AN EVALUATION OF THE IDENTIFICATION OF LEARNING DISABLED STUDENTS IN COLORADO

Lorrie A. Shepard and Mary Lee Smith

Abstract. **The purpose of the study was to evaluate the process whereby children are identified as learning disabled (LD) and the validity of subsequent educational placements: (a) a representative sample of 1,000 cases were analyzed both quantitatively and qualitatively; (b) extensive questionnaires were administered to professionals involved in LD identification; (c) definitions and assessment procedures were evaluated. Only 28% of the LD cases met strict criteria for LD; another 15% showed weak signs of the handicap. The remaining 57% currently in LD programs were better described by other indicators such as non-English dominant (7%), other handicaps (10%), slow learners (11%), minor behavior problems (4%), etc. The costs of LD identification were found to be excessive; each year almost half of the resources available for LD pupils was spent on assessment and staffing. Policy implications are discussed.**

The purpose of the study was to evaluate the procedures and results of the identification of children with learning disabilities (LD)[1] in Colorado. The identification process was meant to include the following steps: referral, assessment, staffing, and placement. The most important question to be addressed was, "Are the LD placements valid?"

The problems inherent in the identification of LD are widely recognized. First, the lack of an adequate definition reflects the continued debate about the theoretical nature of the construct. Thus, although the National Advisory Committee's (1968) definition incorporated in PL 94-142 reflected the best professional consensus at the time, it is vague and ambiguous (Hammill, 1974). Further, the promulgation of this definition may, in fact, create and perpetuate some misconceptions about the nature of LD (cf. the NJCLD position paper, 1981; Hammill, Leigh, McNutt, & Larsen, 1981). The definitional problems compound the measurement problems that exist whenever psychologists attempt to assess a construct. Practical guidelines for identification cannot be made unambiguous or unequivocal. For example, Thurlow and Ysseldyke (1979) stated that criteria used in identifying learning disabled children are both "highly variant and nebulous" (p. 2).

To date, limited attention has focused on the type of research undertaken in the present investigation, that is, determining not only if the school-identified cases of LD show sufficient evidence of LD, but also the definitional, assessment, and staffing practices associated with the validity of placement. In the past, researchers took the population of LD as given and proceeded to study the characteristics of the subjects in this category. For example, in a widely disseminated article, Clements (1966) summarized the symptoms of minimal brain dysfunction

LORRIE A. SHEPARD, Ph.D., is Associate Professor and Chairman, Research and Evaluation Methodology, School of Education, University of Colorado.

MARY LEE SMITH, Ph.D., is Associate Professor, Research and Evaluation Methodology, School of Education, University of Colorado.

Shepard, L.A. & Smith, M.L. (1983). "An Evaluation of the Identification of Learning Disabled Students in Colorado." Learning Disabled Quarterly, 6: 115-127.

which he gleaned from one hundred publications; he did not, however, seek any evidence as to which indicators discriminated between normal and LD children. Authors who have examined the characteristics of children in this category find that LD pupils are enormously diverse and often do not meet definitional criteria (Kirk & Elkins, 1975; Norman & Zigmond, 1980).

The recent work of Ysseldyke and his colleagues is exceptional because these authors attempted to test not only whether LD students evidenced "LD characteristics", but also whether they were distinguishable from other underachievers. Thus, Ysseldyke, Algozzine, Shinn, and McGue (1982) found a 96% overlap between the scores of LD and underachieving students on 49 psychometric variables. Based on a comparison of the characteristics of LD children and the federal definition, the authors concluded that as many as 40% might be misclassified. Similarly, numerous studies and reports suggest that many nonhandicapped children may be misidentified in the LD category (Poplin, 1981). However, to date, a systematic attempt has not been made to determine what subgroups of the LD school population are validly placed and what these students' characteristics are.

The present study was mandated by the Colorado state legislature and commissioned by the Colorado Department of Education. Among the political factors leading to the commissioning of this research were the steadily increasing numbers of students identified as LD (Davis & Smith, 1981). Key legislators on the Colorado House Education Committee were also familiar with variability in criteria from district to district and, therefore, questioned the validity of increasing placements in the LD category.

The present evaluation study was designed to examine the identification process and the validity of placement decisions derived from that process. The following specific questions were developed to guide the research:

1. What definitions and criteria guide the identification of LD pupils?
2. How many children are identified as LD? Are differences in the percentages of pupils identified across districts associated with differences in the characteristics of LD children?
3. What tests are used in the assessment of LD children? Are the tests reliable, valid, and fair?
4. What is the role of clinical judgment in identifying children as LD? How valid are these judgments?
5. What are the characteristics of children who are called LD? Can a potentially diverse population of pupils be described in terms of more homogeneous, identifiable subgroups? How valid are the diagnoses of LD?
6. What costs are involved in meeting both legal and professional requirements for identification of LD children? Are increments in costs associated with increases in validity?

METHODS

The evaluation design had several parts including two large-scale surveys. The first consisted of an analysis of the case files of individual pupils currently classified as LD. The second was a survey of professionals involved in the identification process. Supplemental investigations included: logical analysis of federal, state, and local criteria for identifying LD, re-analysis of existing state data for correlates of prevalence rates, structured interviews with directors of special education, and a cost analysis of the identification process. Complete details of the methods and analyses are given in the final evaluation report (Shepard & Smith, with Davis, Glass, Riley, & Vojir, 1981).

Study of Pupil Case Files

A probability sample of 1,000 LD pupil files (3.8% of the Colorado population) was selected.[2] This sample was randomly subdivided into 200 cases to be used in a qualitative study and 800 for quantitative analysis (10 cases were lost due to clerical errors or logistical problems).

Quantitative analysis. The 790 cases selected for quantitative analysis were read on site by trained coders. An eight-page case file record was developed for coding information such as the following:

- demographic characteristics; e.g., age, race
- program characteristics; e.g., number of hours
- referral information; by whom and reasons
- previous history in special education

- additional services
- who attended staffings and wrote reports
- narrative and coded basis of handicap
- tests administered
- IQ and achievement scores
- evidence of processing disorders/by specialist
- behavioral indicators
- other sources of academic problems
- medical indicators
- exit criteria

The first 100 cases were read again at the end to check for drift in the ratings. Results of a reliability study revealed a 96% concordance rate among the three coders.

Qualitative analysis. The 200 case files selected for qualitative analysis were copied[3] and read by the authors. In carrying out this analysis, we followed the procedures of Glaser and Strauss (1967) who suggested that large amounts of narrative data can be analyzed by deriving categories (patterns or variables) from sections of the data, refining and defining them with other parts of the data, and finally deriving hypotheses that explain the relationships among categories.

Having studied 20 files to gain a general understanding, we derived five research questions (what Glaser and Strauss call *categories*) and used them to investigate the remaining case files: (a) *statistical significance of the discrepancy;* (b) *consistency* (evidence that when different clinicians evaluated the same child they produced consistent, reliable patterns of that child's abilities and behavior or acknowledged the contradictions and attempted to resolve them in favor of the more consistent or valid data sources); (c) *marginality of placement* (whether, based on the significance of the ability/achievement discrepancy, the history of the case, and the patterns of ability test scores, the child's problem was severe or mild); (d) *necessity for special education;* (e) *cluster* (subcategories we referred to as clusters or "implicit definitions" of learning disabilities; these included subgroups of clinical LD as well as non-LD groups better characterized by labels such as "slow learner" or "second language interference"). Methods of triangulation and assumptions of the qualitative analysis are discussed further in Chapter 2 of Shepard and Smith (1981).

The results of the qualitative analysis of cases were presented in two forms. First, a table of frequencies of the cases falling into each category was presented along with findings and hypotheses from the entire analysis. Second, 10 case histories were written (Smith, 1981) of students whose LD files were selected from the 200 qualitatively analyzed. These 10 comprise a quota sample, chosen to be typical (not representative in the statistical sense) of the range of cases. The case histories are not included in this abbreviated article but can be read in the full report.

Survey of Professional Opinions and Practices

Representative samples of 674 LD teachers, 217 social workers, 176 school psychologists, 240 speech/language specialists, and 499 school principals were selected and surveyed by mailed questionnaires. Return rates ranged from 74% to 83%. In addition to followup mailings, non-returns in an a priori 20% random subsample (called the *core sample*) were contacted by phone to obtain answers to key questions or to learn the reason for non-response. Late return and core sample versus non-core returns were examined to estimate the differences in opinions reflected in the non-responses.

Five separate questionnaires[4] (from 6 to 11 pages in length) were tailored for the groups of professionals who typically participate in assessment and staffing. The principals' questionnaire was the shortest and focused primarily on policy issues. These subjects were asked questions regarding the definition of LD (PCD), the identification process (including various professionals' influence and parental participation), funding formulae, costs (measured in personnel time), and exit criteria. The specialists' questionnaires included all of the above issues except those pertaining to fiscal policy. In addition, these surveys included questions elaborating on the definition of LD and operational indicators. School psychologists, speech/language specialists, and LD teachers were given lists of tests and were asked to indicate how frequently they used each test as well as their rating of the reliability and validity of each.

Interviews with Directors of Special Education

A representative sample of 22 directors of special education were interviewed using a formal protocol. Interviews lasted from 50 to 90

minutes. Some directors preferred to respond in writing to some questions. Detailed questions and standard probes addressed the following general issues: definition of LD and operational criteria, incidence of LD (causes of over- and underidentification), the funding formula for reimbursement from the state, assessment and staffing procedures, and types of instructional interventions for LD students.

Other Substudies and Data Sources

To supplement the two large-scale studies of LD pupils and professionals, information was also gathered from other sources. Existing data on the prevalence of various handicaps from the Colorado Department of Education were used as well as comparable nationwide data from USOE publications. Existing state fiscal data were especially important for crosschecking the cost analysis of the identification process. A conceptual analysis was made of the essential elements in the Colorado and federal definitions of LD. Written documents obtained from sampled districts were analyzed following this conceptual framework.

ANALYSIS AND RESULTS

Definitions of Learning Disabilities

The Colorado definition of perceptual or communicative disorders (PCD) closely parallels the federal definition of specific learning disabilities. Although the Colorado label might suggest some distinction between PCD and LD, the essential elements in both definitions are: a significant discrepancy between intellectual ability and achievement; an inferred dysfunction in one or more of the basic psychological or learning processes; and an exclusion (i.e., ruling out) of other known causes of learning difficulties such as mental retardation, visual or hearing handicaps, emotional disturbance, or cultural deprivation. The limitations of this definition are well documented (Hammill et al., 1981). Nevertheless, for the purposes of the evaluation, it was important to know what policies were in effect (school year 1980-81) and to what extent the existing formal criteria guided placement decisions. The analysis of formal documents from local districts revealed little deviation from the state definition and suggested criteria. Therefore, the large differences later observed in numbers and severity of LD from district to district were more likely due to informal policies and the beliefs of the individual professionals involved.

Questionnaire results indicated considerable variability among professionals in the extent to which their own views matched the legal definitions of LD. The definitional elements with the greatest agreement were those pertaining to significant discrepancies and psychological-neurological deficits. In contrast, the greatest divergence of professional opinion was found in the interpretation of exclusionary signs. For example, from 12% to 24% of professionals (psychologists and LD teachers, respectively) noted that linguistic differences were a positive rather than a negative indicator for LD. These professionals did not consider language differences as a possible reason for exclusion (as the theoretical construct implies). Responses to this question paralleled those for cultural deprivation, socioeconomic disadvantage, excessive absences, and inadequate teaching. In each instance, up to a quarter or more of the professionals surveyed were willing to count *for* the disorder signs that should have argued *against* it; i.e., evidence that a student's problem was due to poor teaching would for these professionals support the diagnosis of LD. Also, a small group of professionals (approximately 10%) interpreted short attention span and general poor health as *sufficient* evidence of LD.

Prevalence

The percentage of LD pupils (of the total number of children enrolled in school) in Colorado has increased from 4.2% in 1975-76 to 5.1% in 1979-80, an increase of 21%. Such an increase is not likely to be due to demographic changes in the state since such changes would also affect other categories of handicap. Instead, LD has grown as a percentage of all handicaps from 36.7% to 46.7% over the same time period. Colorado's percentage of learning disabled children is slightly above the national average which is also increasing.

Special education units throughout the state vary widely in the percentage of their enrollments identified as LD, from 2.11% to 8.56% in 1979-80. It was suggested that this variability across units could be interpreted in two ways: (a) either there are true differences in the incidence of the disability or (b) differences are due to local policies and practices that systematically and arbitrarily produce varying

rates of identification. Findings from national studies (Thouvenelle, Rader, & Hanley, 1982; Ysseldyke et al., 1982) support the latter conclusion.

Correlational analyses were conducted to determine what district characteristics were associated with high or low prevalence rates. Many of the correlations among district characteristics were uninformative. Since LD children account for nearly half of all handicapped children, it is not surprising that the percentage of LD children was highly correlated with the total percentage handicapped. The percentage of LD students also has modest positive correlations with the speech and emotionally disturbed categories. An interesting finding for one school year was a negative correlation between percentage LD and both percentage EMR and socioeconomic level (of the district) suggesting that the LD label may be used for some EMR children in high socioeconomic districts. Not much was made of this finding, however, because it did not hold up across years. A more stable and potentially illuminating finding was the modestly strong correlation ($r = .51$) between district size and percentage of valid LD placements; that is, large districts tended to have greater *percentages* of valid placements. As explained in later analyses, individual cases were called "validly LD" if they matched any one of seven legal or clinical criteria for LD. On the basis of anecdotal and interview data, we conjectured that large districts have higher percentages of valid placements because they tend to have more alternative programs such as Title I and bilingual education. Therefore, in large districts special education is more often used only as a last resort. This finding may have important implications for policy decisions intended to improve the validity of LD identification.

Assessment Practices

A profile of typical assessment practices was obtained both from LD pupil files and from the surveys of professionals. The data included simple counts of the tests given most often and the number administered to each student. They also included professionals' knowledge of test validity, knowledge of statistics used in test score interpretation, and professional opinions about the role of clinical judgment. The details of this component of the evaluation process are reported in a separate article by Davis and Shepard (1983). Therefore, only the most salient findings, especially those relevant to the cost and validity of LD placements, are listed below:

1. The number of formal tests given to identify LD pupils was often excessive;
2. Although some pupils were given too many tests, one-quarter of the population received too little testing;
3. Most of the tests used in the diagnosis of LD were technically inadequate;
4. Many clinicians were unaware of the technical inadequacy of the tests they used;
5. Specialists often selected technically inadequate measures even when more valid instruments were available. These choices tended to follow traditional habits associated with each professional group;
6. Substantial numbers of specialists did not understand simple statistics used in test score interpretation;
7. Although many specialists (28% to 44%) believed that test results should be secondary to clinical judgment in the diagnosis of LD, they often did not take basic steps to ensure the validity of their judgments.

Staffing Procedures

The staffing process comprises four distinct phases: referral, assessment, staffing conference, and placement. According to Colorado Department of Education data (1980), 42,195 pupils were referred for the staffing process during the academic year 1978-79 (the last year for which these data were available). Of this number, 78% (32,792) were assessed. Among those assessed, 80% (26,088) were staffed. In turn, 83% of those staffed were placed and received special services. Thus, slightly more than half of those originally referred were eventually served. As for the remainder of the students referred, one can speculate that either their problems were alleviated or spontaneously improved before the next phase in the staffing process took place, or they were judged (by preliminary assessments or in mini-staffing) to be ineligible for services. The above figures represent all handicapping conditions, not just LD.

Referral. According to the analysis of LD case files, the person most likely to make the referral for special help was the classroom teacher. Seventy-six percent of the LD cases (on which referral data were available) were thus referred. In 8% of the cases the child was re-

ferred by his/her parent.

Assessment. The assessment phase of the identification process was discussed in the previous section. Although the majority of cases involved six or seven tests (and were in compliance with regulations for unbiased assessment), there were also cases at the extremes for which no tests or excessive numbers of tests were administered. A relevant finding from the interview data was that in most districts the selection of tests is left to the discretion of individual professionals. To date, limited efforts have been made to control redundant testing.

Staffing conference. The federal and state regulations governing staffings are extensive. Despite some evidence of poor record keeping, we found that Colorado districts were largely in compliance with the rules as to who should be present at staffings. Thus, the typical LD staffing was attended by seven or eight professionals. A serious negative consequence is that parents were sometimes intimidated (reported by 70% of the professionals).

The typical LD staffing meeting lasted for 45 minutes with a reported range of from 15 minutes to 3 hours. According to our survey, more than half of the time is spent in determining "handicapping condition and placement," while a quarter of the time is spent writing the IEPs or planning instruction.

The majority of professionals involved in the identification of LD pupils are satisfied with the staffing process. They believe that the amount of time spent is sufficient to ensure accurate findings and satisfy due process requirements, but is not excessive or wasteful of professionals' time.

Characteristics of LD Pupils in Colorado

What are the characteristics of students who have been evaluated and placed as LD by the current procedures? How many of them match the officially legislated definitions of LD and the definitions of learning disabilities described in the professional literature? To answer these questions, we randomly subdivided the representative sample of LD case files and analyzed them using either quantitative or qualitative methods.

Quantitative analysis. The 790 coded cases were characterized using variables such as IQ, verbal-performance IQ discrepancy, weak and strict significant ability-achievement discrepancy, below-grade-level achievement, hyperactivity, major or minor behavioral problems, and non-English language dominance.[5]

As in previous studies (Kirk & Elkins, 1975; Norman & Zigmond, 1980), the characteristics of the LD population can be described by considering the above variables individually. Some of the more salient findings are presented here briefly. However, we give a minimum amount of attention to the results of each variable singly, because we believe that a more illuminating analysis is provided by *combinations* of variables in the next section.

Based on IQ test data important findings are the following: at the time of initial assessment and staffing, 26.8% of LD pupils were placed in LD programs without any IQ test data; 28.5% (or 39% of those with tests) had IQs below 90, i.e., below the average range; 8.3% had IQs of 80 or below.

Using the strict criterion for a significant difference ($\alpha = .05$), only 23% of the LD population demonstrated a significant discrepancy on even one of their math or reading achievement tests. For 40 percent of the LD pupils, however, an IQ test or an achievement test was either missing or achievement tests were administered that do not have normative data. Therefore, it was impossible for us or the staffing committee to calculate a discrepancy.

Many LD pupils were not functioning below grade level as measured by standardized achievement tests. In kindergarten and first grade, in fact, averages for some tests were above the students' years and months in school. In grades 2 and 3 the average grade equivalent score was only about four months below grade level. By the junior-high and high-school grades, the achievement lags were more substantial, on the order of 4.5 years for high-school aged pupils.

The comparison of single LD pupil characteristics with each definitional requirement separately gives only a crude and potentially misleading picture of the overall validity of LD placements. Because there is not a single, concrete way to operationalize the construct of LD, it is important to consider combinations of indicators to see if the *pattern* of signs justifies placement in LD even when no one indicator is significant by itself. A set of computer algorithms was developed to assign the 790 coded cases to mutually exclusive clusters on the basis of each subject's most salient characteristics. The hierar-

TABLE 1

Quantitative Identification of Subgroups in the Colorado LD Population Presented in Major Categories

	Percentage of LD Cases
Other Handicaps	
EMR	2.6%
Emotionally Disturbed	7.5%
Hearing Handicapped	0.2%
	10.3%
Learning Disabilities (True LD)	
Significant Ability/Achievement Discrepancy	20.5%
High-Quality Evidence of Processing Deficit	4.7%
Brain Injured	0.6%
Hyperactive	2.0%
Weak Significant Discrepancy *and* Verbal/Performance Discrepancy	3.6%
Weak Significant Discrepancy *and* Medium-Quality Processing Deficit	1.1%
Medium-Quality Processing Deficit *and* Verbal/Performance Discrepancy	6.6%
Medium-Quality Processing Deficit only	3.5%
	42.6%
Other Learning Problems	
Language Interference	6.6%
Slow Learners	11.4%
Environmental Causes	2.2%
Below-Grade-Level Achievement	6.1%
Minor Behavior Problems	3.7%
	30.0%
Other	
Poor Assessment (no IQ and no achievement tests)	6.4%
Miscellaneous (including normal — more than half are from high SES districts)	10.6%
	17.0%

Note: The standard errors computed for these percentages were generally 1% or less; the strict significant discrepancy percentage (20.5%) had an estimated error of 2%.

chical method of clustering and the criteria used to create each subgroup are explained in Shepard and Smith (1981).

The results of the subgroup analysis are summarized in Table 1. Of the school-identified population of LD, only 43% had evidence of LD. The remaining cases are more appropriately identified as "other handicaps" (10%) or "other learning problems" (30%) which do not imply a handicap (e.g., "second language interference" or "slow learner"). Six percent of the cases were not classified because they lacked IQ and achievement test data; 11% had data but could not be classified often because their performance levels were normal.

Qualitative analysis. A randomly parallel sample of 200 pupil cases was analyzed using qualitative methods. The following précis of results is summarized by research question.

1. *Significance of discrepancy.* Sixteen percent of the cases lacked sufficient data. Of all the cases, 39% demonstrated statistically significant discrepancies. Forty-five percent had nonsignificant or simply random differences between their ability scores and *any* of their achievement scores.

2. *Marginal placements.* If the pupil's symptoms were mild, if the clinical signs of his/her disability were equivocal, and if we judged that his/her need for special education service was not obvious, he/she was counted as a marginal placement. Out of the cases studied, 35% were categorized as marginal placements.

3. *Consistency.* In all but the smallest districts, the student considered for LD placement is evaluated by three, four, five, or even more professionals. The same characteristics evaluated frequently by one professional are evaluated by others. No one would expect complete agreement among these different clinical appraisals. However, some consistency is expected; for example, separate assessments of intellectual potential should yield about the same results; identification of specific strengths and weaknesses ought to converge and not conflict. If a child's weakness lies in reading, more than one test of reading achievement ought to reflect that. Of the cases studied, the clinical evidence of 68% was judged to be of poor consistency. The remaining cases showed good consistency; that is, the tests and clinicians presented evidence that converged on a coherent picture of the pupil's problems and characteristics.

4. *Cluster.* The typology of symptoms and characteristics was developed and the cases were classified by type. Their test results, histories, and reported symptoms were studied carefully and judged to fall into one of the following clusters.

a. Operational LD — The official definition of learning disabilities consists of achievement significantly below ability. Pupils are excluded from this definition if their ability-achievement discrepancy can be explained by language, emotional, or cultural disability. Of the cases studied, 21% fell into this cluster.

b. Clinical LD — According to the professional literature, children with learning disabilities may not evidence a significant discrepancy between their ability and achievement either because they have compensated for their specific disability through their stronger, intact abilities or because their disability depressed not only their achievement scores but their ability test scores. Children in this cluster, therefore, had no significant discrepancies between ability and achievement, but did show convincing evidence of an intrinsic disorder that was consistent across tests, clinicians, and time. Five percent of the cases fell into this cluster.

c. Slow learners — Profiles of the learning disabled tend to contain both significantly high and low scores on separate abilities. For slow learners, in contrast, all their separate abilities are approximately the same and lower than those of children their own age. Of all the files of students placed as LD and sampled for qualitative analysis, 13% were classified as slow learners. An additional 1% was classified as mentally retarded, having an IQ of less than 75 and no evidence of a learning disability.

d. Emotionally disturbed — Twenty-two percent of all the LD cases studied in the qualitative analysis were judged to be emotionally disturbed rather than

learning disabled. That is, the evidence about their psychological maladjustment was utterly convincing; evidence of any sort of intrinsic disorder was weak or nonexistent. Some were victims of child abuse or severe family problems. They were said to be highly anxious or deeply withdrawn, abusive, hyperactive, or emotionally unstable. In some cases, the parents or professionals resisted the label "emotionally disturbed" and opted for LD as the preferred label or treatment.

e. Language problem — Twelve percent of the students represented in the LD files are of the following type: They are of Hispanic or Indian descent. Some native language is spoken in their homes. Their verbal abilities are significantly lower than their performance or quantitative abilities. They are evaluated by school clinicians to be LD because of low achievement, differences between language and nonlanguage achievement, and difference between their scores on the verbal and performance scales of the *Wechsler Intelligence Scale*. This latter characteristic is mistakenly judged as a marker of perceptual disorder whereas it probably signifies language interference and, therefore, signals the need for intensive training in English.

f. Hearing, vision, or health — Some students have been categorized as LD although their primary problem relates to visual or hearing acuity or because they suffer from epilepsy or a physically based motor problem. As a result of the qualitative analysis seven percent of the files fell into this cluster.

g. Miscellaneous — Twenty percent of the cases did not fall into the above clusters. Therefore, they are included here. For example, 8% demonstrated no discernible handicaps or problems and no characteristics that allow us to typify them. Perhaps they were simply students who scored the lowest in classes or schools of above average children. Some appeared more like underachievers (3%) than like children with handicaps. Some seemed to be slower to develop than children of their own age, while not being outside the normal distribution of developmental rates (5%). And some had problems that appeared to be attributed more to teaching problems, the classroom situation, or teacher-pupil conflicts than to any psychological characteristics of the children themselves (4%).

5. *Necessity for special education.* All the students whose cases were chronicled in the LD files were judged by school committees to need help from special education. Yet, as we read the cases it was readily apparent that some were more in need of such assistance than others. For some, all their achievement scores lagged years behind their classmates' while others were only a month behind in spelling. Some needed only a little more flexibility on the part of their classroom teacher. Some needed perhaps only to change from open-space to a self-contained class. Some needed only their parents' expectations to become more realistic. Based on our judgment, 60 percent of the pupils surveyed needed special education help. Eighteen percent needed no help beyond that which a classroom teacher ought to be able to give. Twenty-two percent were judged to need a kind of help different from that which is typically available in LD programs — psychotherapy or intensive English training; some needed tutorial help in basic skills.

Costs

Two separate strategies were used to estimate the cost of the LD identification process from initial referral to the staffing conference. First, a typical sequence of events was described based on data from the analysis of pupils' files. For example, an average of six formal tests are administered and an average of six professionals usually attend the staffing for an average of 50 minutes. Time estimates were made for each element in the sequence. Then, median professional salaries (derived from a sample of three districts) were used to estimate costs on an hourly basis. From this procedure the average cost of identification was conservatively estimated to be $525. (This figure does not include transportation, personnel benefits, or materials costs.)

An independent source of cost data was ob-

tained from the reports that professionals submit to the Colorado Department of Education each year summarizing time spent in different activities. The proportion of time spent with LD pupils as well as the proportion of time spent in assessment and staffing were used to extract the fraction of the total special education costs spent on LD identification. This amount was divided by the total number of LD pupils for the same year (1978-79), yielding an average cost of $505.

The two figures arrived at by different routes confirm each other. The figure based on proportion of professional time, however, was amortized over the total LD population, not just those being assessed and placed for the first time that year. Therefore, the cost of identification is an ongoing "bite" from the special education budget, not just a one-time expense. The total amount of special education resources available to LD pupils each year (district + state) is only an average of $1204 (also 1978-79). Thus, just under half of all the special education resources for LD pupils is spent *each year* on identification and placement.

CONCLUSIONS

The major findings of the present evaluation were the following:

1. The numbers of LD pupils in Colorado are large and growing.
2. Most tests used in the assessment of LD do not meet technical standards for reliability and validity. Furthermore, clinicians who substitute professional judgment for test scores often have not been trained to test hypotheses and look for consistency in the evidence.
3. The identification procedures in Colorado districts are largely in compliance with state and federal requirements. Due process is satisfied. The staffing decisions are almost always made by teams of more than two people (more often seven or eight professionals). But sometimes parents are intimidated by the number of professionals.
4. The elaborate identification procedures have a negative consequence, however. The costs in professional time and dollar resources are excessive. Each year almost half of the special education funds available for LD students (state and local combined) are spent on assessment and identification.
5. Approximately 60% or more of the pupils currently identified as LD do not match the legal definitions or the definitions presented in the professional literature.

The overidentification of pupils in the LD category was the single most important finding from the study. The data from the pupil files, questionnaire responses, and much anecdotal evidence lead us to two competing explanations as to why so many students are labelled LD when they are not. In the following sections we discuss the "children-with-special-needs" versus the "problem-children" explanations.

Nonhandicapped Children with Special Needs

Many of the "nonhandicapped" children have serious problems in school and need special help. This is especially true for pupils in the language interference group. Because they are ethnic minorities, however, it is important that they not be called handicapped. They may lag seriously behind in school because their first language is not English or because they have trouble adapting to the mores of the school. Similarly, children who come from poor homes and who miss more than a month of school per year (classified in the environmental problem group) have enormous academic problems. Many of the students we identified as slow learners, whose IQs are in the range from 70 to 89 (without signs of LD), experience tremendous difficulty learning in school even though their progress is entirely consistent with their potential.

The evidence indicates that staffing committees classify as handicapped children of the above types. They are not handicapped, yet they need extra attention. Currently there is no way to provide such attention other than through an LD label. A few directors, especially from rural areas, specifically stated that special education was the only recourse for students who in a large district might have been in Title I or bilingual programs. Thus, we conjectured that the substantial correlation between district size and percentage of valid placements ($r = .51$) occurs because special education in large districts is more often a placement of last resort since these districts offer more alternative programs.

Removing Problem Students from the Regular Classroom

A less commendable motive can also be

described for misidentification, namely, removal of troublesome and hard-to-teach children from the regular classroom. Thus, some of the LD cases who did not demonstrate any of the indicators of LD and did not qualify for other handicapped subgroups scored above grade level on nationally normed tests. Some of them showed minor behavior problems as their only abnormal characteristic. Some had complete files, but not a single indicator of LD or other learning or behavior problems.

Coles (1978) proposed the radical thesis that labeling students learning disabled is a way of blaming them for the failure of schools to provide adequate education for all. For the 20% to 25% of LD cases who showed no signs of a handicap or who were not seriously below grade level, it is more reasonable to propose that the disorder lies in the school environment rahter than in the child. However, in the qualitative analysis "teaching problems" were mentioned by specialists as a possible source of the problem in less than 1% of the cases. We did not have the opportunity to observe the characteristics of children who were referred but not placed in LD. It is possible that some of these cases were treated by making adaptations in the regular classroom. A few districts have built into their identification process the requirement that alternatives be tried before special education placement. Nevertheless, the sizable number of LD cases showing no signs either of LD or of other serious learning problems suggests that the issue of problems within the school setting itself is not raised often enough.

Recommendations

Rules and criteria can be improved. They cannot, however, force valid placements. As with many psychological constructs, the validity of LD identification cannot be reduced to simplistic statistical rules. Minimal criteria for the reliability and discriminant validity of both formal and informal assessments can be established, but ultimately the integration of separate pieces of diagnostic information must rest on professional judgment. The findings of this study indicate the need for better training of clinicians. The validity of LD placements is also likely to be enhanced if clinicians feel that more rigorous adherence to the definition of LD will not deny services to the most extreme cases of students who are not handicapped but lag far behind in school. Alternative programs for nonhandicapped pupils will help to improve the validity of LD identification. At the same time, clinicians eager to meet the needs of children will have to address the issue of the extra costs and potential harm that results when normal children are called handicapped because regular education is unprepared to serve the full range of normal behaviors and learning styles.

POLICY IMPLICATIONS

The Special Education Advisory Committee to the Colorado State Board created a task force to respond to the Shepard and Smith study. That task force drew the following conclusions:

> The inescapable inference that should be drawn from this study is that too many children have been labeled inappropriately by well intentioned professionals operating in a system that attempts to meet the needs of too many children who are unsuccessful in regular classrooms.
>
> The problem is not just a special education problem. The general education community needs to provide broader alternatives for marginal students other than special education. The problem is not just a Colorado problem. It is being experienced nationally by states and school districts *(Report of the Special Education Committee Task Force,* 1981, p. 1).

Because the central concern of the Advisory Committee Task Force had focused on what to do with the "marginal learner," the Colorado Commissioner of Education created a second task force to address this issue. The task force on marginal learners concluded that not every student who was "not having a successful school experience" needed to go through the comprehensive identification process needed to ensure the educational and constitutional rights of the handicapped (*Task Force Report,* 1981). Instead, a consultive teacher model was proposed whereby special education teachers would work with a child in the context of the regular classroom rather than labeling him/her as handicapped. This solution is responsive to both the needs of the marginal learner and the excessive costs of LD identification. However, it may be exceedingly difficult to establish it as policy because heretofore legal requirements forbid the use of special education resources for "nonhand-

icapped" pupils.

It is too soon to assess the impact of the evaluation study. Davis and Smith (1981) provide a summary of the legislative hearings where the study results and responses to the study were presented. One result in response to the finding of excessive costs of identification has been the drafting of a bill to remove the Colorado requirements for assessment which are far in excess of the federal requirements. Alternatives to address the problem of overidentification are still being debated.

REFERENCES

Applied Management Sciences. FEDAC supporting statement and instruments for a study to evaluate procedures undertaken to prevent erroneous classification of handicapped children. Department of Education, Contract No. 300-79-0669, July 1980.

Clements, S.D. *Minimal brain dysfunction in children* (NINDS Monograph No. 3, U.S. Public Health Service Publication No. 1415). Washington, DC: U.S. Government Printing Office, 1966.

Cochran, W.G. *Sampling techniques*. New York: Wiley & Sons, 1963.

Coles, G.S. The learning disabilities test battery: Empirical and social issues. *Harvard Educational Review*, 1978, *48*, 313-340.

Davis, W.A., & Shepard, L.A. Specialists' use of tests and clinical judgment in the diagnosis of learning disabilities. *Learning Disability Quarterly*, 1983, *6*, elsewhere in this issue.

Davis, W.A., & Smith, M.L. *The history and politics of a state-wide evaluation: The Colorado study of learning disabilities*. Paper presented to the joint meeting of the Evaluation Research Society and the Evaluation Network, Austin, October 1981.

Glaser, B.G., & Strauss, A.L. *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine Publishing Co., 1967.

Hammill, D. Learning disabilities: A problem in definition. *Division for Children with Learning Disabilities Newsletter*, 1974, *4*, 28-31.

Hammill, D.D., Leigh, J.E., McNutt, G., & Larsen, S.C. A new definition of learning disabilities. *Learning Disability Quarterly*, 1981, *4*, 336-342.

Kirk, S.A., Beery, P.B., & Senf, G.M. A survey of attitudes concerning learning disabilities. *Journal of Learning Disabilities*, 1979, *12*, 239-245.

Kirk, S.A., & Elkins, J. Characteristics of children enrolled in the Child Service Demonstration Centers. *Journal of Learning Disabilities*, 1975, *8*, 630-637.

National Advisory Committee on Handicapped Children. *Special education for handicapped children. First Annual Report*. Washington, DC: U.S. Department of Health, Education, and Welfare, January 31, 1968.

National Joint Committee for Learning Disabilities. *Learning disabilities: Issues on definition*. Unpublished manuscript, 1981. (Available from Drake Duane, NJCLD Chairperson, c/o The Orton Dyslexia Society, 8415 Bellona Lane, Towson, Maryland 21204.)

Norman, C.A., & Zigmond, N. Characteristics of children labeled and served as learning disabled in school systems affiliated with Child Service Demonstration Centers. *Journal of Learning Disabilities*, 1980, *13*, 16-21.

Poplin, M.S. The severely learning disabled: Neglected or forgotten? *Learning Disability Quarterly*, 1981, *4*, 330-335.

Report of the Special Education Advisory Committee Task Force on Perceptual-Communicative Disorders in Colorado. Colorado Department of Education, Denver, Colorado, April 1981.

Shepard, L., & Smith, M.L., with Davis, A., Glass, G.V, Riley, A., & Vojir, C. *Evaluation of the identification of perceptual-communicative disorders in Colorado*. Laboratory of Educational Research, University of Colorado, Boulder, Colorado 80309, February 1981.

Smith, M.L. Case histories of perceptual-communicative disordered children in Colorado. In L. Shepard, & M.L. Smith, with A. Davis, G.V Glass, A. Riley, & C. Vojir, *Evaluation of the identification of perceptual-communicative disorders in Colorado*. Laboratory of Educational Research, University of Colorado, Boulder, Colorado 80309, February 1981.

Task Force Report on Insuring All Students an Appropriate Education. Colorado Department of Education, Denver, Colorado, 1981.

Thouvenelle, S., Rader, J.R., & Hanley, T.V. *A study to evaluate procedures undertaken to prevent erroneous classification of handicapped children*. Contract Number: 300-79-0669, U.S. Department of Education, unpublished draft data, 1982.

Thurlow, M.L., & Ysseldyke, J.E. *Current assessment and decision-making practices in model programs for the learning disabled* (Research Report No. 11). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities, 1979.

Ysseldyke, J.E., Algozzine, B., Richey, L., & Graden, J. Declaring students eligible for learning disability services: Why bother with the data? *Learning Disability Quarterly*, 1982, *5*, 37-44.

Ysseldyke, J.E., Algozzine, B., Shinn, M.R., & McGue, M. Similarities and differences between low achievers and students labeled learning disabled. *The Journal of Special Education*, 1982, *16*, 73-85.

FOOTNOTES

[1]In Colorado, learning disabilities are called perceptual-communicative disorders (PCD). The label was changed in 1973 for this and several other categories of handicap to be less pejorative. Nevertheless, the definition and criteria for PCD are equivalent to the federal definition and guidelines for LD. It is acknowledged, however, that some specialists in Colorado (20% according to our survey data) may be influenced by the label to use perceptual-motor models for diagnosis and treatment.

[2]A stratified two-stage cluster sampling design was used for both the study of case files and the survey of professionals. The sampling frame, weighting procedures, and formulae for calculating standard errors (Cochran, 1963) are given in technical appendix A (Shepard & Smith, 1981).

[3]All identifying information in the pupil files was removed by clerks on site.

[4]Several items regarding the definition of LD were taken with permission from Kirk, Beery, and Senf (1979). Some questions were modified from Applied Management Sciences (1980).

[5]The definitions and coding rules for these variables are given in Shepard and Smith (1981).

This research was funded in part by the Colorado Department of Education.

Requests for reprints should be addressed to: Lorrie Shepard, Laboratory of Educational Research, Box 249, University of Colorado, Boulder, CO 80309.