5-1987

# Expert System Technology and Concept Instruction: Training Educators to Accurately Classify Learning Disabled Students

Mary Anne Prater
*Utah State University*

EXPERT SYSTEM TECHNOLOGY AND CONCEPT INSTRUCTION:

TRAINING EDUCATORS TO ACCURATELY CLASSIFY

LEARNING DISABLED STUDENTS

by

Mary Anne Prater

A dissertation proposal submitted in partial fulfillment
of the requirements for the degree

of

DOCTOR OF PHILOSOPHY

in

Special Education

Approved:

UTAH STATE UNIVERSITY
Logan, Utah

1987

ACKNOWLEDGEMENTS

I wish to acknowledge those who have provided support and encouragement throughout my doctoral studies. In particular I wish to thank my chairperson, Dr. Joseph Ferrara for his insightful ideas and enthusiasm from the conception to the completion of this research. I wish also to thank the other members of my committee, Dr. Alan Hofmeister, Dr. Ron Thorkildsen, Dr. Margaret Dyreson, and Dr. Hyrum Henderson, for their individual and collective contributions.

I would also like to acknowledge and thank those who arranged for students at their individual universities to participate as subjects in this study--namely, Dr. Richard Baer, Utah State University; Dr. Eric Rudrud, St. Cloud State University; and Dr. Linda Reetz and Dr. Ray Thompson, The University of South Dakota.

In addition I am indebted to the contributions made by my fellow-doctoral students. In particular, to Margaret Lubke and Elizabeth Martindale for their moral and professional support throughout the evolution of not only this research but of the doctoral studies as a whole. I wish to also acknowledge Dr. Jim Parry who proceeded me in this endeavor for modeling completion of his degree.

And to my parents, Herman and Barbara Prater, my thanks and grateful appreciation for providing the continual and needed encouragement and support which was vital to the completion of this degree.

Mary Anne Prater

# TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

# LIST OF TABLES

LIST OF TABLES (continued)

# LIST OF FIGURES

Figure | page

# ABSTRACT

Expert System Technology and Concept Instruction:
Training Educators to Accurately Classify
Learning Disabled Students

by

Mary Anne Prater, Doctor of Philosophy

Utah State University, 1987

Major Professor: Dr. Joseph M. Ferrara
Department: Special Education

Many learning disabled student being served by the public school systems have been inaccurately classified. Training and research efforts are needed to assist members of the multidisciplinary team in making more accurate learning disabilities classification decisions.

CLASS.LD2, a computer-based expert system, was designed to assist multidisciplinary teams by providing second-opinion advice regarding the appropriateness of a learning disabilities classification for individual student cases. The existing expert system, CLASS.LD2, was combined with strategies for effective concept instruction to create an instructional package entitled LD.Trainer.

The purpose of this study was (a) to develop a computer-based instructional package combining expert system

technology and strategies for effective concept instruction and (b) to test the effectiveness of the instructional package against another system application. The training application against which the instructional package was compared consisted of users running consultations with the original expert system.

Of specific interest was (a) the effectiveness of both training programs across experienced and inexperienced teachers, (b) the performance of the experienced as compared with the inexperienced teachers regardless of the training program used, (c) whether an interaction between level of experience and training program occurred, (d) which training program was more effective for the experienced teachers, and (e) which training program was more effective for the inexperienced teachers.

Ninety-seven students from three universities served as subjects and were randomly assigned to one of the two treatment groups. Subjects who completed the LD.Trainer materials scored statistically ($p < .Ø5$) and educationally higher (SMD = + Ø.96) on the posttest than those who ran CLASS.LD2 consultations. Statistical and educational significance were also obtained across the experienced and inexperienced subjects when considered alone. An interaction, although not statistically significant ($p < .Ø5$), was obtained between group and experience level.

Although there exist many similarities between the processes of building expert systems and concept analysis, incorporating both to develop an effective training tool had not previously been demonstrated. Results of this study indicated that the two fields, successfully combined, can create an effective and efficient training tool.

(192 pages)

# INTRODUCTION

Public Law 94-142 mandates that a multidisciplinary team, after gathering performance data on a potentially handicapped student, meet together to make decisions about the student's educational program. These decisions include whether the student is handicapped, and if so, what handicapping condition; what goals and objectives are appropriate for the student; and where the student could be most appropriately and least restrictively served (Code of Federal Regulations, 1980).

Research indicates that "the special education team decision-making process, as currently employed in public school settings, is at best inconsistent" (Ysseldyke et al., 1983, p. 77). For example, Ysseldyke, Algozzine, Richey, and Graden (1982) report that educational decision makers "use assessment data to support or justify decisions that are made independent of the data" (p. 42).

Because of the poor decisions made by multidisciplinary teams, Ysseldyke (1983) has estimated that, as many as half of the students labeled learning disabled (LD) have been inappropriately classified. These inaccurate decisions may also account, in part, for the observed 84 percent increase in the number of learning disabled students identified during the past few years (Hofmeister, 1983).

The definition of learning disabilities is unclear and an often debated issue (Sabatino, 1983). In practice, the range of characteristics of students being served as LD is very broad. "[H]eterogeneity is the rule rather than the exception" (Keogh, 1983, p. 22). In fact, several researchers are presently investigating subtypes of learning disabilities hoping to improve research samples, treatment alternatives, and diagnosis of LD (McKinney, 1984).

Although the debate continues regarding LD definitions, the definition presently used in the American public school systems is based on Public Law 94-142. In order to receive funds for special education, each State must design, present, and have approved, their plan for implementing P. L. 94-142 regulations. As part of the plan, the diagnostic definition for each handicapping condition must be stated.

Even with efforts to clarify definitions for LD, there remain problems with the multidisciplinary team decision-making process. That is, the team members fail to consistently consider objective data (Ysseldyke, et al., 1982), and the quality of that data (Ysseldyke, Algozzine, Regan, & Potter, 1980). These problems clearly contribute to the present overidentification of learning disabled students (Ysseldyke, et al., 1983). Because of poor decision-making processes in the school systems and lack of clear definitions in the literature, one is not surprised at the enormous numbers of LD students who are presently being served but who are inappropriately classified.

In their review of over five years of research on LD assessment and decision-making, Ysseldyke et al. (1983) conclude that future efforts must involve training members of the multidisciplinary team to become better decision makers. Clearly this includes training them to more accurately identify learning disabled students.

General dissatisfaction and ineffectiveness of traditional modes of training contribute to this specific problem. Institutions of higher learning have "assumed a central, primary, comprehensive, and continuing responsibility for the integrity and vitality of society's knowledge base" (Smith, 1978, p. 3). Yet, problems such as inadequate curricula, lack of faculty training, and failure to keep up-to-date equipment and facilities contribute to the failure of universities to adequately teach the content of this ever-changing knowledge base (Carnegie Council, 1979).

The specific lack of adequate teacher preparation at the university level places additional burden on inservice programs of local and state education agencies. Like universities, local and state education agenices have been criticized for using ineffective models of instruction. For example, inservice instruction typically is didactic, very general in scope, and lacks effective feedback (Borg, Kallenback, Kelly, Langer, & Gall, 1970). Effective models of preservice and inservice instruction should consider the immediate needs and interests of teachers and be based on a programmatic approach to handling real-life educational decision-making problems (Wang, Vaughan, & Dytman, 1985).

The LD classification decisions made by multidisciplinary teams are real-life decision-making problems in the public schools. In order to assist educators in making accurate LD classification decisions, Ferrara and Hofmeister (1984) developed an expert system entitled CLASS.LD2. Expert systems are computer-based programs which replicate human decision-making processes (Barr & Feigenbaum, 1981). Based on user responses and "inferences" made by the system, CLASS.LD2 provides the user with advice regarding the appropriateness of a learning disabilities classification (Ferrara, Parry, & Lubke, 1985). Expert systems, such as CLASS.LD2, may be used to provide second opinions as a consultant would (Hofmeister & Ferrara, 1986).

In addition to providing second opinions, expert systems may be applied to training situations (Prater & Ferrara, 1986). Previous attempts at applying existing expert systems for training purposes have involved the development of sophisticated front-ends to the original system and have taken the form of intelligent computer-assisted instructional programs (Clancey & Letsinger, 1981). For example, MYCIN, a well-known medical expert system, was adapted for instructional purposes (Davis, Buchanan, & Shortliffe, 1975). Initially, MYCIN was programmed to assist physicians with diagnosing bacterial diseases. Later the MYCIN data base was used to develop an intelligent computer-assisted instruction (ICAI) program entitled NEOMYCIN. It was designed to teach physicians and medical students diagnosis of bacteriological diseases in the form of ICAI (Clancey & Letsinger, 1981).

Expert systems may be employed as training tools without complicated and sophisticated modifications (Prater & Ferrara, 1986; Prater & Lubke, 1986). For example, users who simply engage in consultations with expert systems may increase their knowledge or decision-making ability regarding the content of the system (Alessi & Trollip, 1985). Or, tools of effective concept instruction may be combined with expert system technology to create an expert system-based training package. That is, the system may be modified to facilitate the presentation of examples and nonexamples so that complex concepts such as "learning disabled students" may be taught (Ferrara, Prater, & Baer, in press).

## Statement of the Need

Expert systems, computer programs designed to replicate the best experts' logic and decision-making processes, can be modified to serve as training tools. These training materials may be developed in such a way as to incorporate knowledge about effective computer-assisted instruction and conceptual instruction. However, no one has developed such a training package. Therefore, the effectiveness of such a training package is unknown.

Effective preservice and inservice programs dealing with the LD classification issues appears vital. CLASS.LD2, the expert system designed to provide second opinions regarding the accuracy of LD classifications, may be effectively employed as a training tool for both experienced and inexperienced teachers.

## Purpose of the Study

The purpose of this proposed study was (a) to develop a computer-based instructional package combining expert system technology and effective concept instruction strategies and (b) to test the effectiveness of the instructional package against another expert system application. The development portion of the study included formative evaluation steps as outlined in the procedures section. Testing the effectiveness of the training package required an experimental research design in order to compare the training package against another training application. The training application against which the instructional package was compared, consisted of users running consultations with the original expert system.

Of specific interest was (a) the effectiveness of both training programs across experienced and inexperienced teachers, (b) the performance of the experienced as compared with the inexperienced teachers regardless of the training program used, (c) whether an interaction between level of experience and training program occurred, d) which training program was more effective for the experienced teachers, and (e) which training program was more effective for the inexperienced teachers.

## Research Questions

The major research questions for this study included:

1. Based on experienced and inexperienced educators' accurate classification of learning disabled students in a selected instructional environment:

a. Is the modified expert system training package, LD.Trainer or the original expert system, CLASS.LD2, more effective?

b. Do experienced and inexperienced teachers perform equivalently regardless of the training method they used?

c. Does an interaction exist between the amount of teaching experience the subjects have (i.e., experienced and inexperienced) and the training method they used?

d. Which training method is more effective with experienced teachers?

e. Which training method is more effective with inexperienced teachers?

## Hypotheses

1. Using Analysis of Variance (ANOVA), there will be no statistically significant ($p < .05$) difference between the posttest performance of those participating in the LD.Trainer and the CLASS.LD2 groups.

2. Using ANOVA, there will be no statistically significant ($p < .05$) difference between the posttest performance of experienced and inexperienced teachers.

3. Using ANOVA, there will be no statistically significant ($p < .05$) interaction between amount of experience and training method.

4. Using Planned Orthogonal Contrasts, there will be no statistically significant ($p < .05$) difference between the posttest performance of experienced teachers in the LD.Trainer and the CLASS.LD2 groups.

5. Using Planned Orthogonal Contrasts, there will be no statistically significant ($p < .05$) difference between the posttest performance of inexperienced teachers in the LD.Trainer and the CLASS.LD2 groups.

# REVIEW OF LITERATURE

Researchers indicate that public school personnel inaccurately classify many of the students presently served as learning disabled (Ysseldyke et al., 1983). The concept of "learning disabled" is complex and difficult to teach (Ferrara et al., in press). Training preservice and inservice personnel to more accurately identify learning disabled students could incorporate empirical and theoretical knowledge about effective concept instruction.

An expert system that provides advice regarding the appropriateness of a learning disabilities classification, CLASS.LD2 (Ferrara & Hofmeister, 1984), has been developed and may be modified as a tool for training. In order to develop and test the effectiveness of the training tool, literature in the following areas was reviewed: concept instruction, traditional and intelligent computer-assisted instruction, and expert system technology.

## Concept Instruction

Concepts are the fundamental structure for thought throughout a human being's lifetime (Klausmeier, Ghatala, & Frayer, 1974). In fact, our whole world can be described in terms of concepts. Although some concepts are learned through observation (Gagne, 1985), conceptual learning is also an integral part of any school curriculum (Markle, 1975). Consequently, psychologists, instructional designers, and educators have been concerned with the teaching and learning of concepts for many years (Woolley & Tennyson, 1972). Before presenting procedural strategies for designing effective concept instruction, a description of what concepts are is discussed.

Definitions of "Concept"

Some authors suggest that formal definitions of concepts vary widely (Martorella, 1972; Klausmeier, et al., 1974). Most recent definitions in the literature, however, include some reference to a set of characteristics which distinguish examples of the concept from nonexamples of that concept. Tennyson and Park (1980), for example, define a concept as a "set of specific objects, symbols, or events which share common characteristics (critical attributes) and can be referenced by a particular name or symbol" (p. 56).

Concepts may be further described as basic or complex. Engelmann and Carnine (1982) consider a "basic concept" as "one that cannot be fully described with other words (other than synonyms)" (p. 10). Good examples of basic concepts are colors. Since they can't be adequately described in words, examples must be used to teach basic concepts. Complex concepts depend upon context and dimensionality (Ferrara et al., in press). With the concept "liberal," for example, variations occur in its meaning depending upon the times and the current social perspective. Although "liberal" may be defined in understandable verbal terms, examples in context must also be presented in order to grasp the full conceptual meaning.

## Properties of All Concepts

Frayer (1970) (cited in Martorella, 1972) suggested that concepts may have six common characteristics: (a) critical attributes, (b) rules for joining attributes, (c) irrelevant attributes, (d) a label, (e) supraordinate concepts, and (f) subordinate concepts. In order to discuss effective concept instruction, each must be defined.

*Critical and irrelevant attributes.* "A critical attribute refers to any attribute that is essential to an example for the example to be classified as a member of a given concept class. An attribute that may be present but is not essential is termed an 'irrelevant attribute.'" (Hofmeister, 1977, p. 98). A variety of labels is used by different authors to describe "critical" and "irrelevant" attributes. Critical, relevant, or defining attributes are synonymous as are noncritical, irrelevant, or variable attributes.

*Rules.* The rules for joining concept attributes are typically divided into four types: conjunctive (and), disjunctive (and/or), conditional (if, then), and biconditional (if and only if) (Bourne & O'Banion, 1971). Although the most common type of concepts are conjunctive

(Merrill & Tennyson, 1977), it is important to clearly understand the type of rule used to connect the attributes so that it may be taught (Engelmann, 1969).

Labels. Although concept labels are arbitrary (Tennyson & Cocchiarella, 1986), they represent the set of characteristics that the examples of the concept have in common and thus, provide a means of communication. In fact, most of the words used in any given language refer to concepts (Merrill & Tennyson, 1977).

Supra- and subordinate concepts. Supraordinate and subordinate concepts are one way of describing the relationship between concepts. In addition to these two, there are also coordinate concepts. If one supraordinate concept can be divided into several subordinate concepts, then the subordinate concepts are coordinate concepts (Merrill & Tennyson, 1977).

## Effective Concept Instruction

Researchers investigating the most effective instructional strategies for teaching concepts have derived empirically-based sets of design strategies (Tennyson & Park, 1980). Merrill and Tennyson (1977); Engelmann and Carnine (1982); Klausmeier (1980); and Eggen, Kauchak, and Harder (1979) have, for example, developed frameworks for development of instructional materials designed to teach concepts. Most recently, Tennyson and Cocchiarella (1986) have presented an instruction design for concept teaching which is an updated extension of the Merrill and Tennyson (1977) model and remains based on "direct empirical validation from a programmatic line of instructional systems research" (p. 40). These models for the development of effective concept instruction, as well as primary research in this area, have been reviewed and a combined procedure is presented.

In order to effectively teach concepts, five components of the instruction needs to be carefully designed. These five components include (a) analysis of the concept, (b) definition of the concept, (c) examples and nonexamples of the concept, (d) the teaching sequence, and (e) the diagnostic classification test.

Concept analysis. The analysis of the concept must include the content structure of the concept, including the broader and prerequisite concepts. For example, one must consider the coordinate, subordinate, and supraordinate structure of the concept(s) being taught. The analysis facilitates several functions. First, the structure helps determine the most effective instructional strategies (Tennyson & Cocchiarella, 1986). Second, the structure may be presented to the learner (Driscoll & Tessmer, 1985a; Markle, 1977). And third, such structures may be used to assess student knowledge (Champagne, Klopfer, Desena, & Squires, 1981).

When coordinate concepts are being taught, one is also teaching the broader concept. That is, when a student "learns the first individual concept, he learns the general operation or procedure for handling all instances of the broader concept" (Engelmann, 1969, p. 51). Engelmann used the example of polar concepts. Once the student learns 'big' from 'small,' it is easier to learn and teach 'tall' from 'short.'

The prerequisite concepts also need to be considered. That is, if the lessons are structured in such a way as to incorporate prerequisite concepts, this will improve storage and retrieval of information (Tennyson & Cocchiarella, 1986).

In addition to the structure of the concept, the critical attributes and the role of context need to be determined. A clear list of the critical and irrelevant attributes will facilitate generation of examples and nonexamples, selection of the appropriate number of examples and nonexamples, and organization and sequencing of the examples and nonexamples throughout instruction. Contextual information will prove helpful when creating a definition of the concept. In addition, the label might be further elaborated by use of the concept (Tennyson & Cocchiarella, 1986).

Concept definition. "Definitions are statements that express rules for classifying" (Gagne, 1985, p. 114). The verbal definition of the concept must communicate all of the critical attributes and the relationships of those attributes to the learner (Carroll, 1964). In addition to the content of the definition, one needs to make certain that the definition is written in vocabulary appropriate to the target population (Feldman & Klausmeier, 1974).

Examples and nonexamples. The examples and nonexamples should be matched on the irrelevant attributes, but differ on the critical attributes. By matching examples and nonexamples on irrelevant attributes one is demonstrating that the irrelevant attributes are not important attributes in distinguishing examples from nonexamples (Tennyson, Woolley, & Merrill, 1972).

It may be possible that learner sophistication and task complexity interact with the need to minimize variation in the irrelevant attributes. For example, Carnine (1980b) discovered that preschoolers who were exposed to maximum differences between examples and nonexamples scored higher than those exposed to minimal differences. Previous research had been conducted on adult learners (Tennyson, et al., 1972). However, when examples and nonexamples differ on irrelevant attributes, one needs to be certain that the student has learned to discriminate based on the relevant, not the irrelevant attributes.

Generally speaking, research supports that both examples and nonexamples need to be used in instruction

(Williams & Carnine, 1981). However, the researchers of one study did conclude that the ratio between relevant and irrelevant attributes within examples of the concept may determine whether negative instances are helpful in teaching the concept (Shumway, White, Wilson, & Brombacher, 1983). In addition, it is possible that successive concepts may be effectively taught with examples only (Tennyson & Cocchiarella, 1986).

Presentation of examples and nonexamples is most effective when the matched pairs vary widely on irrelevant attributes (Tennyson, et al., 1972; Carnine, 1976). Divergency with respect to both irrelevant attributes and contexts is necessary (Merrill, Reigeluth, & Faust, 1979).

To adequately teach concepts one must not only be concerned with teaching discrimination between examples and nonexamples, but also generalization beyond the examples and nonexamples used in instruction (Carnine, 1980a; Driscoll & Tessmer, 1985b). Generalization refers to accurately classifying a new example which is novel or differs in some way from previously encountered examples (Markle & Tiemann, 1970).

Teaching sequence. The teaching sequence should include eight components: instructional objectives, definition and label, appropriate number and sequence of examples and nonexamples, both expository and interrogatory examples, elaboration of the critical attributes, strategy help, immediate feedback, and "embedded refreshment."

The instructional objectives should define the purpose of the lesson and should be written in observable terms. The understanding of a concept can be demonstrated in three ways. They include (a) when given instances of the concept the student identifies which are examples; (b) when given instances the student groups them into concepts; or (c) when given the concept label the student identifies or produces the definition (Merrill, 1983). Although labeling, recognizing, or recalling a definition may be, under some circumstances, a desirable instructional objective, it should not be confused with understanding a concept or demonstrating classification behavior (Merrill & Tennyson, 1977, Markle, 1977).

Presenting highly meaningful labels, or labels which the students understand, facilitates concept learning (Fredrick & Klausmeier, 1968). The definition can be used to recall for the student component elements or a framework of the concept (Tennyson & Boutwell, 1974). Presenting the definition to the learners can economize the teaching sequence by reducing the number of examples needed to learn the concept (Engelmann, 1969; Tennyson & Park, 1980) and has been demonstrated to be more effective than only demonstration of examples and nonexamples (Anderson & Kulhavy, 1972; Merrill & Tennyson, 1977). The definition should be presented to the learner before the examples and nonexamples (Tennyson & Park, 1980).

Including the appropriate number of examples and nonexamples is often considered a matter of judgment, the rule being: include enough examples to adequately represent the concept and enough nonexamples to clearly differentiate the concept from other similar concepts (Eggen et al., 1979). It has been suggested that the optimal number of examples and nonexamples is related to: the number of critical and irrelevant attributes of the concept and/or learner characteristics, such as age, prior knowledge, aptitude, on-task performance, and cognitive style (Tennyson & Park, 1980; Tennyson & Rothen, 1977)

When students are told or shown examples and nonexamples of the concept and simultaneously given the identifying concept name, these examples and nonexamples are called "expository instances" (Merrill, et al., 1979). In addition, students can be presented with an example or nonexample and asked to recall or match it to the concept name. This is referred to as "interrogatory" or "inquisitory instances" (Merrill & Tennyson, 1977). Both types of instances are necessary for effective concept instruction (Tennyson & Cocchiarella, 1986).

The order in which the examples and nonexamples are presented is also important. Several researchers have concluded that a "best example," an example which is average, central, or prototypical, should be presented first

(Tennyson & Cocchiarella, 1986; Tennyson, Youngers, & Suebsonthi, 1983; Klausmeier, et al., 1974). However, because the first examples used in instruction are the hardest and take the longest time to learn, Engelmann (1969) has suggested presenting first examples which are trivial or hardest to identify.

One should use some kind of attention-focusing device to "direct the student's attention to the critical attributes present in a specific example; to potentially confusing variable attributes present in a specific example or nonexample; and to the absence of the critical attributes in a specific nonexample" (Merrill & Tennyson, 1977, p. 83). Attention can be directed through the use of highlighting, coloring, underlining, or printing in bold print. When presenting information verbally, the instructor can change verbal intonation or stress to emphasize the attributes.

Strategies--such as elaboration, retrieval, chunking, and mnemonics (Atkinson, 1975; Torgeson & Kail, 1980)--can be incorporated as part of the presentation form to assist students in gaining mastery of the concepts being taught (Merrill et al., 1979). Similarly, routines or algorithms can be taught to students (Engelmann & Carnine, 1982).

Feedback is important for interrogatory or practice items (Merrill et al., 1979). Rather than feedback as knowledge of results, it is suggested that feedback as attribute elaboration be used because it will assist most students to see **why** an example is an example (Merrill & Tennyson, 1977). Otherwise, if only correctness is given, students will probably fail to understand why.

"Embedded refreshment" is a term used by Tennyson and Cocchiarella (1986) to describe a design strategy used for recalling specific prerequisite knowledge. Typically, advance organizers and reviews have been used. Tennyson and Cocchiarella also recommend the use of pretests and/or a procedure whereby students are presented with embedded refreshment only if they are unable to solve an interrogatory problem.

Diagnostic classification test. The test is used to assess students' abilities to discriminate between examples and nonexamples and to generalize to new instances. This test should incorporate novel examples and nonexamples that are representative of the concept (Harris, 1973; Markle & Teimann, 1970; Engelmann & Carnine, 1982). The items on the test should represent a valid and divergent sample of the concept (Merrill & Tennyson, 1977) and should be sequenced randomly throughout the test (Merrill et al., 1979).

Three major classification errors have been described by several authors. They include: overgeneralization, undergeneralization, or misconception (Markle & Teimann, 1970; Merrill & Tennyson, 1977; Park, 1981). Assessing the type of classification error facilitates selection of the most appropriate remedial technique.

Summary

An overview of some key instructional design strategies for concept instruction have been presented. Both research findings and instructional design theory have been reviewed. There exists, however, a great deal more literature in the area of concept instruction. For example, some researchers have been studying the influence of learner characteristics such as advising students regarding their progress throughout instruction (Tennyson & Buttrey, 1980) or providing learner control in determining when to stop instruction (Tennyson, 1980; Tennyson, 1981).

## Computer-Assisted Instruction

Application of computer software to assist concept instruction is not new. Some research is being conducted in the areas of concept instruction applications to intelligent computer-assisted instruction (Park, 1981; Park 1984) and artificial intelligence-based instruction (Tennyson, 1986; Prater & Ferrara, 1986; Ferrara, et al., in press).

Educational software may be divided into three categories (a) conventional computer-assisted instruction (CAI) (b) simulations, and (c) intelligent computer-assisted instruction (ICAI) (Harmon & King, 1985).

Conventional Computer-
Assisted Instruction

Most of the educational software present today may be considered computer-assisted instruction (CAI). In CAI programs, the computer is the primary deliverer of the instruction. This should not be confused with computer-managed instruction in which the computer only manages the delivery of instruction (Burke, 1982).

Burke (1982) defines traditional CAI as:

> [T]he direct use of the computer for the facilitation and certification of learning - that is, using the computer to make learning easier and more likely to occur (facilitation), as well as using the computer to create a record proving that learning has occurred (certification) (p. 16).

Traditional CAI software may be divided into two categories (a) tutorial instruction and (b) drills (Alessi & Trollip, 1985; Burke, 1982). Tutorial programs present information and guide the student through the learning process, while drill programs primarily incorporate practice exercises. Generally speaking, "CAI programs follow a script-based model of instruction controlled by the program developer. Program developers determine the amount of information presented to the student, the sequence of instructional content, and the specific questions or problems to which the student must respond" (Thorkildsen, Lubke, Myette, & Parry, 1985, p. 5).

CAI packages have been implemented in a variety of ettings and with a variety of content areas (Clark, 1983).

In some sense CAI programs are like books. That is, they provide instructional information to the student (Wenger, 1985). The software presents information, asks for a response, and then provides the user feedback (Harmon & King, 1985).

The drill and practice model is not an effective way to instruct concepts, particularly complex concepts, because they are multidimensional and dependent upon context. Traditional CAI is not designed to manipulate all the necessary facets and to vary the outcome which can facilitate teaching complex concepts (Ferrara et al., in press).

### Simulation

The second type of software, simulation programs, provide some kind of simulation or projected data with which the student interacts. The student is presented with a series of questions and then is provided an outcome based on the responses. Therefore, the student may observe in a simulated situation how the outcome can vary depending upon their responses to the questions (O'Shea & Self, 1983). This type of software "seeks to teach concepts by allowing the student to gain experience in a series of simulated situations" (Harmon & King, 1985, p. 239).

Non-computerized simulations have been used in fields such as teacher and administrator preparation programs for many years (Wolfe & Macauley, 1975) and computerized simulations have been demonstrated successfully (Flake, 1975). Now "[t]here is an enormously wide range of topics on which research based on computer simulation is progressing" (Priest, 1981, p. 285). For example, computerized simulations have been developed in areas such as statistical concepts (Stockburger, 1980), logic circuits (Steinberg, Baskin, & Hofer, 1986), genetics (Kinnear, 1986), and conservation and energy (Cartwright & Neikkinen, 1981).

It has been suggested, however, that many of the simulations developed may be inadequate because (a) they deal only with learning specific related skills rather than cognitive strategies, (b) the learners lack the necessary prerequisite knowledge, (c) no remediation or recall of information is provided, and (d) failure to incorporate the simulation into any course curriculum (Breuer & Hajovy, in press).

An example of a simulation type of program is a program entitled STEAMER (Hollan, Hutchins, & Weitzman, 1984). STEAMER acts as an instructional tool for training naval steam propulsion engineers. The program simulates actual objects which could be used in teaching, but at great expense. That is, mistakes with a real steam turbine could be dangerous and expensive (Goodall, 1985). STEAMER provides an "interactive inspectiable simulation based on computer graphics" (Wenger, 1985, p. 38) by displaying a functional and representative model of the propulsion.

Although conventional CAI may be inadequate, simulations hold promise for complex concept instruction. Simulations can be designed to be multifaceted and based upon varied outcomes.

### Intelligent Computer-Assisted Instruction

The third type of educational software involves intelligent systems, also called intelligent tutoring, intelligent teaching systems, or intelligent computer-assisted instruction (ICAI). ICAI programs use artificial intelligence techniques so that, unlike CAI programs, students can "interact with the computerized tutor rather than just respond to the tutor's directives" (Thorkildsen et al., 1985, p.5). The present ICAI systems exist primarily for the function of experimentation (Sleeman & Brown, 1982). Three possible forms of ICAI are discussed below and include: (a) learner modeling, (b) learner modeling and instructional variables, and (c) expert systems.

_Learner modeling._ An intelligent computer-assisted learner modeling program, collects information about the student's work, hypothesizes what the student knows, and analyzes the student's thinking processes. The computer program then utilizes this information in selecting appropriate teaching sequences and strategies for that particular student (Alessi & Trollip, 1985; O'Shea & Self, 1983). Consequently, the "value" of using the learner modeling for CAI programs may be to create individualized instruction (Suppes, 1979).

ICAI developers using the student modeling approach also strive to capture and represent the knowledge of an expert in the content area being taught so that the program can interact dynamically with the user, making decisions by referring to the knowledge (Wenger, 1985; Yazdani & Lawler, 1986). Unlike CAI, intelligent tutoring systems may be programmed to answer unexpected questions, draw new inferences, and consequently modify their presentation to meet the needs of the user (Harmon & King, 1985).

DEBUGGY is a system designed by staff at the Xerox Palo Alto Research Center and represents the learner modeling approach to intelligent computer-assisted instruction. It is being designed strictly for research purposes under the assumption that student errors represent "bugs." Consequently, correction of the bug will result in improved performance (Harmon & King, 1985). The DEBUGGY system sets out to explain why a student is making a mistake, as opposed to simply identifying the mistake (Roberts & Park, 1983).

Harmon and King (1985) state that:

> The program depends on a detailed cognitive analysis of the types of errors that students can make. This analysis takes quite a bit of thought and effort, but once it is done, it makes it possible to develop a program that can interact with any particular student to figure out exactly what problems that student is having. (p. 242).

Acknowledged shortcomings of learner modeling systems discussed in the literature include the following:

1. The instructional material produced according to the student's query or mistake is often at the wrong level. That is, the system assumes too much or too little student knowledge (Sleeman & Brown, 1982).

2. This approach models only one particular conceptualization of the domain which may or may not be appropriate to teach. The systems are not designed to discover and work within the student's own conceptualization (Roberts & Park, 1983; Sleeman & Brown, 1982).

3. This method uses only one instructional strategy regardless of the student's individual differences (Tennyson, in press).

4. The system's tutoring and critiquing strategies only occur post hoc, following student errors and misconceptions (Sleeman & Brown, 1982).

5. User interaction is too restrictive, limiting the student's expressiveness (Sleeman & Brown, 1982).

6. The extreme labor-intensive nature of development is a major concern (Roberts & Park, 1983). The amount of time and effort required to develop a system which incorporates only a small amount of information is enormous.

7. This approach has been demonstrated in only a few highly-structured content areas (i.e., mathematics, electronics) (Roberts & Park, 1983).

8. The hardware and software requirements to run ICAI programs is generally prohibitive for the individual consumer (Roberts & Park, 1983).

9. The "core problem" of learner modeling systems may be "the complexity of actual users of the instructional systems" (Yazdani & Lawler, 1986, p. 200). This complexity is summarized by the total of 130 'bugs' discovered in the domain of place-value subtraction.

Advantages of ICAI learner-modeling have also been discussed. The following possible outcomes of ICAI research appear in the literature:

1. The ability to isolate the following teaching/learning characteristics: (a) student characteristics, (b) instructional strategies, (c) subject matter being taught, and (d) nature of communication between teacher and student (Roberts & Park, 1983).

2. Insights can be gained into how people learn by "providing an immediate, powerful analysis of student response patterns" (Roberts & Park, 1983, p. 11).

3. Formalization and experimentation with problem-solving strategies can be an outgrowth of the ICAI research (Clancey, Shortliffe, & Buchanan, 1979).

4. The exposure to a variety of examples in ICAI usually exceeds what actual experience would provide (Clancey et al., 1979).

Learner modeling and instructional variables. Tennyson and his colleagues at the University of Minnesota have developed a computer program entitled the Minnesota Adaptive Instructional System (MAIS) which incorporates both the learner model and effective concept instruction variables. MAIS is programmed to assess student knowledge (i.e. learner modeling) with respect to the information to be learned (i.e. knowledge base). Therefore, the conditions for optimal instruction can be created based on the results of these assessments.

The MAIS has been used in research settings to investigate the effectiveness of certain instructional features and has successfully taught physics concepts to university (Tennyson, 1980) and high school students (Tennyson, 1981), as well as psychology (Tennyson & Buttrey, 1980) and biology concepts (Tennyson & Park, 1984) to secondary-aged students.

Unlike most of the current ICAI development which has stemmed from the computer science field, it is Tennyson's and his colleagues' goal to make contributions to instructional theory and practice by empirically testing variables which influence concept instruction (Tennyson, in press). Therefore, unlike other ICAI research development, their goal is not only to develop software, but to investigate effective instructional design features of such software (Tennyson, in press).

Expert systems. Expert systems may also comprise a form of intelligent computer-assisted instruction. In fact, most ICAI development has involved the development of methods to enhance the components of an expert system environment (Tennyson, in press). Expert systems comprise one component of the artificial intelligence field. Artificial intelligence (AI) is an area of computer science concerned with the development of computing systems that replicate certain human characteristics which are commonly associated with intelligent behavior, namely--understanding, learning , language, reasoning, and solving problems (Barr & Feigenbaum, 1981).

Expert systems may be described as computer programs which replicate experts' knowledge of a domain (Sowizral & Kipps, 1986). Programmers who develop expert systems seek to replicate the problem-solving or decision-making processes conducted by those knowledgeable and experienced in the particular field.

Human experts use two types of knowledge: "facts, or assertions, about their area of expertise . . . and . . . rules of inference that allow them to reason within that domain" (Sowizral & Kipps, 1986, p. 28-29). Facts are usually contained in specialized textbooks and journals, whereas the rules of inference are often learned by practical experience (Kidd, 1984). The rules of inferences are also called heuristics or rules of thumb (Waterman & Jenkins, 1986). Both types of knowledge, facts and rules of inferences, are used to develop expert systems (Stefik et al, 1983). Although different techniques exist representing expert knowledge, most programmers code the knowledge into a set of "if-then" rules (Thompson & Thompson, 1985; Waterman & Peterson, 1986). The inference engine executes these sets of rules (Hofmeister, 1986; Sowizral & Kipps, 1986). The knowledge base, therefore, remains explicit or separate from the inference engine (Hofmeister, 1986; Waterman & Jenkins, 1986).

Expert systems may be designed by "picking the brains" of a small group of experts (Waterman & Peterson, 1986). Hypothetical situations are usually presented to the experts and they are asked to make a decision and then justify that decision. Through this process the expert system developers can analyze the experts' decision processes and create the necessary set of facts and heuristic rules.

Once a prototype of the expert system has been developed, it is then systematically tested. This requires both formative and summative evaluation which emphasize inter- and intra-reliability, as well as, validity of the system (Hofmeister, in press; Parry, 1986b).

Areas identified as appropriate for application of expert systems involve diagnosis, interpretation, prediction, planning, instruction, monitoring, and design (Hayes-Roth, Waterman, & Lenat, 1983). All of these areas are appropriate applications in the field of education (Hofmeister & Ferrara, 1986). In particular, the following areas have been suggested as potential applications to educational expert system development: (a) diagnosis of exceptional learners, (b) recommendations regarding due process procedures, (c) skill assessment, (d) behavioral intervention recommendations, (e) selection and evaluation of instructional materials, (f) improved instructional effectiveness suggestions, (g) staff evaluation, (h) student retention, (i) student course of study counseling, and (j) curriculum development and revision (Ragan & McFarland, in press).

Although there exists many potential educational applications of expert system, relatively few systems have been developed for education (Thorkildsen, et al., 1985; Parry, 1986a). Examples of educationally relevant expert systems that are at least at the prototype stage include:

1. The Computer-Assisted Planning for Educational Resources (CAPER) provides instructional programming recommendations for students prior to special education placement (Haynes, Pilato, & Malouf, in press).

2. The Smart Needs Assessment Program (SNAP) recommends the type of training regular educators who

are serving handicapped students need (Haynes et al., in press).

3. Mandate Consultant (Parry, 1986a) provides advice regarding the individualized education program development procedures mandated by P. L. 94-142 and Utah Rules and Regulations (Parry, 1986b).

4. Behavior Consultant (Ferrara, Serna, & Baer, 1986) recommends behavioral techniques for modifying inappropriate classroom-type behaviors (Serna, 1986).

5. CLASS.LD2 (Ferrara & Hofmeister, 1984), provides second-opinion advice regarding the appropriateness of a learning disabilities classification based on Utah and federal regulations related to P. L. 94-142, as well as expert opinion (Hofmeister & Ferrara, 1986)

The validity of CLASS.LD2 was assessed using actual student file information. Of 264 students, disagreement between the multidisciplinary team and the advice obtained from CLASS.LD2 occurred in 78 cases. These 78 were then evaluated by three experts and their decisions were compared with the advice obtained from CLASS.LD2. Results indicated that (a) The CLASS.LD2 advice was in agreement with the experts more often than with the multidisciplinary teams; (b) CLASS.LD2 decisions significantly correlated with the experts' decisions; and (c) In the six cases in which the experts unanimously disagreed with the expert system, CLASS.LD2 conformed more strictly with Utah and federal rules and regulations (Martindale, Ferrara, & Campbell, 1986).

Expert systems are designed primarily to solve problems for the user (Thorkildsen et al, 1985). This, however, is not its only function. For example, the system can be used as a "tool that guides and simulates decision making by its ability to explain the lines of reasoning it uses to arrive at each decision it makes" (Waterman & Jenkins, 1986, p. 95).

Expert systems contain "practically all existing knowledge in certain well-defined areas . . . [and] . . . [t]he program is therefore an 'expert' in that field of knowledge" (Alessi & Trollip, 1985, p. 45). Because the system contains knowledge about a particular topic, as well as logical connections of this information, it could be used to not only provide expert advise, but allow students to converse with it (Alessi & Trollip, 1985).

The knowledge base of an existing and validated expert system can be used to develop an ICAI program. That is, the expert system contains information (i.e. the rules, attributes, examples, and values) which can guide the instructional design analysis (Ragan & McFarland, in press).

Researchers have developed and modified expert systems for use as training tools. As discussed earlier, MYCIN, a medically-based expert system, was adapted for instructional purposes as an intelligent computer-assisted instructional program entitled NEOMYCIN. (Davis et al., 1975; Clancey &

Letsinger, 1981). In certain ways, NEOMYCIN is simply MYCIN rearranged for tutorial purposes. NEOMYCIN contains all of the knowledge base and the inference engine of the MYCIN system. In addition, it contains all the actual case experiences accumulated through consultations with MYCIN. Consequently, NEOMYCIN has access to several hundred examples from which to draw for instruction. Also included in NEOMYCIN, which is missing from MYCIN, is an additional inference engine designed to manage the tutorial portions of the interactions with the learners (Harmon & King, 1985).

Initially, NEOMYCIN selects a case at random. The learner attempts to diagnosis the patient's problems by analyzing the data given by the computer. When using MYCIN, the physician is asked questions about the patient by the system and the system provides the concluding diagnosis. With NEOMYCIN the process is opposite, the learner asks the system the questions and identifies the diagnosis. At the same time, the system solves the problem in the way MYCIN would. NEOMYCIN, in this process, develops a decision tree and everytime the student asks the system a question, NEOMYCIN compares the MYCIN decision tree with the route the student is taking. If the student asks a question that is obviously irrelevant or unnecessary, NEOMYCIN will ask the student why that information is desired and that it is irrelevant and why. When the student is ready to make a diagnosis, NEOMYCIN will compare the student's diagnosis with MYCIN's. If they are in disagreement, NEOMYCIN informs the student that the diagnosis is incorrect and why (Davis et al., 1975; Harmon & King, 1985).

Other applications of expert systems for training purposes have been advocated. These applications include (a) exposure to the system by running consultations, (b) incorporating the system as part of a training package, and (c) modifying the existing system into a concept instruction training tool (Alessi & Trollip, 1985; Prater & Ferrara, 1986). The purpose for training in each of these applications is to train the students to replicate the decisions produced by the system. Although these training applications have been advocated, no empirical studies examining the effectiveness of expert systems as training tools have been located.

## Expert System Technology and Concept Instruction

The development of expert systems and effective concept instruction have many similarities. That is, the processes of concept analysis and knowledge engineering are very similar. For example, as one analyzes a concept for instructional purposes, the critical attributes must be identified and the definition must be created. As a knowledge engineer interviews experts, the critical attributes of their decisions must be identified and rules must be created. These and other similarities appear in Table 1.

Previous use of expert system technology for training applications has been to attach a sophisticated and costly front-end tutorial program to an existing expert system (i.e. NEOMYCIN) or to develop a simulation-type expert system for the purpose of training (i.e. STEAMER). It is suggested that expert systems may be designed to provide second-opinion advice and at the same time provide training applications without development of a sophisticated front-end. This approach would involve relatively small modifications of the expert system and the development of printed materials. Effective concept instruction would be used in the development of the system and materials.

## Summary

Several types of computer-assisted instruction are presently being used in instructional settings. A distinction can be made between conventional computer-assisted instruction, simulations, and intelligent tutoring systems. Expert system technology, a form of artificial intelligence, has been applied to both intelligent tutoring systems and simulations for training purposes.

Training applications of existing expert systems have only been demonstrated by the sophisticated modification of MYCIN into the intelligent computer-based instructional tool

Table 1

Concept Analysis and Knowledge Engineering Compared

| Concept Analysis | Knowledge Engineering |
|---|---|
| Identification of critical and variable attributes of the concept. | Identification of critical and variable attributes of the expert's decisions. |
| Creating the concept's definition in terms of rules. | Creating rules to represent the expert's decisions. |
| Creating examples and nonexamples of the concept. | Creating cases against which the prototype system can be tested. |
| Defining the teaching sequence. | Defining the appropriate order of the rules. |
| Testing the learners' ability to accurately classify examples and non-examples of the concept. | Testing the system's ability to replicate the expert's decisions. |

entitled NEOMYCIN. Additional training applications of expert systems which are less costly have been advocated but have yet to be tested empirically. One of the training applications advocated is the combination of concept instruction and expert system technology. Knowledge of effective concept instruction and computer-based instruction could be combined to create an effective expert system-based training package.

# METHODS

The methods section includes a discussion of the (a) subjects, (b) materials, (c) data and instrumentation, (d) research design, (e) procedures, and (f) analysis of data. The formative evaluation portion of this study is discussed in the Materials and Analysis of Data Sections.

## Subjects

### Target and Accessible Population

The target population is defined as the population to which the results of a study can be generalized, whereas the accessible population are those subjects from the target population who are available to the researcher (Bracht & Glass, 1968). Generalizing from the accessible to the target population requires knowledge about the characteristics of both populations.

Undergraduate and graduate students in regular education, special education, communication disorders, and psychology comprised the target population. The accessible population were those who volunteered for participation in this study. Characteristics of the target population were defined by the researcher and compared against the demographic information obtained on the accessible population. This comparison assisted in assessing the external validity of the study.

Sample

The accessible population, or those who volunteered for participation, and the sample of subjects consisted of the same group. Ninety-seven students from three universities--Utah State University, St. Cloud State University (Minnesota), and the University of South Dakota--participated as subjects in the study. These universities were selected because of the researcher's contact with faculty and the willingness of faculty to participate. Subjects included experienced teachers and inexperienced teacher trainees in regular education, special education, communication disorders, and psychology.

For those subjects at Utah State University a one-hour, tuition-free course was offered through the special education or psychology department. Announcements were made in the previous quarter's classes and flyers were distributed throughout the College of Education. At St. Cloud and The University of South Dakota, students who were registered for a course in special education assessment were given the option of participating as subjects in the study. Completion of the training fulfilled one of several assignments from which they could select. Once the pool of subjects was defined, each subject was randomly assigned to one of the two experimental groups.

## Materials

Training materials included the LD.Trainer package and representative special education student files for the CLASS.LD2 group. The LD.Trainer package underwent four formative evaluation stages. A description of the process by which the materials were developed or selected and the formative evaluation plan of LD.Trainer follow.

### LD.Trainer

First, the selection of the content to be covered was made. Second, strategies of effective concept instruction were applied. Third, the actual modification of the expert system, CLASS.LD2 was completed; and fourth, the materials were formatively evaluated and revised.

Selection of content. In order to provide a group of information to the subjects at one time, the knowledge base of CLASS.LD2 was broken into conceptual components. Because CLASS.LD2 contained information relating to more than that in the Utah Rules and Regulations (Utah State Board of Education, 1981) and because it was desirous to keep the amount of material and time spent by those in the study at a minimum, the developer of LD.Trainer selected those components most needed to comply with the Utah Rules and Regulations in classifying a learning disabled student. From these components three lessons were developed with subcomponents as appear in Table 2.

Table 2

Lessons and Sublessons

| Lesson Number | Lesson Title |
|---|---|
| 1 | Discrepancy and IQ |
| 1a | 40% below actual grade placement |
| 1b | 40% below expected grade placement |
| 1c | IQ score |
| 1d | Summary of discrepancy and IQ |
| 2 | Other Handicapping Conditions Exclusions |
| 2a | Sensory impairments |
| 2b | Physical and health impairments |
| 2c | Communication and behavior disorders |
| 2d | Missing data |
| 2e | Summary of other handicapping conditions |
| 3 | Other Exclusions |
| 3a | Economic and environmental |
| 3b | Cultural |
| 3c | Alternative services |
| 3d | Summary of other exclusions |

Thirteen sublessons were developed using a model of effective concept instruction. For each lesson, definitions and examples and nonexamples of that definition were created. The examples and nonexamples were matched on irrelevant attributes. Critical attributes within the examples, nonexamples, and definitions were highlighted. A chart describing the attributes of the LD.Trainer materials as compared with the CLASS.LD2 materials appears in Appendix A.

Each lesson in the LD.Trainer materials was divided into two components: instruction and practice. During both, subjects entered information into the computer as prescribed by the written materials and viewed how the outcomes (advice and confidence factors) varied just by manipulation of the value of one or two of the critical variables. In order to use the CLASS.LD2 expert system in this manner, modifications in the expert system were made. A discussion of these modifications and the format for the instruction and practice portions of each lesson follow.

*Modification of CLASS.LD2*. The expert system, CLASS.LD2, was modified for training purposes. These changes resulted in a simulation-type of software. The steps for modifying CLASS.LD2 included the following:

1. Hypothetical student data were created which represented a matched example and nonexample of the concept being taught. That is, the example and nonexample were matched on the irrelevant attributes, but the critical

attributes (the attributes being taught in that particular lesson) varied. In Lesson 2.a, for example, all attributes were held constant except for the decibel hearing loss. One example of a learning disabled student had a 12 decibel hearing loss in the better ear and the nonexample student had a 40 decibel hearing loss in the better ear. Variation of this critical attribute, hearing loss, although all the other attributes were held constant, made one student eligible, and the other not eligible for a learning disabilities classification.

2. Using the hypothetical data generated in step #1, a consultation was initiated with CLASS.LD2 and a record of the consultation stored. M.1 (Teknowledge, 1986), the authoring tool used to develop CLASS.LD2, is programmed to store information in what is called the dynamic memory or the cache. The cache contains all the conclusions the system has derived either from asking the user or "inferring" the answer based on the programmed rules. Therefore, once a record of the consultation is stored, the cache contains expressions such as the following:

'IQ test score' = 87 cf 100 because 'you said so'

age = 157 cf 100 because 'you said so'

'grade placement' = 8.2 cf 100 because 'you said so'

'basic reading' score = 3.4 cf 100 because 'you said so'

'discrepancy actual = 58.5366 cf 100 because rule-685 'discrepancy estimated' = 48.4653 cf 100 because rule-720

3. The expressions representing the critical attributes and those expressions used by the system to "infer" the critical attributes being taught were deleted from the cache. For example, Lesson 1.b was designed to present the expected discrepancy concept. In order to compute the expected discrepancy, the following information is needed: IQ score, age, grade placement, and test score. Therefore, in the example presented in step #2, the following lines were deleted from the caches for Lesson 1.b: IQ test score, age, grade placement, basic reading score, and discrepancy estimated. Then, the system had information relating to everything about this instance EXCEPT those five values.

4. This modified cache was saved and used as instances and practice items corresponding to the LD.Trainer written materials. When the modified cache was reloaded, because the system did not have values for the expressions deleted, the system asked the user the questions corresponding to these expressions. Therefore, in the example presented in item #3, the user inputted different IQ scores, yet kept the other four values constant, and created an example and a nonexample of a learning disabled student.

5. A program with menus to load the various lessons, instances, and practice items was developed. The menu files, modified caches, and the authoring tool (M.1) fit onto two

floppy disks. A graduate student in computer science completed this step of the development.

Instruction. During the instructional part of each lesson, the objectives were presented first with the information one would need to respond to the questions listed. Then the definition or definitions were stated with the critical attributes in bold print. Next, brief instances, one example and one nonexample matched on the variable attributes, were presented. Included were explanations of why each was either an example or a nonexample. A sample lesson appears in Appendix B.

Next, instances were worked on the computer. A brief description of an example or nonexample was provided and a chart giving the values of the critical attributes for that instance listed. The subject used the computer by selecting the appropriate lesson and then the instance or practice item from the main and lesson menus, respectively. The LD.Trainer computer system's instances and practice items corresponded with those provided in the printed materials.

Each lesson that was not a summary lesson contained four instances, two matched examples and nonexamples. The summary lessons contained six instances in the instruction. Only those questions pertaining to the lesson appeared on the computer. So, for example, in Lesson 1.D which deals with the appropriate IQ level, the CLASS.LD2 questions for every instance that appeared on the screen included: "What is the student's most recent IQ score?" and "What test was used to measure the student's level of intellectual functioning?". If the IQ score was below the required criterion for an LD classification, the following additional questions appeared: "Is there data, in writing, to support the judgment that the student's true IQ score is above _____?" and "What is your best estimate of the student's true IQ score?". By keeping all of the other variables constant (i.e. grade placement, age, test score), the subjects walked through the process of changing the IQ and related data to create examples and nonexamples of a learning disabled student.

Practice. The second portion of each lesson included the practice section. For each lesson, both the subcomponent and the review lessons, two practice items were presented. The subjects were given a brief description of the student, similar to the instances in the instructional component. But rather than responding to the questions on the computer, the subjects first made a decision about the appropriateness of a learning disabled classification and justified their decision by writing it on the practice pages. Then they inputted the information into the computer and compared their answer with the computer's. Some feedback was also provided in the written materials (i.e., "You should have obtained the following advice because..."). The subjects then inputted different values for that practice item and again could compare the outcomes based on variations in the critical attribute values.

Formative evaluation. Three stages of formative evaluation of the LD.Trainer package occurred. First, after the instances for the lesson materials were generated, novice undergraduates received the instances along with the definition (Appendix C). They attempted to identify the examples (learning disabled student) from the nonexamples (not learning disabled student) based on the definition only. This step is recommended by Merrill and Tennyson (1977) as a way of estimating difficulty level of each instance. Merrill and Tennyson recommend including a range of difficulty with the instances in order to control for possible undergeneralization.

During the second formative stage, two education university faculty members at the University of South Dakota evaluated the proposed plan for development of the lessons and the process by which it was intended to be carried out (Appendix D). They were given the process plan, as well as a sample lesson. This enabled them to not only evaluate the process but the lesson format as well. Modifications were made in both the process and the format based on the evaluators' comments.

The last formative stage included a pilot test of the LD.Trainer package. Four undergraduate students in special education and related fields completed some of the lessons and evaluated them in terms of clarity, correctness, appropriate number of instances, and use of the computer to learn these skills (Appendix E). Because all of the lessons followed the same format (i.e. same number of instances, definitions presented in a similar manner, similar directions), it was appropriate to pilot test a sampling of the materials.

During the pilot test the subjects not only evaluated the materials, but completed them as prescribed by implementation of the LD.Trainer package. Therefore, the pretests and posttests were given and the results analyzed in terms of effectiveness of the training. Undergraduate students were offered one free credit hour in special education as an incentive to participate in the training.

All of the materials, except the practice exercises, were bound by lesson and color coded with different colored covers for each lesson. The practice exercises were then duplicated and distributed as needed by the subjects. The practice exercises were not included in the bound materials because they were expendable.

## CLASS.LD2 Files

Sixteen special education student files provided the materials for the CLASS.LD2 group. Forty-three teacher-selected special education student files from three elementary, one middle, and one high school in a large Utah school district which contained either a current IEP, or a clear statement that the student did not qualify for special education, provided the accessible population of files. The files which met these specifications were stratified by learning disabled or not learning disabled (i.e. behavioral disorders, special needs, or didn't qualify for special education) and by elementary or secondary education. Four files from each category were randomly selected to be included in the packet of files for the CLASS.LD2 group. This process created an equal number of examples and nonexamples and provided a representation across varying ages.

Each set of the 16 files was randomly ordered and bound in two books similar to the lessons for LD.Trainer. The covers of the books were color coded.

### Computer Hardware

In order to complete either experimental training, IBM compatible computers with dual disk drives were necessary. The LD.Trainer required at least 512k memory and the CLASS.LD2 program at least 256k memory.

## Data and Instrumentation

### Pre and Posttest

A domain-referenced test was developed which served as both the pretest and the posttest for this study (Appendix F). All subjects completed the same test. The test was developed in such a way that the subject's performance was keyed to the different elements of the definition of a learning disabled student. This facilitated assignment of the subjects in the LD.Trainer group to those lessons they most needed for remediation.

The test consisted of 12 instances, examples or nonexamples, of a learning disabled student. All components as defined by the Utah Rules and Regulations necessary to make the decision regarding a classification were given. These included: IQ test and score; area of deficiency; grade level and area score OR percent discrepancy between test score and grade placement and between test score and expected grade placement; sensory, health, and physical information; behavioral and communication status; and cultural, economic, and environmental background. Attempts were also made to include a variety of instances, for example by varying gender and representing a range of ages.

The subjects were asked to identify whether or not the student could be classified as learning disabled and to justify their answer (i.e. why or why not). In each nonexample only one critical attribute made the LD classification inappropriate. That is, for each nonexample, all of the critical attributes made an LD classification appropriate except one. This facilitated keying the test to the lessons in LD.Trainer and to determining misconceptions as part of this study.

In addition to the 12 LD or not LD instances, two of the instances contained an additional question. One question related to the appropriateness of the data (i.e. missing information on vision and hearing tests) and the other to the appropriateness of special education service (i.e. the student meets LD requirements but the team concludes that all regular classroom interventions have yet to be tried). These two concepts were included as part of the CLASS.LD2 expert system and the LD.Trainer package.

Test-retest reliability. The reliability of the domain-referenced test was assessed through the test-retest procedure (Ebel, 1979). Six graduate students who did not participate in the study completed the test twice. A Pearson product moment correlation between scores on the two administrations of the test was computed. This provided an estimate of the test-retest reliability. Modifications would have been made in the test if the test-retest reliability were low. This was, however, not necessary.

Validity. In order to assess the validity of the domain-referenced test, the developer ran each test item through a consultation with the CLASS.LD2 expert system. Then the correct outcomes as determined by the developer were compared with the outcomes obtained by the expert system. Because the expert system provided not only learning disabilities advice but confidence factors associated with each piece of advice, only those test items corresponding to (a) learning disabilities advice with 95% confidence (the highest confidence possible) or (b) learning disabilities not appearing at all were included in the test.

Inter-reader reliability. Because the subjects were asked to justify their responses as to whether the student was LD or not LD, the scoring of the tests required some subjective judgment. Therefore, three readers scored all of the tests. When there was disagreement, the item was scored according to the majority (two out of three). Also, percent of agreement across all the items and correlations on test scores across the three readers were computed.

### Demographic Information

Each subject was asked to complete a demographic form which included the area and state of certification; and the type, level, and number of years of teaching and administrative experience (Appendix G). This information was used to distinguish between those subjects with and without teaching experience and to describe the characteristics of the subjects.

## Research Design

In order to answer the research questions, the pretest-posttest control group design advocated by Campbell and Stanley (1963) was used. However, rather than one experimental and one control group, two experimental groups were used. The experimental groups consisted of the LD.Trainer group and the CLASS.LD2 group. The subjects assigned to the LD.Trainer group completed those training materials and the CLASS.LD2 group ran consultations on the expert system, CLASS.LD2, with special education student files given to them.

Subjects in the two experimental groups were further divided into two groups, experienced and inexperienced teachers. Therefore, four groups were created. The experimental design is depicted in Table 3.

Table 3

The Pretest-Posttest Group Design

| | | Pretest | Training | Posttest |
|---|---|---|---|---|
| Experienced Teachers | LD.Trainer | O | $X_1$ | O |
| | CLASS.LD2 | O | $X_2$ | O |
| Inexperienced Teachers | LD.Trainer | O | $X_1$ | O |
| | CLASS.LD2 | O | $X_2$ | O |

Although Campbell and Stanley (1963) state that this true experimental design controls for eight major threats to internal validity, random assignment is necessary in order to assure that differences between the groups may be attributed only to chance assignment. For example, even though subjects were randomly assigned to one of two treatment groups, selection could be a threat to internal validity if the subjects in each group happened to differ on a variable of importance. Random assignment does not prevent selection from being a threat to internal validity. Analysis of the demographic information assists in making this conclusion.

## Procedures

After the accessible population was defined, each subject was assigned a number and then, using a table of random numbers, each was randomly assigned to one of the two groups: the LD.Trainer group or the CLASS.LD2 group.

Each subject was given the pretest and asked to respond to each item to the best of their ability. They were verbally assured that they were not expected to know each of the items initially. For those in the LD.Trainer group, the items were corrected immediately and the appropriate lessons assigned.

After completing the pretest the subjects then selected a computer and were given their appropriate packet of materials. For the LD.Trainer group this consisted of the following: two LD.Trainer disks, all three lessons bound separately, an assignment sheet, and a directions sheet. The CLASS.LD2 group was given two CLASS.LD2 disks, two bound books of student files, an assignment sheet, and a directions sheet.

Instructions

Those participating in the LD.Trainer group received two sets of instructions, verbal and written. During the verbal directions the researcher explained that each lesson they would be completing covered only a portion of the LD definition (i.e. one of the three requirements for LD classification). They were verbally directed in getting started. These directions also were given in written form for them to refer to when needed (Appendix H). The researcher was available to answer direct questions, pertaining to either computer difficulties or questions regarding the materials.

The subjects participating in the CLASS.LD2 group were also given both verbal and written directions. They were told to respond to the questions on the computer using the information given to them in the files. They were also instructed in the use of the commands WHY, SHOW, and LIST and encouraged to use them throughout the demonstration.

These three commands provide the user with additional information about the consultation (Appendix I). For example, when the WHY command is used with any question posed by the expert system, an explanation of why the question is being asked will be given to the user. When one types the SHOW command, all of the information in the cache up to that point in the consultation is shown. And the LIST command can be used to list any rule.

The written directions for the CLASS.LD2 group included mechanical information on starting the system, running a consultation, saving the cache, and repeating the process (Appendix J). These steps were also discussed verbally. The researcher was available to answer questions posed by the subjects in this group. Again the questions related to either computer problems or difficulty interpreting the file information.

Record Sheets

Each subject was given a record sheet. The LD. Trainer record sheet included the lessons the subjects were to complete. They were asked to self-record the amount of time spent on each lesson by listing the beginning and ending time on the record sheets (Appendix K).

The CLASS.LD2 subjects were also given a record sheet and asked to self-record their beginning and ending times for each file. On this form they also recorded the team decision (off of the IEP or the referral form) and the computer's advice and confidence. Since each of the files was bound in a notebook in a random order, the subjects recorded the file number next to the other information (Appendix L).

The subjects were told that they must either complete the materials assigned to them (eight LD.Trainer lessons or 16 CLASS.LD2 files) or they must spend the amount of time given to them to work on the materials (four to four and one-half hours). They were encouraged to take breaks when they felt they needed them.

After the directions had been given, the subjects worked independently. The researcher was available to answer any questions or to assist the subjects if necessary.

Upon completion of the materials or at the end of the time period, the subjects were administered the posttest. The conditions remained the same as during the pretest. That is, they were not given any additional prompts or materials from which to work.

## Analysis of Data

### Formative Evaluation

The results of the three steps of formative evaluation were analyzed. The first stage involved undergraduate students identifying instances as examples or nonexamples based on the definition only. This provided an estimate of difficulty level of the instances as advocated by Merrill and Tennyson (1977). The percent of respondents correctly identifying each instance was computed. The following scale was then applied and an estimate of difficulty assigned to each instance: 49% to 0% - difficult item, 79% to 50% - moderate item, and 80% to 100% - easy item. Approximately equal numbers in each category were used in each LD.Trainer lesson.

The second formative evaluation stage involved two special education faculty members from the University of

South Dakota. They evaluated (a) the process plan for development of the LD.Trainer materials, and (b) a sample lesson. The sample lesson was also demonstrated on the computer. No formal forms or analysis of their comments were made. However, their evaluation was used to improve the development process and the format of the materials.

The third, and final stage, of the formative evaluation involved a pilot study. Four undergraduate majors in special education or related fields completed the training. Upon completion of each lesson they were asked to evaluate the lesson using a Likert-type scale in terms of number of instances, correctness, clarity, and use of the computer (Appendix D). The rating scale number corresponding to each lesson and each question was tabulated. If two of the four respondents assigned a one or a two to the items corresponding to instructions, definition, and written materials (strongly disagree or disagree that the instructions, definitions, and written materials were clear, understandable, and correct), that lesson was modified. If two of the four assigned a one or a two (not enough items) or a four or a five (too many items) to the number of instances and practice items questions, modifications were also made.

## Experimental Design

Statistical significance. When one obtains statistical significance the results may be defined as an unlikely chance occurrence assuming the null hypothesis to be true or that a true difference between the groups exists (Ferguson, 1981). Once the results of a study are obtained and an inferential test of significance applied, researchers can determine whether or not the result is a likely chance occurrence at the level at which alpha, or the level of significance, was set.

The results of this study were analyzed using a 2 by 2 analysis of variance (ANOVA) test of statistical significance. The first three research questions and hypotheses were answered through this statistical technique. That is, the main effects for the groups and the level of experience addressed the first two hypotheses and the interaction effect addressed the third.

Because of the mathematical development of ANOVA, certain assumptions are made. These include (a) the population from which the sample was drawn has a normal distribution across the dependent variable, (b) the population variances are equal (homogeneity of variance), and (c) the various factors of the total variance are additive or, in other words, independence of observations exists (Ferguson, 1981; Glass & Hopkins, 1984).

Glass, Peckham, and Sanders (1972) reviewed studies examining the consequences of not meeting the ANOVA assumptions. They concluded the following: (a) ANOVA is "robust" enough to compensate for nonnormality, particularly if the sample size is large, nondirectional (two-tailed) tests are employed, and the population distribution is not highly skewed; (b) Violation of the homogeneity of variance does not occur if the sample sizes in each cell are equal; and (c) The independence of observations assumption is met when interventions are administered individually.

The first and third ANOVA assumptions were not a concern in this study because (a) the sample size was fairly large (i.e., N = 97), (b) a nondirectional test was used, and (c) the training was conducted individually. Because it was not possible to obtain equal sample sizes in each cell, the amount of departure from homogeneity of variance, was also tested (Glass & Hopkins, 1984).

In addition to the ANOVA, a priori multiple comparison tests were conducted to address the fourth and fifth research questions and hypotheses. The statistical test selected for this analysis, assuming that the appropriate assumptions are met, was planned orthogonal contrasts because they are the most powerful test of mean differences (Glass & Hopkins, 1984).

In order to use planned orthogonal contrasts, each contrast must be orthogonal, or independent, to every other contrast. The value of two contrasts are orthogonal when "the products of the corresponding contrast coefficients sum to zero" (Glass & Hopkins, 1984). Another restriction of planned orthogonal contrasts is that all the desired cells) contrasts. Since in this study there are four cells (i.e., three or less possible contrasts) and only two contrasts will be made, this restriction is met.

For each statistical test, alpha was set at the .Ø5 level ($p < .05$). That is, the researcher failed to reject the null hypothesis if the results obtained were a likely chance occurrence at the .Ø5 level. One is never able to truly accept a null hypothesis, only fail to reject it (Weinberg, Schumaker, & Oltman, 1981). The null hypothesis was rejected if the results were an unlikely chance occurrence at the .Ø5 level.

Educational significance. In addition to the statistically significant results, the data were examined in terms of educational significance. Educational or practical significance may be viewed as the value, benefit, or cost of the results obtained (Shaver, 1985). Educational significance is important to consider because statistical significance only indicates whether the results are a chance occurrence and are a function of sample size. Computing educational significance assists in the interpretation of the results.

In order to obtain an estimate of educational significance, standardized mean differences (SMDs) were computed for each hypothesis. The formula used in this study appears in Table 4. This is the appropriate formula for computing SMDs in a pre/posttest design because (a) an adjustment for possible pretest differences is included in

Table 4

Standardized Mean Difference Formula for Pretest-Posttest Designs.

$$SMD = \frac{[(\bar{X}_{1post} - \bar{X}_{1pre}) - (\bar{X}_{2post} - \bar{X}_{2pre})]}{[(s_{1pre} + s_{2pre} + s_{2post}) / 3]}$$

Key:

$\bar{X}_{1pre}$ = pretest mean of the LD.Trainer group

$\bar{X}_{1post}$ = posttest mean of the LD.Trainer group

$\bar{X}_{2pre}$ = pretest mean of the CLASS.LD2 group

$\bar{X}_{2post}$ = posttest mean of the CLASS.LD2 group

$s_{1pre}$ = pretest standard deviation of the LD.Trainer group

$s_{2pre}$ = pretest standard deviation of the CLASS.LD2 group

$s_{2post}$ = posttest standard deviation of the CLASS.LD2 group the numerator and (b) the denominator includes a comprehensive pooled standard deviation of untreated conditions (K. White, personal communication, October, 17 1986).

## Summary

Validated design strategies for concept instruction and expert system technology were combined to create an instructional package entitled, LD.Trainer. The materials for LD.Trainer were formatively evaluated and then tested against another form of expert system training.

A pretest and posttest experimental research design was employed to determine whether LD.Trainer or CLASS.LD2 was the most effective method of training experienced and inexperienced teachers to accurately classify learning disabled students in a selected instructional environment. The results were examined in terms of statistical and educational significance.

# RESULTS AND DISCUSSION

Within this chapter the following are discussed (a) description of subjects, (b) test-retest reliability and inter-reader agreement, and (c) verification of the independent variable. In addition, each of the five hypotheses, as well as possible rival hypotheses, are evaluated; plus internal and external validity strengths and concerns are discussed.

## Description of Subjects

Ninety-seven university students from three universities participated as subjects in this study. This included 42 students from Utah State University, 34 from St. Cloud State University in Minnesota, and 21 from the University of South Dakota. Fifty students completed the LD.Trainer and 47 completed the CLASS.LD2 materials. The difference between the size of the two groups was attributed to three subjects (two at Utah State and one at the University of South Dakota) who registered for participation, were randomly assigned to a group, but who never began the training. All three of these subjects had been randomly assigned to the CLASS.LD2 group. The breakdown of university by group appears in Table 5.

Table 5

Number of Subjects by University and Group

| Group | University | | | Total |
|---|---|---|---|---|
| | Utah State | St. Cloud | So. Dakota | |
| LD.Trainer | 21 | 17 | 12 | 50 |
| CLASS.LD2 | 21 | 17 | 9 | 47 |
| Total | 42 | 34 | 21 | 97 |

Table 6

Number of Subjects by Experience Level and Group

| | Group | | Total |
|---|---|---|---|
| | LD.Trainer | CLASS.LD2 | |
| Experienced* | 22 | 20 | 42 |
| Inexperienced | 28 | 27 | 55 |
| Total | 50 | 47 | 97 |

*Note*. *Experienced was defined as having taught at least one year.

Experience

Of the ninety-seven subjects, forty-two (42) had taught for at least one year and fifty-five (55) were inexperienced. The breakdown of experience with the randomly assigned treatment group is listed in Table 6. Of those who had at least one year of experience, type of experience by assigned group appears in Table 7.

Table 7

Number of Subjects by Group and Type of Experience

| | Type of Experience | | | |
|---|---|---|---|---|
| Group | Special Ed.* | Other** | None | Total |
| LD.Trainer | 11 | 11 | 28 | 50 |
| CLASS.LD2 | 8 | 12 | 27 | 47 |
| Total | 19 | 23 | 55 | 97 |

Note. *Special education category included psychologists and social workers. **Other category included elementary, secondary teachers and administrators.

The total years of experience was also examined. For those with at least one year of experience the mean number of years of experience was 11.25 with a standard deviation of 6.18 years. The years of experience ranged from 1 to 34. The total years of experience is broken down by assigned group in Table 8.

Table 8

Total Years of Experience by Group

| | Group | | Total |
|---|---|---|---|
| | LD.Trainer | CLASS.LD2 | |
| M | 11.05 | 11.47 | 11.25 |
| SD | 4.52 | 7.73 | 6.18 |
| Range | 3 - 19 | 1 - 34 | 1 - 34 |
| N | 21* | 19* | 40* |

Note. The figures in this table relate only to the subjects who had at least one year of experience. *Two experienced subjects, one in each group, did not report total number of years of experience.

Certification

Type of certification of subjects by group appears in Table 9. Of the 97 subjects, 14 were certified in special education, 30 had other certifications (i.e., elementary education, secondary education), and 53 were not certified.

Table 9

Type of State Certification Obtained by Group

| | Group | | Total |
|---|---|---|---|
| Type of Certification | LD.Trainer | CLASS.LD2 | |
| Special Education | 8 | 6 | 14 |
| Other* | 15 | 15 | 30 |
| None | 27 | 26 | 53 |

Note. *Other category included elementary, secondary, and administrative certification.

Education

The subjects reported (a) the last postsecondary degree they obtained, (b) their present class standing, and (c) their major subject. Most of the subjects had not obtained a postsecondary degree (N = 42) and equal numbers of subjects had a bachelors or a masters degree (N = 23). This information by groups appears in Table 10. Most of the subjects were seniors (N = 28). In addition, 25 subjects were not presently seeking a degree. All of these subjects had obtained at least a bachelors degree and were working toward an administrative credential. The number of subjects by class standing and group appears in Table 11. Table 12 lists the major subject in college by groups. Most of the subjects were regular education majors (N = 51).

Table 10

Highest Postsecondary Degree Obtained by Group

| | Group | | Total |
|---|---|---|---|
| Highest Postsecondary Degree Obtained | LD.Trainer | CLASS.LD2 | |
| None | 20 | 22 | 42 |
| Associate | 5 | 3 | 8 |
| Bachelors | 15 | 8 | 23 |
| Masters | 9 | 14 | 23 |
| Total | 49* | 47 | 96* |

*Note*. *One subject did not respond to this question.

## Research Design

As discussed in the methods section, the pretest-posttest design (Campbell & Stanley, 1963) was used for this study. The independent variable was interaction with one of the two sets of training materials, LD.Trainer or CLASS.LD2. The dependent variable was scores on the posttest.

Table 11

Class Standing by Group

| | Group | | Total |
|---|---|---|---|
| Class Standing | LD.Trainer | CLASS.LD2 | |
| Freshman | 1 | 1 | 2 |
| Sophomore | 2 | 3 | 5 |
| Junior | 11 | 8 | 19 |
| Senior | 13 | 15 | 28 |
| Masters | 11 | 4 | 15 |
| Doctorate | 1 | Ø | 1 |
| Not degree seeking | 1Ø | 15 | 25 |
| Total | 49* | 46* | 95* |

Note. *Two subjects, one in each group, did not respond to this question.

Table 12

Major Subject in College by Group

| Major Subject in College | Group | | Total |
|---|---|---|---|
| | LD.Trainer | CLASS.LD2 | |
| Special Education* | 22 | 16 | 38 |
| Regular Education | 25 | 26 | 51 |
| None | 1 | 3 | 4 |
| Total | 48** | 45** | 93** |

Note. *Special education category included communication disorders and psychology. **Four subjects, two in each group, did not respond to this question.

## Dependent Variable - Pre and Posttest

A pre/posttest was designed to measure the effectiveness of the training. Details on the development and content of the test were discussed in the Methods Section. In order to estimate reliability, a test-retest procedure was used. In addition, three readers were used to score the subjects' responses.

### Test-Retest Reliability

A measure of test-retest reliability was assessed using six graduate students who did not participate as subjects in the study. Each student completed the two tests approximately one week apart. Each item was scored as either correct or incorrect and the total correct for each test were summed. Then a test-retest reliability coefficient was computed and resulted in $r = +0.88$. In addition, agreement between the subjects' responses on each item of the two tests were totalled and divided by the total possible agreement. This resulted in a mean agreement of 86% with a range of 100% to 75%.

### Inter-Reader Agreement

Because of the subjective nature of the justification responses on the test, three graduate students scored each test. This allowed for an agreement between two out of three to be used in the final scoring. Of the three readers, two were blind regarding the intent of the study and the third was the researcher.

Percent of agreement was computed two ways. First, the percent of total agreement was computed as follows. The total number of items scored across all subjects was multiplied by three to obtain the total possible number of agreements. Then every disagreement by one of the three readers was subtracted from the total resulting in the total number of agreement. This total represented 98.0% of the total possible agreement. The second method resulted in a more conservative percentage and consisted of using only those items on which all three agreed. This total was divided by the total possible. This resulted in a score of 94.0% agreement across all three readers.

In addition to percent of agreement, correlations were computed between the three readers on both the pretest and the posttest. These correlations were computed using the raw score of total items correct each reader had assigned. The resulting correlations were very high, ranging from + 0.97 to + 0.99. Interestingly, the correlations were equal between the three readers on the pretest and posttest. The correlation matrices appear in Tables 13 and 14.

## Verification of the Independent Variable

Researchers often neglect verifying that the independent variable actually took place (Shaver, 1983). Efforts were made to provide such verification in this

Table 13

Correlation Matrix of Pretest Scores Assigned by Three Readers

| | Pretest | | |
|---|---|---|---|
| | Reader #1 | Reader #2 | Reader #3 |
| Reader #1 | 1.00 | | |
| Reader #2 | .97 | 1.00 | |
| Reader #3 | .98 | .99 | 1.00 |
| N = 97 | | | |

Table 14

Correlation Matrix of Posttest Scores Assigned by Three Readers

| | Posttest | | |
|---|---|---|---|
| | Reader #1 | Reader #2 | Reader #3 |
| Reader #1 | 1.00 | | |
| Reader #2 | .97 | 1.00 | |
| Reader #3 | .98 | .99 | 1.00 |
| N = 94* | | | |

Note. *Three subjects had not completed the training and taken the posttest prior to the two outside readers' corrections of the other subjects' tests. Therefore, three posttests were corrected by the researcher only.

study. Because, however, the subjects independently completed the training materials, the verification came primarily from self-report.

Subjects reported on their record sheets the beginning and ending times for each lesson or file. This provided an estimate of the amount of time each subject spent interacting with the materials. It also provided a check that they indeed did work through the materials. In addition, those in the LD.Trainer group wrote their responses on the practice items. These were checked for completion. The CLASS.LD2 group recorded on the record sheet (a) the student classification assigned by school personnel which they obtained from the case study file and (b) the computer's conclusions. The record sheets were also checked to ensure that the subjects did actually run the consultations.

## Hypotheses

Based on the five research questions, five hypotheses were tested. They included:

1. Using ANOVA, there would be no statistically significant ($p < .05$) difference between the posttest performance of those participating in the LD.Trainer and the CLASS.LD2 groups.

2. Using ANOVA, there would be no statistically significant ($p < .05$) difference between the posttest performance of experienced and inexperienced teachers.

3. Using ANOVA, there would be no statistically significant ($p < .05$) interaction between amount of experience and training method.

4. Using Planned Orthogonal Contrasts, there would be no statistically significant ($p < .05$) difference between the posttest performance of experienced teachers in the LD.Trainer and the CLASS.LD2 groups.

5. Using Planned Orthogonal Contrasts, there would be no statistically significant ($p < .05$) difference between the posttest performance of inexperienced teachers in the LD.Trainer and the CLASS.LD2 groups.

Hypothesis 1: Between Groups

The first hypothesis read that there would be no statistically significant ($p < .05$) difference between the posttest performance of those participating in the LD.Trainer and the CLASS.LD2 groups. Before examining statistical significance, the posttest data are described.

Descriptive data. The descriptive data for both the pretest and posttest by groups appear in Table 15. There were negligible differences between the two groups' mean performance on the pretest. And on the posttest, those in the LD.Trainer group ($\overline{X} = 16.54$) scored higher than those in the CLASS.LD2 group ($\overline{X} = 13.08$). There also appeared to be more variability in the LD.Trainer than in the CLASS.LD2 group (sd = 4.42 and sd = 3.57, respectively).

Table 15

Pretest and Posttest Scores by Group

| | Group | | Total |
|---|---|---|---|
| | LD.Trainer | CLASS.LD2 | |
| Pretest | | | |
| *M* | 10.30 | 10.19 | 10.25 |
| *SD* | 3.34 | 3.58 | 3.44 |
| Variance | 11.15 | 12.82 | 11.83 |
| Range | 0 - 18 | 0 - 18 | 0 - 18 |
| Posttest | | | |
| *M* | 16.54 | 13.08 | 14.87 |
| *SD* | 4.42 | 3.57 | 4.37 |
| Variance | 19.54 | 12.74 | 19.10 |
| Range | 10 - 25 | 2 - 21 | 2 - 25 |

Statistical significance. To determine whether the difference observed between groups was an unlikely chance occurrence, an analysis of variance (ANOVA) was used as a test of statistical significance. Before it was computed, however, the data were analyzed to ensure that the assumptions of ANOVA were met.

The assumptions for ANOVA include (a) the population from which the sample was drawn has a normal distribution across the dependent variable, (b) the population variances are equal (homogeneity of variance), and (c) independence of observations (Ferguson, 1981; Glass & Hopkins, 1984). As discussed in the Methods section, Glass, Peckham, and Sanders (1972) concluded that (a) ANOVA is "robust" enough to compensate, for nonnormality if the sample size is large, nondirectional tests are used and the population distribution is not highly skewed; (b) Violation of the homogeneity of variance only exists if the sample sizes in the cells are not equal; and (c) The independence of observations assumption is met if the interventions are administered individually. Consequently, the first and third assumptions are not a concern in this study. The second assumption, however, homogeneity of variance needed to be evaluated because the sample sizes in the cells were not equal (Table 6). Using the F-ratio to test the difference between the variance of the LD.Trainer (19.54) and the CLASS.LD2 (12.74) groups indicated that this difference was not statistically significant (F = 1.53; F critical at .05 = 1.70).

Because all of the assumptions were met, a two-way ANOVA using the factors of group and experience was computed. This allowed not only Hypotheses #1, but Hypotheses #2 and #3 to also be tested. The ANOVA table appears in Table 16.

Table 16

Analysis of Variance Table

| Source | SS | df | MS | F-ratio | p |
|---|---|---|---|---|---|
| Group | 82.222 | 1 | 82.222 | 5.06 | 0.03 |
| Experience | 0.500 | 1 | 0.500 | 0.03 | 0.86 |
| Interaction | 32.191 | 1 | 32.191 | 1.98 | 0.16 |
| Error | 1511.069 | 93 | 16.248 | | |

The F-ratio obtained for the group factor was statistically significant at the $p < .03$ level. Consequently, the null hypothesis, that no difference existed between the groups, was not accepted. There appeared to be a statistically significant difference between the performance of those in the two comparison groups. However, these results must be interpreted cautiously. They only indicate that the difference between the groups was an unlikely chance occurrence assuming the null hypothesis to be true and given the sample size of 97.

Educational significance. The difference between the performance of the two groups was also evaluated in terms of educational significance. Using the formula in Table 4, a standardized mean difference (SMD) of +0.96 was obtained. That is, taking the pretest performance into consideration, the subjects in the LD.Trainer group scored on the average almost one standard deviation above the mean performance of the CLASS.LD2 group. Although there are no set standards against which to compare SMDs, a SMD of one-third to one-half a standard deviation in educational research is considered good (Joint Committee on Standards for Educational Evaluation, 1981). Therefore, the difference between the two groups' performance was substantial.

## Hypothesis 2: Between Experienced and Inexperienced Subjects

The second hypothesis read that there would be no statistically significant difference between the posttest scores of experienced and inexperienced teachers. Again, the descriptive information is presented first.

Descriptive data. The performance of the experienced and inexperienced teachers on the pretest and the posttest appears in Table 17. The experienced teachers scored higher on both the pretest and the posttest with the largest variance occurring in the experienced group on the posttest.

Table 17

Pretest and Posttest Scores by Experience Level

| | Experienced | Inexperienced | Total |
|---|---|---|---|
| Pretest | | | |
| *M* | 10.45 | 10.09 | 10.25 |
| *SD* | 3.25 | 3.61 | 3.44 |
| Variance | 10.56 | 13.03 | 11.83 |
| Range | 4 - 18 | 0 - 18 | 0 - 18 |
| Posttest | | | |
| *M* | 15.00 | 14.76 | 14.87 |
| *SD* | 5.08 | 3.78 | 4.37 |
| Variance | 25.81 | 14.29 | 19.10 |
| Range | 2 - 25 | 5 - 25 | 2 - 25 |

Statistical significance. The results of the ANOVA test (Table 16) were also used to test Hypothesis #2. The F-ratio obtained for the experience factor was very small (F-ratio = 0.03) and not statistically significant ($p < 0.86$). Therefore, the null hypothesis cannot be rejected and the difference obtained between experienced and inexperienced teachers on the posttest may be considered a likely chance occurrence assuming the null hypothesis to be true and given a sample size of 97.

Educational significance. Again, a SMD was computed to compare the difference between experienced and inexperienced teachers performance. The formula in Table 4 was computed by substituting the LD.Trainer group by the experienced teachers and the CLASS.LD2 group by the inexperienced teachers. This resulted in a SMD = - 0.04. This can be interpreted to mean that taking the pretest scores into consideration, the experienced teachers scored on the average one twentieth of a standard deviation below the mean performance of the inexperienced teachers. This difference is negligible.

## Hypothesis 3: Interaction

The third hypothesis was that there would be no statistically significant ($p < .05$) interaction between experience and training method. Again, the descriptive information is presented first.

Descriptive data. The descriptive data for experience and assigned group on the pretest and posttest are listed in Table 18. In addition, the mean posttest scores across experience and group are graphed in Figure 1. This graph demonstrates that the subjects in the LD.Trainer group who were experienced scored higher on the posttest ($\overline{X}$ = 17.27) than those who were inexperienced ($\overline{X}$ = 15.96). And the opposite effect occurred for the CLASS.LD2 group. The experienced teachers scored lower ($\overline{X}$ = 12.50) than those who were inexperienced ($\overline{X}$ = 13.52).

Statistical significance. The ANOVA test (Table 16) was used to assess whether the interaction observed between experience and group was statistically significant. The obtained F-ratio equalled 1.98. This was not statistically significant at the pre-specified .05 level ($p < 0.16$). Therefore, this result may be considered a likely chance occurrence given a sample size of 97 and assuming the null hypothesis to be true. The null hypothesis cannot be rejected.

### Hypotheses 4 and 5: Group Difference Between Experienced and Inexperienced

Hypothesis #4 was that there would be no statistically significant ($p < .05$) difference between the posttest performance of experienced teachers in the LD.Trainer and the CLASS.LD2 groups. The fifth hypothesis read that there would be no statistically significant ($p < .05$) difference

Table 18

Pretest and Posttest Scores by Experience Level and Group

| | Group | | Total |
|---|---|---|---|
| | LD.Trainer | CLASS.LD2 | |
| Pretest | | | |
| Experienced | | | |
| *M* | 11.09 | 9.75 | 10.45 |
| *SD* | 3.36 | 3.04 | 3.25 |
| Range | 6 - 18 | 4 - 16 | 4 - 18 |
| Inexperienced | | | |
| *M* | 9.68 | 10.52 | 10.09 |
| *SD* | 3.24 | 3.96 | 3.61 |
| Range | 0 - 16 | 0 - 18 | 0 - 18 |
| Posttest | | | |
| Experienced | | | |
| *M* | 17.27 | 12.50 | 15.00 |
| *SD* | 4.78 | 4.24 | 5.08 |
| Range | 10 - 25 | 2 - 21 | 2 - 25 |
| Inexperienced | | | |
| *M* | 15.96 | 13.52 | 14.76 |
| *SD* | 4.10 | 3.00 | 3.78 |
| Range | 10 - 25 | 5 - 19 | 5 - 25 |

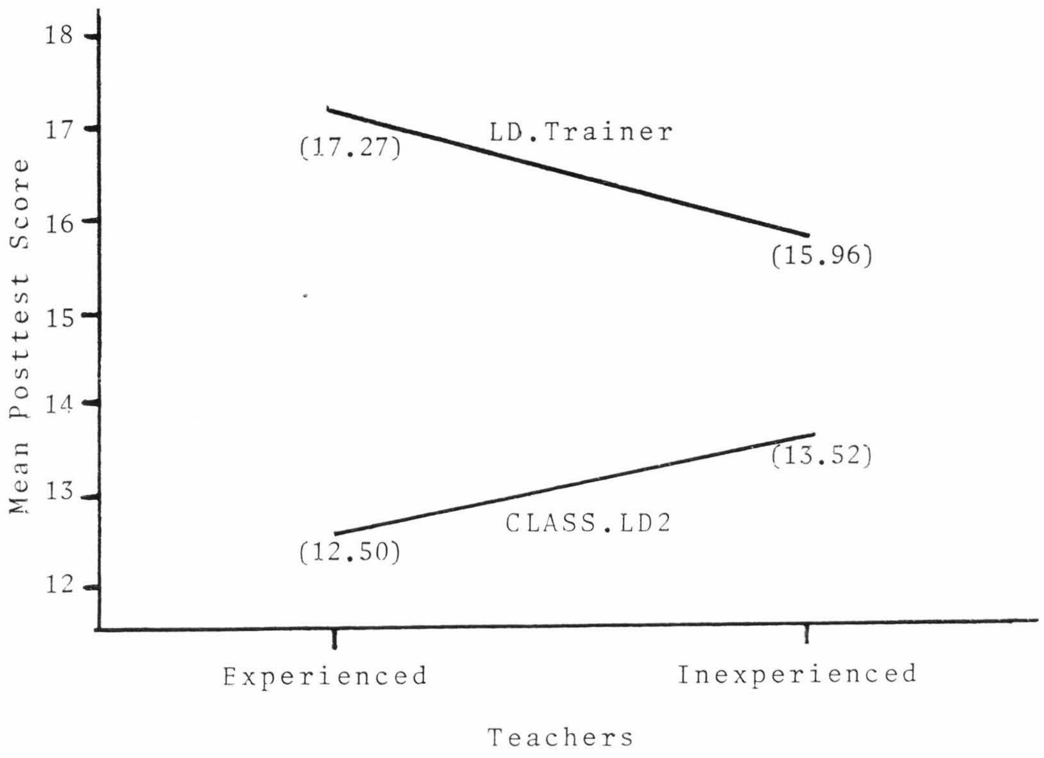


Figure 1. Interaction Between Experience Level and Group on Posttest Scores

between the posttest performance of inexperienced teachers in the LD.Trainer and the CLASS.LD2 group.

_Descriptive data_. The descriptive data appears in Table 18. Experienced teachers who completed the LD.Trainer materials scored higher ($\overline{X}$ = 17.27) than those who completed the CLASS.LD2 materials ($\overline{X}$ = 12.5Ø). As with the experienced teachers, the inexperienced subjects in the LD.Trainer group scored higher ($\overline{X}$ = 15.96) than those in the CLASS.LD2 group ($\overline{X}$ = 13.52).

_Statistical significance_. Planned orthogonal contrasts were used to determine whether the observed differences were statistically significant. In order to use planned orthogonal contrasts, each contrast must be orthogonal, or independent, to every other contrast. In addition, only k - 1 or less contrasts can be used (Glass & Hopkins, 1984). Both of these restrictions are met. The procedure used to test independence appears in Table 19.

The formulas and computations, as well as the critical value of t, appear in Table 2Ø. Both contrasts were statistically significant at the $p < .Ø5$ level. That is, for both the experienced and the inexperienced subjects, those in the LD.Trainer group scored statistically significantly higher than those in the CLASS.LD2 group. Therefore, the null hypothesis is rejected. This indicates hat these differences are an unlikely chance occurrence

Table 19

Independence of Contrasts

| Contrasts | | Constants Assigned to Means | | | |
|---|---|---|---|---|---|
| Number | Means | $\bar{X}_1$ | $\bar{X}_2$ | $\bar{X}_3$ | $\bar{X}_4$ |
| 1 | $\bar{X}_1 - \bar{X}_2$ | 1 | -1 | Ø | Ø |
| 2 | $\bar{X}_3 - \bar{X}_4$ | Ø | Ø | 1 | -1 |
| | Cross Products | Ø | Ø | Ø | Ø |

Note. $\bar{X}_1$ = Mean of the Experienced LD.Trainer Group

$\bar{X}_2$ = Mean of the Experienced CLASS.LD2

$\bar{X}_3$ = Mean of the Inexperienced LD.Trainer

$\bar{X}_4$ = Mean of the Inexperienced CLASS.LD

Table 20

Planned Orthogonal Contrast Results

| Contrasts | $\overline{X} - \overline{X}$ |
|---|---|
| Experienced | 17.27 - 12.50 = 4.77 |
| Inexperienced | 15.96 - 13.52 = 2.44 |
| **Standard Error of Contrast** | **(MS error) x (2/n)** |
| Experienced | (16.248) x (2/42) = 0.880 |
| Inexperienced | (16.248) x (2/55) = 0.769 |
| **Obtained t values** | **Contrast/Standard Error of Contrast** |
| Experienced | 4.77/0.880 = 5.42 |
| Inexperienced | 2.44/0.769 = 3.17 |
| Critical t value = 1.98, $p < .05$, df = 93* | |

*Note*. *For planned orthogonal contrasts the degrees of freedom for the critical t value are the degrees of freedom associated with the mean square error (Glass & Hopkins, 1984).

assuming the null hypothesis to be true and given the sample sizes of 42 and 55.

Educational significance. The effect size of mean differences was computed using the formula in Table 4. This resulted in standardized mean differences in the experienced subjects of + Ø.97 and + 0.96 for the inexperienced subjects. These results indicate that, taking into account pretest performances, those in the LD.Trainer group, regardless of being experienced or inexperienced, scored on the average almost one standard deviation above the mean performance of the corresponding group using the CLASS.LD2 materials. For educational research these differences are substantial (Joint Committee on Standards for Educational Evaluation, 1981).

## Rival Hypotheses

Subjects in the LD.Trainer group were assigned eight lessons and those in the CLASS.LD2 group were assigned to run consultations on sixteen files. In addition, they were told not to spend more than the time given them (four to four and one-half hours) completing the materials. If they had spent the allotted time working through the materials and had not completed the lessons or files assigned to them, they were to stop working and take the posttest.

Number of Items Completed

Table 21 lists the number of items completed by those in each group. Because subjects in the LD.Trainer group were assigned eight lessons and the CLASS.LD2 group sixteen files, the total number possible differed by group and direct comparison is not possible.

Table 21

Number of Lessons or Files Completed

| | Group | |
|---|---|---|
| Number of Items Completed | LD.Trainer* | CLASS.LD2** |
| *M* | 7.50 | 14.53 |
| *SD* | 1.22 | 2.53 |
| Range | 4 - 11*** | 7 - 16 |

*Note.* *Each subject in the LD.Trainer group was assigned eight (8) lesson. **Each subject in the CLASS.LD2 group was assigned 16 files. ***One subject completed 11 lessons within the time allotted.

Amount of Time Spent

A rival hypothesis to the effectiveness of the training may be that the subjects in the LD.Trainer group spent more time working on the materials. Consequently the self-reports on amount of time spent were analyzed. The descriptive data appears in Table 22. In addition, a t-test between the mean number of minutes spent was computed and indicated that this difference was not statistically significant.

Table 22

Total Amount of Time Spent in Minutes by Group

| | Group | | Total |
|---|---|---|---|
| Total Time Spent | LD.Trainer | CLASS.LD2 | |
| *M* | 244.42 | 242.19 | 243.34 |
| *SD* | 50.03 | 48.50 | 49.05 |
| Range | 112 - 352 | 158 - 396 | 112 - 396 |

t = .223

p = .824

Internal Validity

Campbell and Stanley (1963) list seven possible threats to the internal validity of any study. They include: maturation, regression, selection, mortality, instrumentation, testing, and history. These possible threats can only be threats to the internal validity if they affect one group to a larger degree than the other. Although Campbell and Stanley state that the use of random assignment and the pretest posttest design controls for all of the threats they identified, random assignment and the design only control for differences between the groups being attributable to chance. These seven threats to internal validity as they relate to this study are discussed.

Maturation does not appear to be a threat to this study because the subjects were adults and the independent variable lasted between one to six weeks. Regression was also not a threat because the subjects were not selected based on extreme scores. History probably was not a threat because the subjects completed the training independently and again, it lasted only a maximum of six weeks.

Selection as a possible threat can be assessed by examining the characteristics of the subjects in each group. Although there were some differences in the pretest scores (Table 13) and in the college majors (Table 10), the differences are negligible. Consequently, selection does not appear to be a threat.

Although three subjects registered for participation, were randomly assigned to the CLASS.LD2 group and then dropped out, none of them began the training. Therefore, mortality does not appear to be a threat. Possible instrumentation problems were controlled by using three readers. Because the scores of two out of three were used in the final analysis, possible experimenter bias or drift was controlled. Testing effects could have created a threat to internal validity because both a pretest and a posttest were used. But because both groups were assessed similarly, testing effects should have not differed across groups.

### External Validity

The results of this study have limited external validity because (a) only volunteers were used as subjects; and (b) the subjects were not randomly selected from the target population, but were, rather, the same group.

## Summary

The subjects for this study consisted of 97 volunteer students from three universities: Utah State University, St. Cloud State University (Minnesota), and The University of South Dakota. Forty-two of the subjects were experienced teachers and 55 were not. Fifty subjects were randomly assigned to the LD.Trainer group and 47 to the CLASS.LD2 group.

The results of this study indicated that the subjects in the LD.Trainer group scored statistically and educationally significantly better than those in the CLASS.LD2 trainer group. This applied to all subjects, as well as the experienced group and the inexperienced group alone. There was, however, no statistically significant difference between the posttest performance of the experienced and the inexperienced subjects. And although an interaction between group and experience was observed, it was not statistically significant.

Replication of the study is important to accomplish in order to ascertain whether the results obtained may be attributed to true differences, rather than chance occurrences. Replication would also increase the generalizability of the results.

# CONCLUSIONS AND RECOMMENDATIONS

## Conclusions

Subjects who interacted with the LD.Trainer materials scored statistically and educationally higher on the posttest than those who ran CLASS.LD2 consultations. These results were obtained across all subjects and across the experienced and inexperienced subjects when considered alone. On the average, subjects in both groups scored higher on the posttest than on the pretest. In addition, an interaction, although not statistically significant ($p < .05$), was obtained between group and experience.

The LD.Trainer materials were designed to incorporate expert system technology and effective concept instruction. Although there exist many similarities between the processes of knowledge engineering and concept analysis, incorporating both to develop an effective training tool had not previously been demonstrated. Results of this study indicated that the two fields, successfully combined, can create an effective and efficient training tool.

The following conclusions may be drawn from the results of this study:

1. Expert system technology and effective concept instruction can be combined to create an effective and efficient training tool.

2. Based on performance on a paper/pencil test, learners completing the training materials and those who simply use the original system learn at least some of the content of the knowledge contained in the system.
3. Learners completing the training materials learn and can accurately apply that knowledge on a paper/pencil test to a greater degree than those who simply use the original system. This applies to both experienced and inexperienced teachers.
4. The training materials appeared more effective (based on a paper/pencil test) with experienced teachers than with inexperienced teachers.

Each of these conclusions is discussed in more detail.

## Demonstrated Model

The results of this study demonstrate that expert system technology and effective concept instruction can be combined to create a cost-effective and efficient training tool in which the knowledge contained in the expert system is taught. If the expert system already exists, creating a training tool using printed materials and minor modifications in the system become relatively simpler, easier, and more cost-effective than the addition of a sophisticated front-end tutor (like NEOMYCIN). Also, the approach used in this study is more instructionally sound. In particular, the examples and nonexamples used in

LD.Trainer, unlike NEOMYCIN, are carefully selected and sequenced and are accompanied with explicit definitions.

Future expert systems could be developed, not as field-consultants, but primarily for training purposes. In fact, because of the similarities between knowledge engineering and concept analysis, the development of the system and the concept training materials could go hand-in-hand.

Although the combination of effective concept instruction and expert system technology has been demonstrated the model used in this study is somewhat dependent upon the context and the structure of the expert system. For example, the knowledge contained in CLASS.LD2 was easily broken into three distinct lessons. In evaluating another expert system--namely, CLASS.BD (Ferrara, Serna, & Baer, 1986), which provides second-opinion advice regarding behaviorally disordered (BD) classifications, the factors used in making BD decisions are not as distinct to one another as the learning disabilities factors and must be taught as a whole. Consequently, the specific model used to devise LD.Trainer cannot be used in the development of BD.Trainer. Some of the model, however, will remain the same, such as the presentation of definitions and examples/nonexamples.

## Both Approaches Teach Some Content

Learners in both the LD.Trainer and the CLASS.LD2 group on the average gained knowledge as assessed through performance on the pretest and posttest. Through informal analysis of the posttest responses, it appeared that those in the CLASS.LD2 group learned primarily the "automatic disclaimers" to a learning disabilities classification. For example, CLASS.LD2 was designed such that if a user responded 'yes' to a question addressing whether the student's primary problem is behavioral, the system automatically stopped the consultation and gave the advice that a student whose primary problem is behavioral, cannot be learning disabled. The same approach was used with physical problems. These two concepts were generally learned by both groups. However, when a disclaimer was concluded by the system, rather than by one user response, those in the LD.Trainer group learned the concept much better than those in the CLASS.LD2 group. For example, in dealing with the cultural disadvantaged disclaimer, CLASS.LD2 was designed to address several subissues relating to the student's cultural background such as bilingualism, proficiency of English, number of years living in the United States, and percent of minorities in the student's school. Although the user may have responded 'yes' to the question addressing possible cultural problems, the system was designed to analyze user responses to these additional questions to conclude whether the cultural problem precluded an LD classification. Subjects in the LD.Trainer group learned to distinguish cultural differences and those in the CLASS.LD2 group did not.

## LD.Trainer More Effective

The application of concept instruction and an existing expert system to the development of training materials is more effective when compared with simple use of the original expert system. This applies to both novices and to those with experiences in the field. If the goal is to train students to replicate the decision-making processes of an existing expert system, the results of this study would suggest that the development of training materials is more effective and probably worth the development costs than simple exposure to the system.

## LD.Trainer More Effective with Experienced Teachers

In this study, experienced teachers scored higher than inexperienced teachers on both the pretest and the posttest. However, experienced teachers in the LD.Trainer scored statistically higher on the posttest than experienced teachers in the CLASS.LD2 group. This is an opposite effect than originally expected.

It was expected that those with prior knowledge and experience would be able to learn from simple exposure and use of the system. However, this was not the case; in fact, the opposite occurred. Experienced teachers may have began the training with more misconceptions than inexperienced teachers, who merely lacked knowledge. This was, in fact, an observation of those correcting the pretests.

Experienced regular educators, in particular, confused learning disabled with mentally retarded students. The LD.Trainer materials were designed to clarify misconceptions and the CLASS.LD2 materials were not.

## Recommendations

The combination of expert system technology and effective concept instruction has been demonstrated to be an efficient and effective means of training. However, additional expert system development and research is necessary in order to draw conclusions beyond the results of this study.

### Development of Expert System Training Tools

Additional expert systems should be developed across content areas, using different instructional components, and using different ICAI approaches.

Across content areas. LD.Trainer was designed to train teachers to accurately classify learning disabled students. Expert systems in other content areas should be modified as training tools using a process similar to the design of LD.Trainer. The effectiveness of these systems should then be studied.

Use of different instructional components. Although the design of LD.Trainer incorporated effective concept instruction variables, some of the design variables have not been adequately demonstrated. Similar to the Minnesota Adaptive Instructional System (MAIS), which is being used to study instructional variables (Tennyson, in press), future expert system training tools could incorporate different design variables. This could enhance knowledge in instructional theory. For example, whether learner sophistication and task complexity interacts with the need to minimize variation in the irrelevant attributes of matched examples and nonexamples remains unanswered. Also, under what conditions concepts can be taught using examples alone has yet to be adequately demonstrated.

Modifications based on misconceptions. LD.Trainer was designed incorporating validated design strategies for concept instruction. The same pattern of procedures were used for each lesson. This resulted in each sublesson receiving approximately equal emphasis. The effectiveness of LD.Trainer could be enhanced by drawing from the pretraining misconcepts and restructuring the emphasis within the content accordingly. As mentioned earlier, pretest readers indicated that many educators had learning disabilities and mental retardation confused. Within the LD.Trainer materials, therefore, the lessons which differentiate between these two handicapping conditions could be designed to more directly distinguish the two.

Other ICAI comparison. In order to assess whether the model suggested by this study is as or more effective than other intelligent computer-assisted instruction requires the methods to be examined simultaneously. This would require the development of two training systems using the same expert system.

For example, an existing expert system could undergo two different modifications, one similar to that demonstrated in this study and the other similar to that used to create NEOMYCIN (Davis et al., 1975). Then development of the two methods in terms of time, cost, and personnel, could be compared. The two sets of training materials could also be compared in terms of effectiveness.

### Additional Studies

Additional studies need to be conducted using LD.Trainer and other future expert system-based training tools. Several possible modifications of this study are suggested.

Replication of the study. In order to ensure that the differences observed between the effectiveness of the LD.Trainer and the CLASS.LD2 materials was not due to rival hypotheses, replication of this study is needed. Additional research using experienced and inexperienced special and regular educators, should be conducted. One question which remains unanswered is whether the differences between the materials have any differential effect between regular and special educators. Replication would also improve the external validity of the present study.

Materials alone. Although the LD.Trainer system incorporated both computers and printed materials, it may be possible that the printed materials could stand alone and be effective tools for training.

It was the developer's experience that the existence of the expert system facilitated development of the LD.Trainer printed materials. In particular, the system's knowledge base contained rules which were well outlined. These rules helped the author create definitions; and once examples and nonexamples were generated, the availability of the system allowed them to be checked for accuracy. Consequently, the printed materials alone contained the most salient information from CLASS.LD2 and could perhaps be an effective training mechanism alone.

Different dependent variables. Replication similar to this study should be conducted using different dependent variables. Of particular interest would be dependent variables of real-life decisions. That is, do the learning disabilities decisions made by trainees prior to and following training differ?

Many factors are operating in the public school system which influence decisions regarding learning disabilities classification. It would be of interest to examine the role effective instruction can play in influencing these decisions. That is, if school personnel could be better trained to more accurately identify learning disabled students, perhaps more accurate decisions could be made in the future.

REFERENCES

Alessi, S. M., & Trollip, S. R. (1985). Computer-based instruction: Methods and development. Englewood Cliffs, NJ: Prentice-Hall.

Anderson, R. C., & Kulhavy, R. W. (1972). Learning concepts from definitions. American Educational Research Journal, 9, 385-391.

Atkinson, R. C. (1975). Mnemotechnics in second-language learning. American Psychologist, 30, 821-828.

Barr, A., & Feigenbaum, E. A. (Eds.). (1981). The handbook of artificial intelligence (1st ed.). Los Altos, CA: William Kaufman, Inc.

Borg, W., Kallenback, W., Kelly, M., Langer, P., & Gall, M. (1970). The minicourse: A micro-teaching approach to teacher education. Beverly Hills, CA: Macmillan Educational Services.

Bourne, L. E., Jr., & O'Banion, K. (1971). Conceptual rule learning and chronological age. Developmental Psychology, 5, 525-534.

Bracht, G. H., & Glass, G. V. (1968). The external validity of experiments. American Educational Research Journal, 5, 437-474.

Breuer, K., & Hajovy, H. (in press). Adaptive instructional simulations to improve learning of cognitive strategies. Educational Technology.

Burke, R. L. (1982). CAI sourcebook. Englewood Cliffs, NJ: Prentice-Hall.

Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research. Chicago: Rand McNally.

Carnegie Council on Policy Studies in Higher Education (1979). Fair practices in higher education: Rights and responsibilities of students and their colleges in a period of intensified competition for enrollments. San Francisco: Jossey-Bass.

Carnine, D. W. (1976). Establishing a discriminative stimulus by distributing attributes of compound stimuli between positive and negative instances. Eugene, OR: University of Oregon Follow Through Project.

Carnine, D. W. (1980a). Relationships between stimulus variation and the formation of misconceptions. Journal of Educational Psychology, 74, 106-110.

Carnine, D. W. (1980b). Three procedures for presenting minimally different positive and negative instances. Journal of Educational Psychology, 72, 452-456.

Carroll, J. B. (1964). Words, meanings, and concepts. Harvard Educational Review, 34, 178-202.

Cartwright, D. D., & Neikkinen, M. W. (1981, March). Developing conservation attitudes and energy concepts in individuals of various cognitive levels, using the Energy Environment Simulator. Paper presented to the Annual Meeting of the Northwest Scientific Association, Corvallis, OR. (ERIC Document Reproduction Service No. ED 201 486)

Champagne, A. B., Klopfer, L. E., Desena, A. T., & Squires, D. A. (1981). Structural representations of students' knowledge before and after science instruction. Journal of Research in Science Teaching, 18, 97-111.

Clancey, W. J., & Letsinger, R. (1981). NEOMYCIN: Reconfiguring a rule-based expert system for application to teaching. Proceedings of the seventh international joint conference on artificial intelligence. Vancouver, British Columbia, Canada: University of British Columbia.

Clancey, W. J., Shortliffe, E. H., & Buchanan, B. G. (1979). Intelligent computer-aided instruction for medical diagnosis. In W. J. Clancey & E. H. Shortliffe (Eds.), Readings in medical artificial intelligence (pp. 256-274). Reading, MA: Addison-Wesley.

Clark, R. E. (1983). Reconsidering research on learning from media. Review of Educational Research, 53, 445-459.

Code of Federal Regulations. (1980). Assistance to states for education of handicapped children. Title 45 - Public Welfare. Washington D. C.: The Office of the Federal Register, 332-388.

Davis, R., Buchanan, B., & Shortliffe, E. (1975). Production rules as a representation for a knowledge-base consultant program. (Report STAN-CS-75--519 Memo AIM-226). Palo Alto, CA: Stanford University, Computer Science Department.

Driscoll, M. P., & Tessmer, M. (1985a). Applications of the concept tree and rational set generator for coordinate concept learning. Florida State University. (ERIC Document Reproductive Service No. ED 257 875)

Driscoll, M. P., & Tessmer, M. (1985b). The rational set generator: A method for creating concept examples for teaching and testing. Educational Technology, 25(2), 29-32.

Ebel, R. L. (1979). Essentials of educational measurement. Englewood Cliffs, NJ: Prentice-Hall.

Eggen, P. D., Kauchak, D. P., & Harder, R. J. (1979). Strategies for teachers: Information processing models in the classroom. Englewood Cliffs, NJ: Prentice-Hall.

Englemann, S. (1969). Conceptual learning. Belmone, CA: Fearon.

Englemann, S., & Carnine, D. (1982). Theory of instruction: Principles and applications. New York: Irvington.

Feldman, K. V., & Klausmeier, H. J. (1974). Effects of two kinds of definition on the concept attainment of fourth and eighth graders. Journal of Educational Research, 67, 219-223.

Ferguson, G. A. (1981). Statistical analysis in psychology and education (5th ed.). New York: McGraw-Hill.

Ferrara, J. M., & Hofmeister, A. M. (1984). CLASS.LD2: An expert system for classifying learning disabilities [Computer program]. Logan, UT: Developmental Center for Handicapped Persons, Utah State University.

Ferrara, J. M., Parry, J. D., & Lubke, M. M. (1985). CLASS.LD2: An expert system for student classification Logan: Utah State University, Artificial Intelligence Research and Development Unit, Developmental Center for Handicapped Persons. [ERIC Document Reproduction Service No. ED 263 734]

Ferrara, J. M., Prater, M. A., & Baer, R. (in press). LD.Trainer: Modification of an expert system for complex conceptual training. Educational Technology.

Ferrara, J., Serna, R. W., & Baer, R. (1986). Behavior Consultant [Computer Program]. Logan, UT: Developmental Center for Handicapped Persons, Utah State University.

Flake, J. L. (1975). Interactive computer simulations for teacher education. Educational Technology, 15(3), 54-57.

Frayer, D. A. (1970). Effects of number of instances and emphasis of relevant attribute values on mastery of geometric concepts by fourth- and sixth-grade children. (Technical Report No. 116). Madison, WI: Wisconsin Research and Development Center for Cognitive Learning.

Fredrick, W. C., & Klausmeier, H. J. (1968). Instructions and labels in a concept-attainment task. Psychological Reports, 23, 1339-1342.

Gagne, R. M. (1985). The conditions of learning (4th ed.). New York: CBS College Publishing.

Gagne, R. M., & Briggs, L. J. (1979). Principles of instructional design (2nd ed.). New York: Holt, Rinehart and Winston.

Glass, G. V., & Hopkins, K. D. (1984). Statistical methods in education and psychology (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall, Inc.

Glass, G. V., Peckham, P. D., & Sanders, J. R. (1972). Consequences of failure to meet assumptions underlying the fixed effects analysis of variance and covariance. Review of Educational Research, 42, 237-288.

Goodall, A. (1985). The guide to expert systems. Oxford, NJ: Learned Information.

Harmon, P., & King, D. (1985). Expert systems: Artifical intelligence in business. New York: John Wiley and Sons.

Harris, R. C. (1973). Concept learning as a function of type, identifiability, and variety of instructional instances. The Journal of Educational Research, 67, 182-189.

Hayes-Roth, F., Waterman, D. A., & Lenat, D. B. (1983). An overview of expert systems. In F. Hayes-Roth, D. A. Waterman, D. B. Lenat (Eds.), Building expert systems (pp. 3-30). Reading, MA: Addison-Wesley.

Haynes, J. A., Pilato, V. H., & Malouf, D. B. (in press). Expert systems for educational decision making. Educational Technology.

Hofmeister, A. M. (1977). Educating the mildly handicapped. Unpublished manuscript, Department of Special Education, Utah State University, Logan, UT.

Hofmeister, A. M. (March 21, 1983). Comments on Part B, Section 618: Evaluation and Part E, research and education of the handicapped, Public Law 94-142. Testimony before the Senate Committee on Labor and Human Resources, Subcommittee on the Handicapped, Congressional Record. Washington, D.C.: Government Printing Office.

Hofmeister, A. M. (1986). What is an expert system? Technical Paper, Artificial Intelligence Research and Development Unit, Developmental Center for Handicapped Persons, Utah State University, Logan, UT.

Hofmeister, A. M. (in press). Formative evaluation in the development and validation of expert systems in education. Computational Intelligence.

Hofmeister, A. M., & Ferrara, J. M. (1986). Expert systems and special education. Exceptional Children, 53, 235-239.

Hollan, J. D., Hutchins, E. L., & Weitzman, L. (1984). STEAMER: An interactive inspectable simulation-based training system. The AI Magazine, 5(2), 15-27.

Joint Committee on Standards for Educational Evaluation (1981). Standards for evaluations of educational programs, projects, and materials. New York: McGraw-Hill.

Keogh, B. K. (1983). Classification, compliance, and confusion. Annual Review of Learning Disabilities, 1, 22.

Kidd, A. (1984). Human factors problems in the design and use of expert systems. In A. Monk (Ed.) Fundamentals of human-computer interaction (pp. 237-247), Orlando, FL: Academic Press.

Kinnear, J. F. (1986, April). Computer simulation and problem solving in genetics. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco.

Klausmeier, H. J. (1980). Learning and teaching concepts: A strategy for testing applications of theory. New York: Academic Press.

Klausmeier, H. J., Ghatala, E. S., & Frayer, D. A. (1974). Conceptual learning and development: A cognitive view. New York: Academic Press.

Markle, M. M. (1975). They teach concepts don't they. Educational Researcher, 4(6), 3-9.

Markle, S. M. (1977). Teaching conceptual networks. Journal of Instructional Development, 1, 13-17.

Markle, S. M., & Tiemann, P. W. (1970). Really understanding concepts: Or in frumious pursuit of the jabberwock. Champaign, IL: Stipes.

Martindale, E. S., Ferrara, J. M., & Campbell, B. W. (1986). Accuracy of CLASS.LD2: An expert system for classifying learning disabled students. Technical Report, Logan, UT: Artificial Intelligence Research and Development Unit, Developmental Center for Handicapped Persons, Utah State University.

Martorella, P. H. (1972). Concept learning: Designs for instruction. Scranton, OH: Intext.

McKinney, J. D. (1984). The search for subtypes of specific learning disability. Annual Review of Learning Disabilities, 2, 19-26.

Merrill, M. D. (1983). Component display theory. In C. M. Reigeluth (Ed.) Instructional-design theories and models (pp. 279-333). Hillsdale, NJ: Lawrence Erlbaum.

Merrill, M. D., Reigeluth, C. M., & Faust, G. W. (1979). The instructional quality profile: A curriculum evaluation and design tool. In H. F. O'Neil, Jr. Procedures for instructional systems development (pp. 165-204). New York: Academic Press.

Merrill, M. D., & Tennyson, R. D. (1977). Teaching concepts: An instruction design guide. Englewood Cliffs, NJ: Educational Technology Publications.

O'Shea, T., & Self, J. (1983). Learning and teaching with computers: Artificial intelligence in education. Englewood Cliffs, NJ: Prentice-Hall.

Park, O. (1981). A response-sensitive strategy in computer-based instruction. A strategy for concept teaching. Journal of Educational Technology Systems, 10, 187-197.

Park, O. (1984). Empirically based procedures for designing a response-sensitive sequence in computer-based instruction: An example from concept teaching strategies. Journal of Computer Based Instruction, 11, 14-18.

Parry, J. D. (1986a). Mandate consultant. [Computer program]. Logan, UT: Developmental Center for Handicapped Persons, Utah State University.

Parry, J. D. (1986b). The development and validation of an expert system for reviewing a special education practice. Unpublished doctoral dissertation, Utah State University, Logan, UT.

Prater, M. A., & Ferrara, J. M. (1986, May). Training applications of expert systems. Paper presented at the Annual Meeting of the Association for Behavior Analysis, Milwaukee, WI.

Prater, M. A., & Lubke, M. (1986). Expert system training applications. Technical Paper, Logan, UT: Artificial Intelligence Research and Development Unit, Developmental Center for Handicapped Persons, Utah State University.

Priest, A. G. (1981). Artificial intelligence and learning: Conference reports. Instructional Science, 10, 277-285.

Ragan, S. W., & McFarland, T. D. (in press). Applications of expert system in education: A technology for decision-makers. Educational Technology.

Roberts, F. C., & Park, O. (1983). Intelligent computer-assisted instruction: An explanation and overview. Educational Technology, 23, 7-11.

Sabatino, D. A. (1983). "The house that Jack built." Annual Review of Learning Disabilities, 1, 23-24.

Serna, R. W. (1986, May). An expert system for behavior analysis. Paper presented at the Annual Meeting of the Association for Behavior Analysis, Milwaukee, WI.

Shaver, J. P. (1983). The verification of independent variables in teaching methods research. Educational Researcher, 12(8), 3-9.

Shaver, J. P. (1985, October). Chance and nonsense: A conversation about interpreting tests of statistical significance, Part 2. Phi Delta Kappan, pp. 138-141.

Shumway, R. J., White, A. L., Wilson, P., & Brombacher, B. (1983). Feature frequency and negative instances in concept learning. American Educational Research Journal, 20, 451-459.

Sleeman, D., & Brown, J. S. (Eds.) (1982). Intelligent tutoring systems. London: Academic Press.

Smith, D. K. (1978). Faculty vitality and the management of university personnel policies. In W. R. Kirschling (Ed.), Evaluating faculty performance and vitality (pp. 1-16). San Francisco: Jossey-Bass.

Sowizral, H. A., & Kipps, J. R. (1986). Artificial intelligence: A Rand perspective. In P. Klahr & D. A. Waterman (Eds.), Expert systems: Techniques, tools, and applications (pp. 27-69). Reading, MA: Addison-Wesley.

Stefik, M., Aikins, J., Balzer, R., Benoit, J., Birnbaum, L, Hayes-Roth, F., & Sacerdoti, E. (1983). The architecture of expert systems. In F. Hayes-Roth, D. A. Waterman, & D. B. Lenat (Eds.), Building expert systems (pp. 89-126). Reading, MA: Addison-Wesley.

Steinberg, E. R., Baskin, A. B., & Hofer, E. (1986, April). Computer simulations and the acquisition of problem solving skills. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco.

Stockburger, D. W. (1980, April). Teaching statistical concepts using microcomputer simulations. Paper presented at the meeting of the Southwestern Psychological Association, Oklahoma City. (ERIC Document Reproduction Service No. ED 190 111).

Suppes, P. (1979). Current trends in computer-assisted instruction. In M. L. Yovits (Ed.) Advances in computers, Vol. 18 (pp. 173-229). New York: Academic Press.

Teknowledge, Inc. (1986). M.1 [Computer Program]. Palo Alto, CA: Teknowledge, Inc.

Tennyson, R. D. (1980). Instructional control strategies and content structure as design variables in concept acquisition using computer-based instruction. Journal of Educational Psychology, 72, 525-532.

Tennyson, R. D. (1981). Use of adaptive information for advisement in learning concepts and rules using computer-assisted instruction. American Educational Research Journal, 18, 425-438.

Tennyson, R. D. (1986). Application of artificial intelligence methods to computer-based instructional design: The Minnesota Adaptive Instructional System. Journal of Instructional Development, 7(3), 17-22.

Tennyson, R. D. (in press). MAIS: An educational alternative of ICAI. Educational Technology.

Tennyson, R. D., & Boutwell, R. C. (1974). Methodology for the sequencing of instances in classroom concept teaching. Educational Technology, 14(9), 45-49.

Tennyson, R. D., & Buttrey, T. (1980). Advisement and management strategies as design variables in computer-assisted instruction. ECTJ - Educational Communication and Technology Journal, 28(3), 169-176.

Tennyson, R. D., & Cocchiarella, M. J. (1986). An empirically based instructional design theory for teaching concepts. Review of Educational Research, 56, 40-71.

Tennyson, R. D., & Park, O. (1980). The teaching of concepts: A review of instructional design research literature. Review of Educational Research, 50, 55-70.

Tennyson, R. D., & Park, O. (1984). Computer-based adaptive instructional systems: A review of empirically based models. Machine-Mediated Learning, 1, 129-153.

Tennyson, R. D., & Rothen, W. (1977). Pretask and on-task adaptive design strategies for selecting number of instances in concept acquisition. Journal of Psychology, 69, 586-592.

Tennyson, R. D., Woolley, F. R., & Merrill, M. D. (1972). Examplar and nonexamplar variables which produce correct concept classification behavior and specified classification errors. Journal of Educational Psychology, 63, 144-152.

Tennyson, R. D., Youngers, J., & Suebsonthi, P. (1983). Concept learning by children using instructional presentation forms for prototype formation and classification-skill development. Journal of Educational Psychology, 75, 281-291.

Thompson, B. A., & Thompson, W. A. (1985, April). Inside an expert system. BYTE Magazine, p. 315-330.

Thorkildsen, R. J., Lubke, M. M., Myette, B. M., & Parry, J. D. (1985). Artificial intelligence: Applications in education. Educational Research Quarterly, 10, 2-9.

Torgeson, J., & Kail, R. V., Jr. (1980). Memory processes in exceptional children. In B. Keogh (Ed.), Advances in special education, Vol. 1 (pp. 55-99). Greenwich, CN: JAI Press.

Utah State Board of Education (1981). Rules and regulations for education programs for the handicapped. Salt Lake City: Author.

Wang, M. C., Vaughan, E. D., & Dytman, J. A. (1985). Staff development: A key ingredient of effective mainstreaming. Teaching Exceptional Children, 17(2), 112-121.

Waterman, D. A., & Jenkins, B. M. (1986). Developing expert systems to combat international terrorism. In P. Klahr & D. A. Waterman (Eds.), Expert systems: Techniques, tools, and applications (pp. 95-134). Reading, MA: Addison-Wesley.

Waterman, D. A., & Peterson, M. A. (1986). Models of legal decision-making. In P. Klahr & D. A. Waterman (Eds.), Expert systems: Techniques, tools, and applications (pp. 95-134). Reading, MA: Addison-Wesley.

Weinberg, G. H., Schumaker, J. A., & Oltman, D. (1981). Statistics: An intuitive approach (4th ed.). Belmont, CA: Brooks/Cole.

Wenger, E. (1985). _AI and the communication of knowledge: An overview of intelligent teaching systems._ Irvine: University of California, Irvine Computational Intelligence Project, Department of Information and Computer Science.

Williams, P. B., & Carnine, D. W. (1981). Relationship between range of examples and of instructions and attention in concept attainment. _Journal of Educational Research_, _74_, 144-148.

Wolfe, R. O., & Macauley, H. K. (1975). Simulation: A multidimensional approach to teacher education. _Educational Technology_, _15_(1), 45-50.

Woolley, F. R., & Tennyson, R. D. (1972). Conceptual model of classification behavior. _Educational Technology_, _12_(4), 37-39.

Yazdani, M., & Lawler, R. W. (1986). Artificial intelligence and education: An overview. _Instructional Science_, _14_, 197-206.

Ysseldyke, J. E. (1983). Current practices in making psychoeducational decisions about learning disabled students. _Annual Review of Learning Disabilities_, _1_, 31-38.

Ysseldyke, J. E., Algozzine, B., Regan, R., & Potter, M. (1980). Technical adequacy of tests used by professionals in simulated decision making. _Psychology in the Schools_, _17_, 202-209.

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

Ysseldyke, J. E., Thurlow, M., Graden, J., Wesson, C., Algozzine, B., & Deno, S. (1983). Generalizations from five years of research on assessment and decision making: The University of Minnesota Institute. Exceptional Education Quarterly, 4, 75-93.

# APPENDICES

# Appendix A

## A Comparison of the LD.Trainer and CLass.LD2 Training Materials

LD.TRAINER AND CLASS.LD2 TRAINING MATERIALS COMPARED

| Events of Instruction* | LD.Trainer | CLASS.LD2 |
|---|---|---|
| Taking account of prior knowledge | Based on performance on the pretest, students were assigned to lessons. | None |
| Gaining and maintaining attention | Gaining attention was accomplished through verbal instructions. No set strategies were employed for maintaining attention. | Gaining attention was accomplished through verbal instructions. No set strategies were employed for maintaining attention. |
| Informing students about objectives | Students were verbally given an overall objective. In addition, each lesson began with written objectives. | Students were verbally given an overall objective. |
| Stimulating recall of prerequisites | None | None |
| Presenting stimulus, eliciting response, and providing feedback (materials and instructor) | The materials were designed such that examples and nonexamples were presented to the students (stimulus), responses were required, and feedback was given. The instructor was available to monitor and answer questions. | The stimulus, response, feedback features were only those available to the students through the expert system. The instructor was available to monitor and answer questions. |

* Adapted from Gagne' and Briggs (1979)

| | | |
|---|---|---|
| Providing learning guidance | Learning guidance was provided through attribute identification and appropriate sequencing of examples. | None |
| Allowing for individual differences | The materials were designed to be used individually thus allowing for students to move at their own pace. | The materials were designed to be used individually thus allowing for students to move at their own pace. |
| Assessing the performance | Student performance was assessed through the posttest. | Student performance was assessed through the posttest. |
| Enhancing retention and transfer | Use of effective concept instruction strategies in the design of the materials facilitated retention and transfer. | None |

# Appendix B

## A Sample Lesson

## from LD.Trainer

# LESSON 1.A

## DISCREPANCY AREA AND ACTUAL ACHIEVEMENT DISCREPANCY

Objective: After completing Lesson 1.A you will be able to:

(a) determine whether a student qualifies as learning disabled under the discrepancy area and actual achievement discrepancy requirements,

(b) quantify how certain you are that the student is learning disabled, and

(c) justify your responses.

Needed Information: (1) Area of deficiency

(2) Grade placement

(3) Test score

LESSON 1.A

INSTRUCTION

DEFINITION: A learning disabled student must score **at least 40% below grade placement** in at least one of the following academic areas:

**basic reading skill**

**reading comprehension**

**calculation**

**mathematical reasoning**

**written expression**

**listening comprehension**

**oral language.**

In order to determine whether the student is 40% below, the following formula is calculated:

**[(Grade Placement - Test Score)/Grade Placement] x 100 = Percent Behind**

EXAMPLE:

Marsha scored at the **4.2** level (fourth grade, second month) on a **written expression** test while in the **8.7** grade (eighth grade, seventh month). **[(8.7 - 4.2)/8.7] x 100 = 51.7%**

Marsha could qualify as a learning disabled student because she scored more than 40% (51.7%) below her grade placement in an appropriate academic area.

NONEXAMPLE:

Rob is also in the eighth grade, seventh month **(8.7).** On the **written expression** test he scored at the sixth grade, third month **(6.3).** **[8.7 -6.3)/8.7] x 100 = 27.6%**

Rob does not qualify as a learning disabled student because he scored less than 40% (27.6%) below his grade placement although it was in an appropriate academic area.

COMPUTER QUESTIONS

The remainder of the examples and nonexamples for Lesson 1.a will involve use of the computer. The first question you will see is:

---

In which of the area(s) listed below is (are) the child's learning deficit(s)? If the child has problems in more than one area list them all separating each area with a comma (eg. l,r, m,c).

listening comprehension
oral expression
written expression
basic reading skills
reading comprehension
mathematics reasoning
calculation
none of the above

---

The next question you will see is:

---

At the time testing was completed, what was the child's grade placement? eg. 2.0; 3.1; 9.4

(Note: A kindergartener on the first day of school would be entered as 0.0).

---

The third question will be:

---

What was the student's grade level score in the area of _____________? eg. 2.0; 3.1; 9.4

---

INSTANCE 1

SHELLY, WHILE IN THE ELEVENTH GRADE (11.Ø), SCORED AT THE FIFTH GRADE, FIFTH MONTH (5.5) LEVEL IN MATHEMATICAL CALCULATION.

Shelly's learning deficit is in calculation. So in response to the first question type 'c' for calculation. Shelly was tested while in 11.Ø, so type '11.Ø' in response to the second question.

You are given her test score as 5.5. Enter that in response to the third and last question.

| Student | Deficiency Area | Grade Placement | Test Score |
|---|---|---|---|
| Shelly | Calculation | 11.Ø | 5.5 |

You should now obtain the advice regarding Shelly. It should read:

---

Learning disabled:

> If the information which you have provided is correct, the student may be classified as learning disabled at the confidence level suggested below.
> advice shown = learning disabled (95%) because
> rule-625

---

The advice shown indicates that Shelly can be classified as learning disabled with 95% confidence. A 5% confidence leeway is allowed because no one is ever perfectly confident about a learning disabilities classification. Shelly's actual discrepancy is 5Ø%, [(11.Ø - 4.4)/11.Ø] x 1ØØ = 5Ø%. That is, she scored 5Ø% below her grade placement in mathematical calculation, an appropriate academic area.

Type 'done' which will return you to the main menu.

| Student | Deficiency Area | Grade Placement | Test Score | Advice and Confidence |
|---|---|---|---|---|
| Shelly | Calculation | 11.Ø | 5.5 | LD 95% |

INSTANCE 2

MATT WAS IN ELEVENTH GRADE (11.Ø) WHEN HE SCORED AT THE 9.5 GRADE LEVEL IN CALCULATION.

Both Shelly (the previous instance) and Matt were in the eleventh grade (11.Ø) when tested in the area of mathematical calculation. Shelly scored at the 5.5 grade level and Matt scored at the 9.5 grade level. Shelly qualified as learning disabled. In order to determine if Matt qualifies enter, 'c', '11.Ø', and '9.5' to the three questions, respectively.

| Student | Deficiency Area | Grade Placement | Test Score | Advice and Confidence |
|---|---|---|---|---|
| Shelly | Calculation | 11.Ø | 5.5 | LD 95% |
| Matt | Calculation | 11.Ø | 9.5 | |

Although Matt's learning deficiency was in an appropriate academic area (i.e. mathematical computation), Matt cannot be classified learning disabled because he scored only 14% below his grade placement [(11.Ø - 9.5)/11.Ø] x 1ØØ = 14%. Shelly, who was in the same grade (11.Ø) scored 5Ø% below the eleventh grade and thus, qualified for learning disabilities. Again, the confidence factor is less than 1ØØ%, although very close to 1ØØ%. Type 'done' which should return you to the main menu.

| Student | Deficiency Area | Grade Placement | Test Score | Advice and Confidence |
|---|---|---|---|---|
| Shelly | Calculation | 11.Ø | 5.5 | LD 95% |
| Matt | Calculation | 11.Ø | 9.5 | Minor Problem 99% |

INSTANCE 3

CINDY, WHILE IN KINDERGARTEN (Ø.9), SCORED AS A PRESCHOOLER WOULD ON A TEST WHICH MEASURES STUDENTS' ABILITIES TO EXPRESS THEMSELVES WHILE SPEAKING. IN FACT, SHE SCORED Ø.1.

Here we are given Cindy's grade placement (Ø.9) and her test score (Ø.1). However, the deficiency area is not listed as it appears in the definition. It states "a test which measures students' abilities to express themselves while speaking." This may be interpreted as 'oral expression.' Therefore, type 'o', 'Ø.9', and 'Ø.1' in response to the questions.

| Student | Deficiency Area | Grade Placement | Test Score |
|---|---|---|---|
| Cindy | Oral Expression | Ø.9 | Ø.1 |

The advice shown should have read learning disabled at 28%. Cindy's learning deficit is in an appropriate area (oral expression) discrepancy is 89% [(Ø.9 - Ø.1)/Ø.9] x 1ØØ = 89%. However, because she is a young child (kindergarten age), the system is programmed to be less certain about a diagnosis of learning disabilities. That is why the confidence factor is low.

Again, type 'done'

| Student | Deficiency Area | Grade Placement | Test Score | Advice and Confidence |
|---|---|---|---|---|
| Cindy | Oral Expression | Ø.9 | Ø.1 | LD 28% |

INSTANCE 4

RAY SCORED 0.7 ON THE CARROW ELECITED LANGUAGE INVENTORY WHILE IN THE NINTH MONTH OF KINDERGARTEN.

Ray was in the ninth month of kindergarten (0.9). The Carrow Elicited Language Inventory was the test on which he scored 0.7. You should infer that the Carrow test measures oral expression. Therefore, your responses should be 'o', '0.9', and '0.7'.

Ray's area of deficiency and grade placement are the same as Cindy's (the previous example). However, Ray scored higher (0.7) than Cindy (0.1). Cindy qualified as learning disabled. Enter the above listed responses for Ray.

| Student | Deficiency Area | Grade Placement | Test Score | Advice and Confidence |
|---|---|---|---|---|
| Cindy | Oral Expressions | 0.9 | 0.1 | LD 28% |
| Ray | Oral Expressions | 0.9 | 0.7 | |

The conclusion should indicate that a learning disabled classification is not appropriate for Ray. As with Instance 3, the confidence factor relating to this conclusion is only 28% because the student is very young (in kindergarten).

| Student | Deficiency Area | Grade Placement | Test Score | Advice and Confidence |
|---|---|---|---|---|
| Cindy | Oral Expressions | Ø.9 | Ø.1 | LD 28% |
| Ray | Oral Expressions | Ø.9 | Ø.7 | Minor Problem 28% |

# LESSON 1.A

## PRACTICE

PRACTICE 1

CHRIS IS A TEN-YEAR-OLD IN FIFTH (5.Ø) GRADE WHO IS HAVING PROBLEMS READING. HIS TEACHER GAVE HIM THE WOODOCK READING TEST AND DISCOVERED HE IS FUNCTIONING AT 2.Ø.

After reading this item determine on your own whether Chris qualifies for a learning disabilities classification.

* * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

Does Chris qualify for learning disabilities classification?

____ yes ____ no

On a scale of 1 to 1ØØ, about how confident are you that Chris can be classified learning disabled? __________

Why does or doesn't he qualify?

** * * * * * * * * * * * * * * * * * * * * * * * * * * * *

Now enter the appropriate responses and compare the computer's advice to your answer.

Chris does qualify as learning disabled. The advice shown should indicate confidence at the 95% level. Chris is functioning 6Ø% below his grade placement in an appropriate area:

$$[(5.\text{Ø} - 2.\text{Ø})/5.\text{Ø}] \times 1\text{ØØ} = 6\text{Ø}\%$$

* * * * * * * * * * * * * * * * * * * * * * * * * * * * *

Suppose Chris had scored at the 2.Ø grade level in Science, now would he have qualified for a learning disabilities classification?

______ Yes ______ No

How certain are you that Chris now qualifies as learning disabled (on a scale from 1 to 1ØØ)? _________

Why would or wouldn't Chris qualify?

* * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

Now respond to the questions on the computer.

Now start Practice 1. Input grade placement and test score the same as before (ie., 5.Ø and 2.Ø), but now enter his learning problem area as Science.

The advice shown should indicate that Chris' problem is not in an appropriate area and thus, a learning disabled classification is not possible.

Listed in the box below are the data inputed for Chris and the advice and confidence factors obtained from the computer. Write what your advice and confidence levels were in the appropriate boxes. Then list the computer's advice and confidence for the second example.

When you have completed the above, type 'done'.

| Deficiency Area | Grade Placement | Test Score | LD or Not LD Confidence | Computers Advice and Confidence |
|---|---|---|---|---|
| Reading | 5.Ø | 2.Ø | | Learning Disability 95% |
| Science | 5.Ø | 2.Ø | | |

PRACTICE 2

ANDY, A PHYSICALLY HANDICAPPED EIGHT-YEAR-OLD (1Ø4 MONTHS), SECOND GRADER (2.2) SCORED 1Ø5 ON THE STANFORD-BINET. HE IS HAVING DIFFICULTY CONCEPTUALIZING STORY PROBLEMS. HIS TEACHER ADMINISTERED A TEST AND DISCOVERED HE WAS PERFORMING AT THE Ø.6 LEVEL.

* * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

First, does Andy qualify as learning disabled?

____ yes ____ no

How certain are you that Andy can be classified learning disabled (from 1 to 1ØØ)?_______

Why does or doesn't he qualify?

* * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

Now, respond to all the questions on the computer. How does the advice shown compare with your response?

Andy is having difficulty "conceptualizing story problems." This should have been interpreted as a mathematical reasoning problem. The computation for deficit is: [(2.2 - Ø.6)/2.2] x 1ØØ = 72.7%. Therefore, Andy may qualify for a learning disabilities placement.

The computer's conclusion you should have received was "learning disabled (95%)."

* * * * * * * * * * * * * * * * * * * * * * * * * * * * *

Suppose that Andy had scored 1.6 on a math reasoning test. Now would he qualify for learning disabilities?

______ Yes ______ No

How confident are you that Andy can now be classified learning disabled? __________

Why or why not does Andy now qualify?

* * * * * * * * * * * * * * * * * * * * * * * * * * * * *

Input the responses on the computer.

Now Andy does not qualify for a learning disabilities classification because he did not perform 40% below grade placement:

$$[(2.2-1.6)/2.2] = 27\%$$

Below is a chart with the data inputed for Andy. Write your responses and the computer's advice in the appropriate boxes.

| Deficiency Area | Grade Placement | Test Score | LD or Not LD Confidence | Computers Advice and Confidence |
|---|---|---|---|---|
| Math Reasoning | 2.2 | 0.6 | | Learning Disability 95% |
| Math Reasoning | 2.2 | 1.6 | | |

# Appendix C

## A Portion of the Instance Difficulty Level Test

NAME (optional)______________________________DATE______________

DIRECTIONS: Please read all the enclosed information and complete the exercises as they are presented. There are six parts to this packet. At the beginning of each part, you will be given a definition. Then you will be given instances representing examples and nonexamples of the definition. It is important that you first study the definition. Then proceed by reading each instance and determine whether it is an example or nonexample. Be certain that you refer only to the specific definition you are on, not to previous definitions. As you are working, if you have any questions, please raise your hand.

Definition Number One

Definition: A learning disabled student must score **at least 40% below grade placement** in at least one of the following academic areas:

**basic reading skills**

**reading comprehension**

**calculations**

**mathematical reasoning**

**written expression**

**listening comprehension**

**oral language.**

Instances:

1. Melanie has difficulty in her science class. When her teacher tested her she found out Melanie is 40% below her grade placement.

_____Learning Disabled _____Not Learning Disabled

2. Zach is thirteen years old. His English teacher noticed he was having difficulty with sentence structure in his compositions. When his teacher gave him the Woodcock-Johnson Test she discovered he was functioning at the third grade lvel (63% below) and his IQ score was 96.

_____Learning Disabled _____Not Learning Disabled

3. Joe was in the ninth grade (9.0) when he scored at the fourth grade, fifth month (4.5) in reading comprehension. His score indicates he is functioning 50% below his grade placement.

_____Learning Disabled _____Not Learning Disabled

4. Bobby is in second grade, second month (2.2) and scored first grade, sixth month (1.6) on the Key Math. His IQ is 105 so he scored 52% below what he was expected to score and 27% below his grade.

_____Learning Disabled _____Not Learning Disabled

5. Jason was given a test of handwriting and scored at the third grade level (3.0). He was in the seventh grade, sixth month (7.6) at the time. This means he scored 61% below his grade placement.

_____Learning Disabled _____Not Learning Disabled

6. Chris is a ten-year-old in fifth grade who is having problems reading. His teacher gave him the Woodcock Reading Test and discovered he is functioning 56% below fifth grade.

_____Learning Disabled _____Not Learning Disabled

7. Cindy, while in kindergarten, scored as a preschooler would on a test which measures students' abilities to express themselves orally. In fact, she scored 70% below her grade placement on this test.

_____Learning Disabled _____Not Learning Disabled

8. Shelly scored 50% below her grade placement (eleventh grade) in mathematical calculation.

_____Learning Disabled _____Not Learning Disabled

9. Ray scored 22% below his grade placement on the Carrow test of oral language.

_____Learning Disabled _____Not Learning Disabled

10. Jill is failing her art class. Her teacher claims she is performing at least 40% behind her class.

_____Learning Disabled _____Not Learning Disabled

11. Matt scored at the 9.5 grade level in calculation when he was in the eleventh grade (11.0). This means he scored 14% below his grade placement.

_____Learning Disabled _____Not Learning Disabled

12. Andy, a physically handicapped, eight-year-old, second-grader, scored 105 on the Stanford-Binet IQ Test. He is having difficulty reasoning his arithmetic story problems. His teacher administered the Key Math Test and discovered he is performing 70% below second grade.

_____Learning Disabled _____Not Learning Disabled

# Appendix D

## Process for Developing the Instructional Materials

PROCESS FOR DEVELOPING THE INSTRUCTIONAL MATERIALS

LD.TRAINER

1. Break lessons into conceptual subcomponents.

2. Define the subcomponent concepts. Content specialists evaluate the definitions.

3. Generate an instance pool by matching examples and nonexamples on variable attributes and varying the critical attributes.

4. Estimate the difficulty level for each instance to ensure a range. Content specialists evaluate the instances and students identify instances as examples or nonexamples from the definition only.

5. Develop instructional materials. For each lesson subcomponent present:

    a. Definition

    b. Instances - matched examples/nonexamples
       easy to difficult
       attribute isolation

    c. Practice - random presentation
       attribute feedback

6. Develop a diagnostic classification test.

# Appendix E

## Lesson Rating Form

LD.TRAINER

LESSON RATING FORMS

LESSON # AND NAME________________________________

DIRECTIONS: Please read each statement below and circle the number corresponding with your feelings about the lesson just completed. Feel free to list any comments you may have.

1. Instructions

The instructions for this lesson were clear and understandable.

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| strongly disagree | disagree | neutral | agree | strongly agree |

Comments:

2. Definitions

The definition provided was clear, understandable, and correct.

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| strongly disagree | disagree | neutral | agree | strongly agree |

Comments:

3. Number of Instances and Practice Items

There was an appropriate number of instances and practice items provided to learn the concept.

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| strongly disagree | disagree | neutral | agree | strongly agree |

Comments:

4. Written Materials

The written materials in general were easy to follow.

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| strongly disagree | disagree | neutral | agree | strongly agree |

Comments:

5. Computer

Using the computer was a better way than lecture to learn this material.

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| strongly disagree | disagree | neutral | agree | strongly agree |

Comments:

# Appendix F

## Pre and Posttest

# CLASSIFICATION OF LD STUDENTS

## TEST

NAME_________________________________________ NUMBER____________

DATE__________________________

Directions: For each of the following items, read the case study and determine whether or not the student qualifies for a learning disabilities classification. A few of the items contain an additional question. For these items determine whether the test data was appropriate or whether a special education placement is warranted. Check the appropriate line. Then justify your response in the space provided.

1. Dennis is an eleven-year-old boy (142 months old) with muscular dystrophy. He is in sixth grade, zero month (6.0). Dennis' condition requires that he attend a special class so that he may obtain occupational and physical therapy, along with individualized academic help. In particular Dennis is having difficulty with written expression. He scored 83% below his expected and 84% below his actual grade placement on a test in this area. Also, he scored 96 on the WISC-R. Dennis exhibits no unusual behavioral problems although his communication difficulties are becoming more and more apparent and may account, in part, for his learning problems. Educational and social history information indicate no cultural, economic, or environmental causes of his learning difficulties.

_____Learning Disabled     _____Not Learning Disabled

Why can or cannot Dennis be learning disabled?

2. Brad is in fourth grade, ninth month (4.9) and ten years, two months old (122 months). On a test of math calculation Brad scored at the first grade, eighth month (1.8) level and on the Slosson IQ Test he scored 1Ø7. Review of Brad's health records indicates no sensory, physical, health, or communication problems. Although Brad has had behavior difficulties in the past, the multidisciplinary team reviewing Brad's assessment data and other records thinks that behavior problems are not the primary cause of his learning difficulties. The school psychologist visited with Brad's parents in their home and reported that there appeared no cultural or economic causes to Brad's learning problems. The psychologist did learn, however, that Brad's natural parents are divorced and that Brad's mother, whom he lives with, just remarried. His mother and step-father reported that this new adjustment has been very difficult on Brad. The team concluded that there is a possibility, although unlikely, that Brad's learning deficit may be improved if he were to be given more individualized help in his regular classroom.

_____Learning Disabled _____Not Learning Disabled

Why can or cannot Brad be learning disabled?

_____Special Education Appropriate _____Special Education Not Appropriate

Why is or isn't special education appropriate for Brad?

3. Danny has had a lot of academic difficulty since he entered school. He is in second grade, first month (2.1) and scored on the first grade, zero month (1.Ø) level in basic reading skills on the Woodcock Reading Mastery Test. Danny is eight years, zero months old (96 months). His IQ was reported as 98. (The Stanford-Binet was administered.) Danny was screened for both vision and hearing. He scored 2Ø/3Ø on the vision acuity test and had a hearing loss of 12 decibels in his right ear. Danny exhibits no other physical or health problems. He communicates and interacts with his peers well. Nothing in his social and educational history record indicates any cultural, environmental, or economic problems.

_____Learning Disabled _____Not Learning Disabled

Why does or doesn't he qualify?

4. Layne is in the fifth grade, zero month (5.0) and is eleven years, four months old (136 months). He was referred to special education for difficulties with oral expression. He was administered the WISC-R and the Test of Language Development and scored 81 and 1.1, respectively. Based on interviews and home visits with the parents, there appear no cultural, environmental, or economic difficulties. Layne scored 20/20 on the visual acuity test and did not exhibit any loss on the hearing test. Layne does have allergy problems which do not appear to influence his learning. No behavior or communication difficulties have been noted.

_____Learning Disabled _____Not Learning Disabled

Why does or doesn't he qualify?

5. Wayne is a tenth grader (10.2) and sixteen years old (193 months). He is having written expression difficulties. A review of his past history indicates disadvantaged home conditions which are probably the cause of his learning difficulties. Wayne scored 62% below his actual and 60% below his expected achievement on the Test of Written Language and 93 on the WAIS. There appear to be no sensory, physical, health, behavior, or communication difficulties.

_____Learning Disabled _____Not Learning Disabled

Why does or doesn't Wayne qualify?

6. Tony is a first grader (1.5) who was referred to special education because of basic reading and reading comprehension difficulties. He scored Ø.1 on the reading test and 1Ø1 on the Slosson. Tony is seven years, five months old (89 months). His hearing test indicated no hearing loss and he scored 2Ø/9Ø with correction on the test of visual acuity. Both tests were conducted by personnel in his school district. Tony hasn't been a behavior problem in the past and appears to communicate with others effectively. Having interviewed Tony's parents, the multidisciplinary team concluded there were no cultural, economic, or environmental reasons for his learning deficit.

_____Learning Disabled _____Not Learning Disabled

Why does or doesn't he qualify?

7. Natalie was having difficulty with listening comprehension so her teacher referred her to the resource program for additional testing. Natalie, at the time, was in grade 3.7 and nine years old (108 months). She scored 68% below her actual and 70% below her expected grade placement in listening comprehension and 114 on the Stanford-Binet. Based on Natalie's educational and social history, there were no cultural, environmental, or economic reasons for her learning deficit. Natalie's classroom teacher completed a behavior rating scale and the psychologist and special educator observed Natalie in the classroom and on the playground. Based on these observations and the rating scale, the team decided that Natalie's primary problem was her inappropriate behavior. To rule out any sensory problem, Natalie's vision and hearing were also tested by the school nurse and the audiologist. Her visual acuity was 20/40 in her better eye. The audiologist interpreted Natalie's hearing results as indicating no hearing problem, although the actual scores were not recorded.

_____Learning Disabled     _____Not Learning Disabled

Why does or doesn't Natalie qualify as learning disabled?

_____Appropriate Data     _____Inappropriate Data

Why is the data appropriate or inappropriate?

8. Trevor is having difficulty with spelling. His teacher referred him to the resource program for possible help. He scored at the third grade, second month (3.2) level in spelling and is in the eighth grade, fourth month (8.4). Trevor also scored 1ØØ on the WISC-R IQ Test. He was thirteen years, eleven months old at the time. There appear no physical, health, sensory, behavioral, economic, environmental, or cultural problems. Trevor is presently in speech therapy for an articulation problem.

_____Learning Disabled _____Not Learning Disabled

Why can or can't Trevor be classified learning disabled?

9. Emily is a seventh grader (7.4) who has math calculation difficulties. She scored on the third grade, second month level (3.2) in calculation and 95 on the Slosson Intelligence Test. She was thirteen years old (156 months) at the time of testing. Emily's hearing and vision screening tests indicate no sensory problems and her health record shows no educationally related physical or health disabilities. Emily does have a communication problem, but the team reviewing her records thinks her communication problem is secondary to her learning problem. Emily's teacher reports that Emily is very well behaved in the classroom. Emily comes from a high-middle class home which appears very conducive to learning.

_____Learning Disabled _____Not Learning Disabled

Why can or can't Emily be classified learning disabled?

10. Holly is a fourteen-year-old (175 months) ninth grader (9.2). She scored 121 on the WISC-R and 6.0 on an appropriate test of basic reading. The teacher who referred Holly requested that her vision be tested. Holly scored within the normal range of vision. There appeared no other sensory, physical, or health problems. Holly's teachers report no behavioral or communicative difficulties. Holly's parents have visited with the principal and her teachers several times because of their concern regarding Holly's reading difficulties. Based on these conversations with her parents, there appears to be no cultural, environmental, or economic problems contributing to Holly's learning difficulties.

_____Learning Disabled _____Not Learning Disabled

Why does or doesn't Holly qualify as learning disabled?

11. Teresa was in second grade (2.Ø) when she was referred for math reasoning difficulties. At that time she was seven years, six months old (7.6). Teresa scored 1.2 on the Key Math and 9Ø on the Stanford-Binet. Although Teresa suffers from asthma, she has not missed any school for health reasons. Teresa's hearing and vision screening tests were scored in the normal range. Also, possible environmental, cultural, economic, behavioral, or communicative problems do not exist.

_____Learning Disabled _____Not Learning Disabled

Why can or can't Teresa qualify as learning disabled?

12. Celeste was in second grade, second month (2.2) when she scored Ø.6 on a test of written expression. She scored 1ØØ on the Slosson Intelligence Test and was seven years, eleven months old (95 months) at the time. Celeste was born in France and has lived in the United States approximately four years. Celeste speaks both French and English fluently for her age. In fact, although French is spoken in the home, Celeste's English was rated better than her classmates' English by her teacher. The school that Celeste attends is comprised of about 1Ø% minorities. There appear no economic or environmental reasons for Celeste's learning difficulty. Also, her health record indicates no sensory, physical, or health problems. Celeste's teacher reports that Celeste exhibits no behavioral or communicative problems.

_____Learning Disabled _____Not Learning Disabled

Why can or can't Celeste be classified learning disabled?

# Appendix G

## Demographic Questionnaire

# DEMOGRAPHIC INFORMATION

Name (optional)__________________________Number_____________

1. Please check your class standing:

| | | |
|---|---|---|
| _____Undergraduate | _____Graduate | _____Not a degree seeking student |
| _____Freshman | _____Masters | |
| _____Sophomore | _____Doctorate | |
| _____Junior | | |
| _____Senior | | |

2. Major subject__________________________

   Minor subject__________________________

3. Last degree obtained____________________________

4. Please check CERTIFICATION(S) obtained and list the state(s):

| | State(s): |
|---|---|
| _____Elementary education | _____________________________ |
| _____Secondary education<br>Subject(s)__________<br>___________________ | _____________________________<br>_____________________________ |
| _____Special education<br>Type(s)____________<br>___________________ | _____________________________<br>_____________________________ |
| _____Administrative | _____________________________ |

5. Please list the number of YEARS of TEACHING experience:

   _____Regular education/Elementary school

   _____Regular education/Secondary school

   _____Special education/Elementary school

   _____Special education/Secondary school

   _____Special education/Special school

   _____Other (please specify)____________________________

6. Please list the number of YEARS of ADMINISTRATIVE experience:

_____Regular education/Elementary school

_____Regular education/Secondary school

_____Special education/Special school

_____Regular education/District

_____Special education/District

_____State

_____Other (please specify)______________________________

7. Please list the state, district, and school in which you taught or had administrative duties during the following years:

| | School | District | State |
|---|---|---|---|
| 1985-86 | | | |
| 1986-87 | | | |

## Appendix H

## Instructions for LD.Trainer

INSTRUCTIONS

LD.TRAINER

Insert LD.TRAINER Disk 1 in Drive 1.

Insert LD.TRAINER Disk 2 in Drive 2.

Turn on the computer.

Main menu - Type the number corresponding to the lesson <R>.

Lesson menu - Type the number corresponding to the instance or practice item <R>.

If you make an error after <R>: > quit <R>

Read material

*Type responses and <R>

If you make an error after <R>: > abort <R>
M.1> go <R>

You will have to restart that item.

After advice shown: > done

* The auto-completion feature allows you to type the first few letters of your responses only.

# Appendix I

## Commands for CLASS.LD2 Group

CLASS.LD2

COMMANDS

Any of these commands can be typed after a question has been asked by the computer:

| TYPE | FUNCTION |
|---|---|
| 'WHY' | An explanation will be given as to why the question was asked. |
| 'SHOW' | You will see all of the values that the memory has up to this point in the consultation. |
| 'RULE - (#)' | If a rule and number appear after you type 'show,' you may view the actual rule by typing 'rule-' followed by the rule number. |

# Appendix J

## Instructions for CLASS.LD2 Group

# INSTRUCTIONS

## CLASS.LD2

Insert CLASS.LD2 Disk 1 in Drive 1

Insert CLASS.LD2 Disk 2 in Drive 2

Turn on the computer

```
A> return <R>
A> <R>
A> ld

M.l> colors off
M.l> go
```

Write down responses.

*Type responses and <R>.

If you make an error after <R>:

```
> abort <R>
M.l> go
```

You will have to restart the consultation

After advice shown:

M.l> savecache a:file___ (# of the file) DO NOT RETURN

Remove CLASS.LD2 DISK 1 from drive a

Insert CLASS.LD2 FILES disk in drive a, then <R>

Wait for M.l prompt M.l>

Remove CLASS.LD2 FILES disk from drive a

Insert CLASS.LD2 DISK 1 in drive a

M.l> exit <R>

```
A> ld <R>

M.l> colors off <R>
M.l> go <R>
```

* The auto-completion features allows you to type the first few letters of your responses only.

# Appendix K

## LD.Trainer Record Sheet

LD.TRAINER

ASSIGNMENT SHEET AND TIME RECORD

STUDENT'S NAME______________________________ NUMBER____________

| LESSON # AND NAME | DATE WORKED ON | TIME STARTED | TIME COMPLETED |
|---|---|---|---|
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

# Appendix L

## CLASS.LD2 Record Sheet

CLASS.LD2 FILES

RECORD SHEET

Name ______________________________ Set______ Number_________

| File # | Date Entered | Team Classification | Computer Advice and Confidence | Time Started | Time Ended |
|---|---|---|---|---|---|
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |
| | | | | | |

# VITA

**Mary Anne Prater**
**Department of Special Education**
**Utah State University**
**(801) 753-7973**
**Logan, Utah 84322**

## Education

Bachelor of Music, 1975, University of Utah.
Graduated Magna Cum Laude
Major: Music Education Minor: Math Education

M. S., 1982, University of Utah.
Major: Special Education

Attended, 1983-84, Arizona State University.
Major: Special Education

Ph.D., 1987, Utah State University.
Major: Special Education Minors: Instructional Tech.
Educational Admin.

## Professional Affiliations

American Association for Artificial Intelligence (AAAI)
American Educational Research Association (AERA)
Association for Behavior Analysis (ABA)
Council for Exceptional Children (CEC)
Behavior Disorders Division
Learning Disabilities Division
Mental Retardation Division
Teacher Education Division

## Professional Experiences

Research Associate. Behavior Consultant, Learning Disabilities, and Multidisciplinary Expert Systems Training Projects, 1985-present. Intelligence Research and Development Unit, Developmental Center for Handicapped Persons, Utah State University, Logan.

Research Assistant. Curriculum Monitoring and Instructional Decision-Making Micro-Computer Project, 1984-1985. Department of Special Education, Utah State University, Logan.

Special Education Supervisor. 1984, Logan School District, Logan, UT.

Undergraduate Practicum Supervisor. 1984, Department of Special Education, Utah State University, Logan.

Instructor. 1984, Department of Special Education, Arizona State University, Tempe.

Special Educator and Special Education Team Leader. 1980-1983, Jordan School District, Sandy, UT.

Parent Training Coordinator. P.L. 94-142 and IEP Projects. 1979-1980, Utah Association for Retarded Citizens. Salt Lake City.

**Publications**

Prater, M. A. (1987) *Expert system technology and concept instruction: Training educators to accurately classify learning disabled students*. Unpublished doctoral dissertation, Utah State University, Logan.

Prater, M. A. (1986). *Effective concept instruction: A procedure for the design of instructional materials*. Manuscript submitted for publication.

Prater, M. A. (1986). *Reliability estimates of criterion and domain-referenced tests*. Manuscript submitted for publication.

Prater, M. A. (1986). *Enhancing reading performance of mildly handicapped students*. Manuscript submitted for publication.

Prater, M. A. (1986). *A comparison of three early intervention indices*. Manuscript submitted for publication.

Lubke, M. M., & Prater, M. A. (1986). *Expert systems: Implications for special education administrators*. Manuscript submitted for publication.

Ferrara, J., Prater, M. A., & Baer, R. (in press). LD.Trainer: Modification of an expert system for complex conceptual training. *Educational Technology*.

Prater, M. A. (1985). Data decisions made by special and regular educators using performance, progress and charted data. In R. P. West & K. R. Young (Eds.), *Precision teaching: Instructional decision making, curriculum and management, and research* (pp. 160-171). Logan, UT: epartment of Special Education, Utah State University.

Johnson, J., Prater, M. A., West, R., Young, R., & Larsen, R. (1985). Precision teaching concepts: A brief review. Logan, UT: Developmental Center for Handicapped Persons and the Department of Special Education, Utah State University.

Scruggs, T. E., Mastropieri, M. A., Levin, J. R., McLoone, B., Gaffney, J. S., & Prater, M. A. (1985). Increasing content-area learning: A comparison of mnemonic and visual-spatial direct instruction. Learning Disabilities Research, 1, 18-31.

Prater, M. A. (1982). Parent participation in the IEP process across the state of Utah. Unpublished master's thesis, University of Utah, Salt Lake City.

**Grants Funded**

Prater, M. A., & Baer, R. (P.I.) (1985). An Artificial Intelligence-Based Behavior Consultant Training Program: Inservice for Regular Educators Serving Handicapped Students. U.S. Department of Education--Office of Special Education and Rehabilitative Services. Utah State University, Logan.

Miller, G. H., & Prater, M. A. (1979) A Training Program for Parents of School-aged Handicapped Students Regarding P.L. 94-142 and the IEP Process. Utah State Department of Education. Utah Association for Retarded Citizens, Salt Lake City.

**Products**

Prater, M. A., & Althouse, B. (1986). LD.Trainer. [Computer program and accompanying printed materials]. An expert system designed to train educators and related school personnel to accurately identify learning disabled students. Developmental Center for Handicapped Persons, Utah State University, Logan.

**Presentations**

Prater, M. A. (1987, April). The modification of a computer-based expert system for training special educators to accurately classify learning disabled students. Paper accepted for presentation at the Annual Meeting of the Council for Exceptional Children, Chicago.

Prater, M. A., Serna, R. W., & Hemphill, H. (1986, June). Expert systems in the assessment of handicapped students. Workshop presented at the Annual Intervention Procedures for Exceptional Children, Utah State University, Logan.

Prater, M. A., & Ferrara, J. M. (1986, May). Training applications of expert systems. Paper presented at the Annual Meeting of the Association for Behavior Analysis, Milwaukee.

Prater, M. A. (1985, December). Intelligent systems as 'if-then' templates in training situations. Paper presented at the Annual High Technology in Higher Education Meeting, Utah State University, Logan.

Prater, M. A. (1985, December). LD.Trainer: An expert system application in special education. Paper presented at Technology Information Exchange, Regional Resource Centers, Utah State University, Logan.

Prater, M. A. (1985, August). CLASS.LD: An expert system. Paper presented at the International Joint Conference on on Artificial Intelligence, Los Angeles.

Prater, M. A. & West, R. (1985, May). Training and supervision of teacher decision-making. Paper presented at the National Precision Teaching Conference, Seattle.

Prater, M. A. (1985, April). The keyword method and the keyword vocabulary method: Increasing learning and memory skills of learning disabled students. Paper presented at the National Meeting of the Council for Exceptional Children, Los Angeles.

Scruggs, T. E., Mastropieri, M. A., Leven, J. R., McLoone, B., Gaffney, J. S. & Prater, M. A. (1985, April). Increasing content-area learning: A comparison of mnemonic and visual-spatial direct instruction. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago.

Prater, M. A. (1984, April). Data decisions based upon anecdotal and graphed data. Paper presented at the Annual National Precision Teaching Conference, Park City, UT.

## Special Skills

Statistical analysis and research methodology skills - Served as tutor and advisor to faculty and doctoral students in research methodology and statistical analysis of data. User and trainer of SYSTAT and STATISTIX, two micro-computer statistical software packages.

Project management software skills - User of SuperProject and SuperCalc3, micro-computer software programs for developing PERT charts, spreadsheets, and other management tools.

Editorial advising and writing - Extensive experience in editing and writing publications and technical reports.

Completed knowledge engineering training - Teknowledge, Palo Alto, CA, January 1986.