (1977) has suggested, the processing differences would have been eliminated in the second session. By contrast, the students in the present study appeared motivated by the task in both sessions, and yet, the results showed that the LD students as a group appeared to have more difficulty in solving the problems.

An analysis of the second-session scores determined statistical differences between the LD and non-LD students on nine variables. Ranked in the describing hierarchical order of the highest differences, these variables are: (a) trial of last error (TLE), (b) resampling of disconfirmed hypotheses (DCH), (c) percent correct (PC), (d) problem solution time (MST), (e) the percentage of problems solved (PPS), (f) retaining confirmed hypotheses (RCH), (g) latency period following a confirmed hypothesis (LPFC), (h) retaining disconfirmed hypotheses (RDH), and (i) response consistency (RC).

In assessing students for learning disabilities, information on one variable appears to serve as a discriminator: mean trial of last error. Phillips (1974) considers TLE as an overall measure of a student's problem-solving efficiency. A student whose strategic behavior maximally incorporates the available information will solve the problem in fewer trials than one who does not. General efficiency most clearly distinguishes LD from non-LD students. Correlated with TLE is MST ($\underline{r}$ = .725). MST is a temporal indicator of problem-solving efficiency (Nuessle, 1972) and signals that efficient solutions also require the least time. The LD students are requiring additional trials and time to solve the problems and have a longer latency period following a confirmed hypothesis. If one were to hypothesize differences between LD and non-LD students, differences would not likely be on LPFC. Latencies following confirmed hypotheses would be expected to be short, while latencies following disconfirmed hypotheses would be longer (LPFD) (Levine, 1969). Given the reported processing problems of the LD student, a difference on LPFD would be expected; yet the above data did not support such a conclusion. (p = 0.273). Presently, the differences on LPFC cannot be fully explained; they are likely due to LD students' coding, recoding and rehearsal of the confirmed hypotheses.

The other dependent variables on which significant differences were found are also associated with the coding, recoding, and rehearsal of information. The second largest performance difference between LD and non-LD students was found on DCH, resampling disconfirmed hypotheses. The LD students were more likely to resample a hypothesis even though it had been previously disconfirmed. This failure to profit from feedback is also evidenced by the significant differences on RCH, retaining a confirmed hypothesis, and RDH, retaining a disconfirmed hypothesis. The data suggest that even explicit feedback concerning the quality of an answer does not consistently benefit the LD student's performance either immediately or within the course of completing the same problem. The concepts of correct ("Yes, this picture has the answer in it") or wrong ("No, this picture," pointing to the alternate to the one selected, "has the answer in it") do not necessarily alter an LD student's task responses. One might speculate that the LD student has externalized the cause of his/her successes and failures to such an extent that chance is operating and the examiner's feedback is ignored. The validity of this assumption seriously undermines the extensive efforts made in training the students and demonstrating possible solutions. It may be concluded that the LD student's behavior lacked the strategic approaches of their non-LD peers.

This conclusion is also supported by the differences between the LD and the NLD groups determined on the response consistency (RC) variable. Just as the experimental group was less consistent in their responses than the control group students, the LD students were found to be less consistent than their non-LD peers. In practice this means that they would verbally <u>state</u> an hypothesized solution and yet <u>choose</u> a different solution. They did not trust their beliefs, nor were they willing to test them in a manner different from non-LD students.

As a consequence of the noted processing errors, the LD students were slower to learn as quantified in the percentage correct variable. A sequence of any four trials provided a sufficient amount of information to arrive at a problem's solution. By subtracting the initial four trials of a problem from the total trials of a problem, one could determine how much information was gained from the first four trials. The higher the percentage of correct answers, the more information the student gained from the initial trials. The LD students demonstrated significantly poorer performance on this PC variable than did the non-LD students.

In summary, the LD students demonstrated significant processing differences compared to their NLD peers in their ability to code, recode, and recall information. They also responded to the trials in a manner suggesting less efficient strategic behavior and a failure to profit from explicit feedback concerning their answers. They were much more likely to take chances on possible solutions rather than following a logical, strategic problem-solving pattern. Thus, they were most clearly distinguished from their non-LD peers by number of errors and length of time required in problem solving. As measured by Torgesen's hypothesis (1977), the LD students did not appear to be passive problem solvers, but very inefficient as reflected in processing and strategic differences discussed above.

<u>The Effect of Reinforcement</u>

The control group members, who were paid a standard fee for participating, demonstrated superior performance compared to that of the experimental group, whose members were paid for each correct answer and problem solved, and who lost money for incorrect answers. Rather than improving performance, the reinforcement/response cost condition had a deleterious effect. This finding is counter to those of Haines and Torgesen (1979) and of Hallahan et al. (1978). However, results similar to those obtained in this study have been reported by Spence (1970).

The procedures of the second session clearly and repetitively presented information about the reinforcement and response cost conditions. On the fifth trial and each trial afterward, all students received a chip for correct answers or lost a chip for incorrect answers. At the end of each problem, students in the experimental group were told how much they had earned for the particular problem. All students were paid at the end of their second session.

It is reasonable to expect that the variable representing the greatest mean group difference and having the highest weight in the discriminant analysis was the response consistency variable (RC). By taking extra chances and gambling on the correct solution, the experimental group's performance deteriorated. Being correct meant more money and apparently the effort to do so distracted them from efficient processing. As Spence observed, reinforcement procedures on a complex task that already requires extensive cognitive effort do not necessarily improve performance. It should be noted that this affected the LD and non-LD students in the same manner. Clearly, reinforcement and response cost schemes are selected for the purpose of changing behavior, but the direction of the change may interact with the selected procedures.

Limitations

Future researchers should carefully weigh the advantages and disadvantages of recording data on a DATAMYTE. Data collection was greatly facilitated by its use, but numerous problems arose immediately afterward. The two most time-consuming and costly tasks were editing the extensive files of data and preparing computer programs for calculating each of the dependent variables. Based on these considerations, paper-pencil response recording and manual data entry would be preferable. The exceptions might be if one were specifically interested in temporal measurements, simply calculated dependent variables, or in repeating measurements over an extended period of time.

Whenever research subjects represent a special population such as the learning disabled, the generalizability of the findings becomes a concern. The subjects in this study were not randomly selected, but were requested to participate. The few refusals indicates the high degree of willingness to participate. Regardless, the validity of these findings requires careful attention as to the marker variables of the participating students.

Two of the variables, response consistency and retention of confirmed hypotheses, had mean group scores in the mid- to upper- 90th percentile ranges. Ceiling effects lend themselves to possible statistical regression. Also, percentile scores are frequently transformed using an arcsine transformation. Such a transformation was intended for these data, but was not completed.

Perhaps the most serious question relates to whether or not the results are experimental artifacts. The task and its procedures placed numerous cognitive demands on the student. Additional research would determine whether these demands were unique to the discrimination learning task or were common to a student's academic learning experiences.

Design Limitations

The internal validity is threatened by the contaminating effects of the students' interaction with each other. In two unknown instances, two students reported their experiment earnings to classmates despite the experimenter's caution not to discuss the experiment with classmates for several days. The results of this contamination are unknown. All students did know that they would receive payment for participation.

Queries from the students indicated a concern about the use of the DATAMYTE in the experiment. Whether this distraction significantly influenced performance is unknown. The instrumentation is judged preferable to a paper-pencil recording procedure using a stop watch to measure time intervals. Hopefully, the distraction of recording responses was minimized by the verbal feedback ("correct" and "wrong") to the student's responses and by answering any questions about the DATAMYTE.

The students completed a number of experimental problems that number varied according to the student's efficiency at solving the problem and the length of the class period used for testing. Hypothetically, the probability of solving one problem was constant as was the cumulative effect of a series of problems. The number of problems administered should not have affected the probability of solving a problem, but may have had a facilitating effect.

An additional design problem concerns the reliability of the temporal measures: mean solution time and the latency period following the confirmed and disconfirmed hypothesis selection. Time measures were recorded automatically by the DATAMYTE as data were entered by the researcher. The researcher was cued by the pointing response of the student, which was not recorded on the audio cassette. Since the student's pointing response was not reproducible based on the recording, reliability measures on the researcher's scoring of this variable were not retrievable.

Implications and Future Research

The results of the present study indicated that the LD adolescents were not passive learners on this task; they demonstrated significant processing and strategic behavior differences from their non-LD peers. These results lead to two obvious consequences: (a) LD adolescents may have difficulty learning regardless of the content, e.g., strategic behaviors, math facts, functional skills, or biology; and (b) the mode of instruction needs to be carefully evaluated. For example, explicit corrective feedback ("correct" or "wrong") was more frequently ignored by the LD students as a group. If educators are going to attempt to meet the individual needs of each student, a reliable, valid procedure for identifying processing differences relevant to a school curriculum needs to be determined. Thus, future research should identify a students' specific processing deficits and correlate them with academic performance. Such careful study will permit a match between a learner's aptitudes and particular instructional models (Glaser & Resnick, 1972). Identification of process deficiencies is a fruitless effort if they have no particular consequence for instructional procedures.

The LD student's unsystematic behaviors require further evaluation. One would expect a change in behavior as a consequence of instructional feedback, and yet such was not apparent from these LD students. Constructs such as locus of control, attribution theory, and success expectancy warrant further study with LD adolescents. This interpretation is supported by the previous research of Adelman (1978) and Pearl, Bryan, and Donahue (1980), who identified LD students as demonstrating lower levels of intrinsic motivation and fewer feelings of internal control over success.

The role of reinforcement and punishment procedures in problem solving also requires careful consideration. While the facilitative effects of rewards are well known, the debilitating effects need further study. These results should caution practicioners that complex, cognitive tasks may require a student's full attention and that motivational strategies need to be carefully controlled. One needs to consider whether or not the motivational techniques have a net effect of reducing or increasing the cognitive strain in problem solving. Level of arousal has been shown to have differential effects on short-term and long-term memory (Kesner, 1973).

A major consequence of this research is the observsation of numerous qualitative differences between the LD students and the non-LD students. In the classroom and in future research, such learner characteristics must be considered along with task demands. Through future efforts of replication and cross validation, the profile of the LD adolescent needs to be more accurately tested and refined. A logical followup study would be the validation of the processing dysfunction with a specific academic task. Ideally, the academic task would bear a close relationship to the cognitive processes and strategic skills required in the discrimination learning task. Such a test is needed to assess the ecological validity of the processing dysfunction vs. the production deficiency models applied to the LD learner.

REFERENCES

Adelman, H. S. The concept of intrinsic motivation: Implications for practice and research with the learning disabled. Learning Disability Quarterly, 1978, 1, 43-54.

Baumeister, A. A. Problems in comparative studies of mental retardates and normals. American Journal of Mental Deficiency, 1967, 71, 869-875.

Byrd, D. The effects of mnemonic training upon problem-solving strategies as a function of developmental memory level. Unpublished master's thesis, Memphis State University, 1979.

Cook, T. D., & Campbell, D. T. Quasi-experimentation: Design and analysis issues for field settings. Chicago: Rand McNally, 1979.

Finn, J. D. A general model for multivariate analysis. New York: Holt, Rinehart, & Winston, 1974.

Finn, J. D. Multivariance: Univariate and multivariate analysis of variance, covariance, regression. Version 5, Release 3. Chicago: International Educational Services, 1976.

Flavell, J. H. Developmental studies of mediated memory. In L. P. Lipsitt & H. W. Reese (Eds.), Advances in child development and behavior, Vol. 5. New York: Academic Press, 1970.

Gholson, B. Developmental stages, hypothesis theory, and individual differences. Paper presented at the NATO International Conference on Cognitive Psychology and Instruction, Amsterdam, Holland, June 1977. (ERIC Document Reproduction Service No. ED 143 695).

Gholson, B., & Danziger, S. Effects of two levels of stimulus complexity upon hypothesis sampling systems among second and sixth grade children. Journal of Experimental Child Psychology, 1975, 20, 105-118.

Glaser, R., & Resnick, L.B. Instructional psychology. In P. Musson & M. Rosenzweig (Eds.), Annual review of psychology (Vol. 23). Palo Alto, Calif: Annual Reviews Inc., 1972.

Haines, D. J., & Torgesen, J. K. The effects of incentives on rehearsal and short-term memory in children with reading problems. Learning Disability Quarterly, 1979, 2, 48-55.

Hallahan, D. P., Gajar, A. H., Cohen, S. B., & Tarver, S. G. Selective attention and locus of control in learning disabled and normal children. Journal of Learning Disabilities, 1978, 11, 231-236.

Hallahan, D. P., Tarver, S. G., Kauffman, J. M., & Graybeal, N. L. A comparison of the effects of reinforcement and response cost on the selective attention of learning disabled children. Journal of Learning Disabilities, 1978, 11, 430-439.

Kemler, D. G. Patterns of hypothesis testing in children's discriminative learning: A study of the development of problem-solving strategies. Developmental Psychology, 1978, 14, 653-673.

Kesner, R. A neural system analysis of memory storage and retrieval. Psychological Bulletin, 1973, 80, 177-203.

Levine, M. Neo-noncontinuity theory. In G. Bower & J. T. Spence (Eds.), The psychology of learning and motivation (Vol. 3). New York: Academic Press, 1969.

Levine, M. A cognitive theory of learning research on hypothesis testing. Hillsdale, New Jersey: Laurence Erlbaum, 1975.

Mims, R. M., & Gholson, B. Effects of type and amount of feedback upon hypothesis sampling systems among 7- and 8-year-old children. Journal of Experimental Child Psychology, 1977, 24, 358-371.

Neimark, E. D. Intellectual development during adolescence. In F. D. Horowitz (Ed.), Review of child development research (Vol. 4). Chicago: University of Chicago Press, 1975.

Nuessle, W. Reflectivity as an influence on focusing behavior of children. Journal of Experimental Child Psychology, 1972, 14, 265-276.

Pearl, R., Bryan, T., & Donahue, M. Learning disabled children's attributions for success and failure. Learning Disability Quarterly, 1980, 3, 3-9.

Phillips, S. Introtacts in children's discrimination learning. Unpublished doctoral dissertation, State University of New York at Stony Brook, 1974.

Phillips, S., & Levine, M. Probing for hypotheses with adults and children: Blank trials and introtacts. Journal of Experimental Psychology, 1975, 104, 327-354.

Phillips, S., Levine, M., & O'Brien, J. T. Transfer in discrimination learning: Do children mediate more than adults? Journal of Experimental Child Psychology, 1978, 26, 161-173.

Richman, S., & Gholson, B. Strategy modeling, age, and information processing efficiency. Journal of Experimental Child Psychology, 1978, 26, 58-70.

Spence, J. T. The distracting effects of material reinforcers in the discrimination learning of lower- and middle-class children. Child Development, 1970, 41, 103-111.

Stanovich, K. E. Information processing in mentally retarded individuals. In N. R. Ellis (Ed.), International review in mental retardation. New York: Academic Press, 1978.

Tarver, S. G., Hallahan, D. P., Cohen, S. B., & Kauffman, J. M. The development of visual selective attention and verbal rehearsal in learning disabled boys. Journal of Learning Disabilities, 1977, 10, 491-500.

Torgesen, J. K. Performance of learning disabled children: A theoretical assessment. Journal of Learning Disabilities, 1977, 10, 27-34.

TABLE 1

Marker Characteristics of Subjects

		Learning Disabled	Non-Learning Disabled	Total
Sex	Male	32	32	64
	Female	10	10	20
Age (months)	Mean	171.64	170.19	170.91
	S.D.	11.73	8.85	10.36
Grade Level	6	1	0	1
	7	8	7	15
	8	20	21	41
	9	13	14	27
Mean IQ		85.78	107.11	96.49
Standard Deviation		11.47	11.83	15.54
Number		42	42	

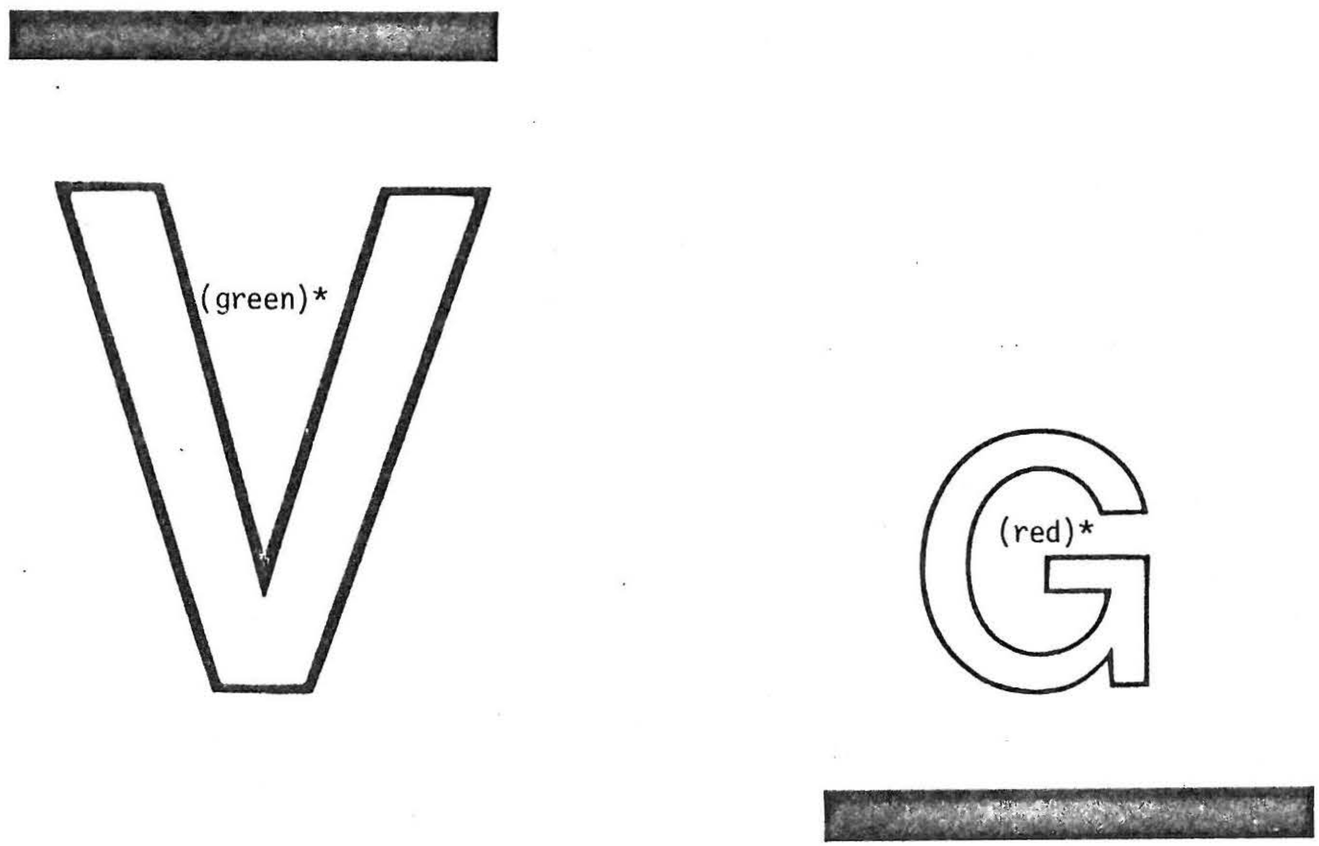

FIGURE 1

Four-Dimensional Stimulus Pattern

*The color of the first letter (V) was green; the color of the second letter (G) was red

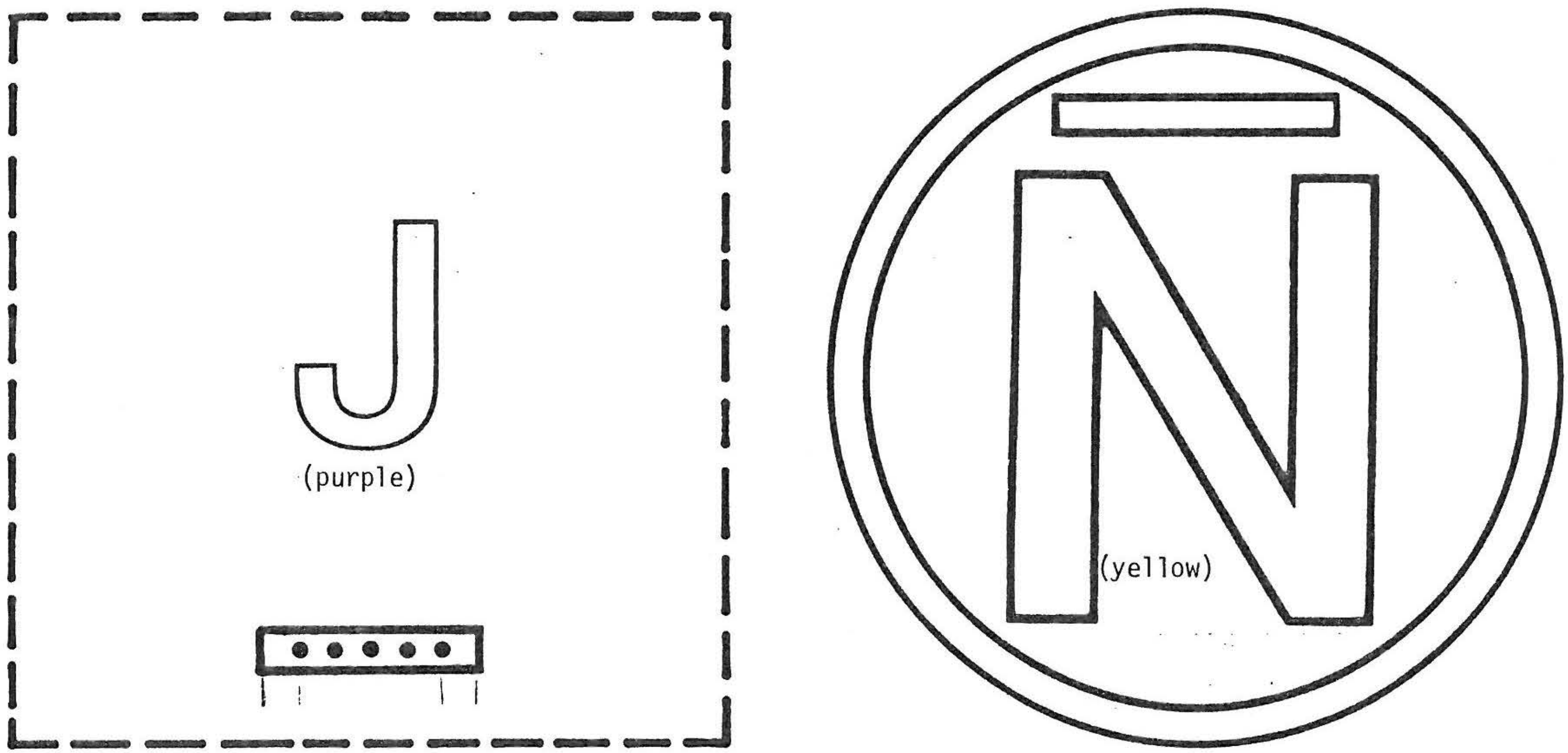

FIGURE 2

Eight-Dimensional Stimulus Pattern

TABLE 3

MANOCOVA Summary Table

Source	*df*	Likelihood Ratio Criterion	Rao's *F* Transformation	P-Value
Grand mean	1	---	---	---
Group (LD and non-LD)	1	.554	1.3167	0.2918
Subjects within group (error)	40	---	---	---
Treatment (experimental and control)	1	.369	2.7945	0.0258
Group x treatment	1	.742	.5674	0.8306
Treatment x subject with group (error)	40	---	---	---
Total	84			

TABLE 2

Inter-Rater Reliability Values

Examiner	No. of Subjects	Total Number of problems trials		Student's Hypothesis		Feedback		Consistency	
				Errors	%age of Agreement	Errors	%age of Agreement	Errors	%age of Agreement
1	5	47	573	2	99%	1	99%	2	99%
2	5	42	471	1	99%	0	99%	5	99%
Total	10	89	1044	3	99%	1	99%	7	99%